IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 26, 2019, accepted September 10, 2019, date of publication September 13, 2019, date of current version October 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2941397

Bidimensional Empirical Mode Decomposition for SAR Image Feature Extraction With Application to Target Recognition

MING CHANG[1], XUQUN YOU[1], AND ZHENGYANG CAO[2]
[1]Shaanxi Provincial Key Laboratory of Behavioral and Cognitive Neuroscience, School of Psychology, Shaanxi Normal University, Xi'an 710062, China
[2]Xi'an ASN UAV Technology Company Ltd., Xi'an 710065, China

Corresponding author: Xuqun You (youxuqun_snnu@163.com)

This work was supported by Humanity and Social Sciences Construction Special Project of Ministry of Education (18JDGC024), the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 31500897), the Fundamental Research Funds for the Central Universities of Ministry of Education of China (Grant No. GK201603123).

ABSTRACT This study introduces a target recognition algorithm for synthetic aperture radar (SAR) images based on the features extracted by bidimensional empirical mode decomposition (BEMD). BEMD provides an adaptive and empirical way to process signals, which generates bidimensional intrinsic mode functions (BIMFs) to describe the details of SAR images. Therefore, the generated BIMFs are complementary to the original image and their joint use could probably improve the recognition performance. In order to fully exploit the discrimination of these components, the joint sparse representation (JSR) is employed during the classification. JSR operates as multi-task learning algorithm, which represents each component numerically while considering their inner correlations. The original image together with the generated BIMFs are simultaneously represented by JSR to determine the target label according to the output reconstruction errors. Experimental results on the Moving and Stationary Target Acquisition and Recognition (MSTAR) data set demonstrate the validity of the proposed method under different operating conditions. In comparison with some baseline algorithms, the superiority of the proposed method is furtherly validated.

INDEX TERMS Synthetic aperture radar (SAR), target recognition, bidimensional empirical mode decomposition (BEMD), joint sparse representation (JSR).

I. INTRODUCTION

As an active sensor transmitting and receiving microwave, synthetic aperture radar (SAR) is able to provide high-resolution observations of the focused ground scene for military or civilian applications. Because of the specific imaging mechanism, it is much more difficult to interpretate SAR images than their optical counterparts. Automatic target recognition (ATR) is a computer-aided decision system, which automatically locates the potential targets and determine their labels in an acquired SAR image [1]. A baseline SAR ATR system usually involves three key steps, i.e., detection, discrimination, and classification [2]. Target detection locates the positions of the potential targets in a large-size SAR image and outputs the regions of interest (ROI). Afterwards, these ROIs are further selected via target discrimination to eliminate false alarms brought by natural clutters like trees, rivers, etc. Finally, the remaining ROIs are classified in the classification stage to determine their target labels. The Semi-Automated Image Intelligence Processing (SAIP) [2] and Moving and Stationary Target Acquisition and Recognition (MSTAR) programs [2] are two typical SAR ATR systems, which perform target classification in the template-based and model-based ways, respectively. For a concrete classification algorithm, it usually consists of feature extraction and classifier. During feature extraction, the original SAR image is transformed into a stable subspace, which differs itself from other classes. So far, many types of features have been employed or designed for SAR ATR including geometrical [4]–[8], transformation [9]–[15], and electromagnetic features [16]–[19]. For the geometrical features, they were used to depict the physical sizes or shape of the target, e.g., target outline [4] and binary target region [5], [6]. These features contain clear physical meaning but are hard to

The associate editor coordinating the review of this manuscript and approving it for publication was Geng-Ming Jiang.

be extracted with high precision because of the speckle noises in SAR images. The transformation features are extracted via the matrix projection or signal processing techniques. The manifold learning methods including principal component analysis (PCA) [9], [10] and non-negative matrix factorization (NMF) [11] were used to project SAR images into low-dimensional feature space. Some signal processing methods such as wavelet analysis [14] and monogenic signal [15] were also demonstrated useful for SAR image feature extraction. Unlike optical images, SAR collects the backscattering field of the target. Therefore, the electromagnetic features can be used for the discrimination of different targets such as the scattering centers, polarization, etc. Several classification schemes were developed for the attributed scattering centers for SAR ATR in the previous works [16]–[19]. The classifiers operate on the extracted features to make decisions on the target labels. At present stage, most classifiers in SAR ATR were inherited from the traditional pattern recognition fields like face recognition, fingerprint indexing, etc. The advanced classifiers including support vector machines (SVM) [4], [5], [20]–[22], sparse representation-based classification (SRC) [23]–[26] were demonstrated effective for SAR ATR. Zhao and Principe first applied SVM to SAR target recognition and validated its superior performance over traditional template-based methods [20]. Afterwards, SVM became a very popular classifier in the field of SAR ATR [4], [5], [21]. Thiagarajan et al. introduced SRC to SAR target recognition and demonstrated its effectiveness [23]. After then, SRC was widely employed to classify different kinds of features to improve the recognition performance [15], [24]–[26]. Other classifiers like adaptive boosting (AdaBoost) [27], modified polar mapping [8], and discriminative graphical model [28] were also validated effective in SAR target recognition. The recently popular deep learning methods also find their application in SAR ATR, which combine feature learning and classification in the same framework. With excellent classification capability, deep learning models turn to be the mainstream in the field of SAR ATR at present with a rich set of novel methods [29]–[45]. These methods mainly differ in the network structures and preprocessing algorithms of the training samples. The embryo of deep learning-based SAR target recognition methods was the one proposed in [29] using the convolutional neural network (CNN). Later, Chen et al. developed the famous all-convolutional networks (A-ConvNet) and significantly improved the overall recognition performance [30]. Other architectures of deep networks were designed recently with good performance using the latest achievements in deep learning, such as cascade coupled CNN [31], gradually distilled CNN [32], enhanced squeeze and excitation network (ESENet) [33], multi-stream CNN [34]. Despite the innovations in network structures, some researchers tried to enrich the available training samples to enhance the classification capability of the trained networks. In [35], Ding et al. augmented the available training samples by image translation and noise addition. In [36], Yan simulated noisy, multi-resolution, and occluded samples to train the designed CNN. In [37], more SAR images of targets of interests were simulated based on their CAD models, which were used to train CNN via transfer learning. In addition, it is also a feasible way to combine deep learning models with traditional classifiers. Wagner combined CNN and SVM for SAR ATR [22]. In [38], CNN worked cooperatively with sparse coding. Cui et al. updated CNN with an assistant classifier, i.e., SVM [39]. In [40], a decision fusion strategy was developed by combination of CNN and scattering center matching in a hierarchal way. Although with good performance in current literatures, it should be noted that deep learning models have defections in the field of SAR ATR. In fact, the practical situations during SAR data acquisition is very complex and there exist many extended operating conditions (EOCs). These EOCs may be caused by the variations of the target (e.g., configuration variance), background environment (e.g., occlusion), and sensors (e.g., noise corruption) [3]. As a result, the deep learning models trained by limited and fixed training samples could hardly handle these nuisances. In this case, it is desired that some novel features can be developed or employed for SAR ATR in order to improve the overall performance.

In this study, we apply bidimensional empirical mode decomposition (BEMD) to SAR ATR. EMD [46] was proposed by Huang et al., which provides an effective tool for adaptive and multi-scale analysis of non-stationary signals. As a generalization to 2D signals, e.g., images, BEMD was proposed and applied to image segmentation, fusion, compression, etc. [47]–[53]. It is demonstrated that the features generated by BEMD, i.e., bidimensional intrinsic mode functions (BIMFs), can better hold the details of the original image such as the contour and local structures. So, they can help interpretate the image for different applications, e.g., object reconstruction, image denoising, and image enhancement. This study introduces BEMD into SAR image feature extraction and target recognition. The BIMFs can capture broader spectral information of the target, which provide complementary information for the original SAR image. Hence, it is assumed that the joint use of the original image and its BIMFs would help improve the recognition performance. Accordingly, the motivations of applying BEMD into SAR image feature extraction and target recognition can be summarized as following points. First, BEMD was demonstrated to be a powerful tool for image processing, such as denoising, enhancement, etc. Therefore, it can be smoothly applied to SAR image feature extraction with stable performance. Second, SAR images contain rich spectral information, which could not be fully exploited by a simple feature extraction algorithm. It is preferred that the multi-level and hierarchical features can be extracted to represent the properties of the targets in SAR images. BEMD provide an available tool for this job, which produces several different BIMFs jointly describing the image. Third, the generated BIMFs from BEMD are actually correlated because they describe the same project. Therefore, these BIMFs not only complement each other but also share inner correlations. In the sense, more discriminative information can be used for separating

different kinds of targets. In the classification stage, the joint sparse representation (JSR) [15], [54]–[56] is used to consider the original SAR image and its BIMFs in a unified framework. JSR is the extension of SRC to multi-task problems, which is able to consider the correlations between different components. In the proposed method, the original SAR image and its BIMFs actually describe the same target so they indeed share some correlations. Therefore, JSR is a suitable classifier to perform the classification job. The main contributions in this study can be summarized as follows: (1) BEMD is introduced to SAR image feature extraction. The extracted BIMFs could complement the original image thus providing more discrimination for the correct target recognition. (1) JSR is employed to jointly classify the original SAR image and its BIMFs. By exploiting the inner correlations of different components, the decision from JSR is more precise so the overall ATR performance can be enhanced.

The remainder of this study is organized as followings. In Section 2, the basic theory of BEMD is reviewed and applied to feature extraction. Section 3 describes the classification scheme of JSR with application to target recognition. The proposed method is evaluated on the MSTAR data set in Section 4 under several operating conditions. Section 5 makes some discussions on the experimental results. Conclusions are drawn in Section 6 to summarize this study.

II. BASIC THEORY OF BEMD

Huang et al. first proposed EMD to adaptively analyze the non-stationary signals [46]. Unlike traditional signal decomposition methods, e.g., wavelet analysis, EMD does not impose any prior assumptions on the data, such as linearity and stationarity. In the past researches, EMD has been numerically validated to be more capable of describing patterns in non-stationary and nonlinear signals.

As a natural generalization of EMD to 2D space, BEMD is capable of describing an image using several BIMFs [47]–[53]. The original image is decomposed into high and low frequency components with some residues. Hence, the generated BIMFs could better reflect the global and detailed information of the decomposed image, which motivates us to extend its use into target recognition in SAR images. The detailed process of sifting the BIMFs of a given image $I(x, y)$ can be summarized as the following steps:

Step 1: Identify the locations of the extrema (local maxima or minima) in the input image $h_{10} = I(x, y)$;

Step 2: Generate the upper and lower 2D envelopes via 2D interpolation based on the maxima and minima point sets, which are denoted as $e_{\text{upper}}(x, y)$ and $e_{\text{lower}}(x, y)$, respectively.

Step 3: Obtain the local mean $e_{\text{mean}}(x, y)$ as equation (1).

$$e_{\text{mean}}(x, y) = \left[e_{\text{upper}}(x, y) + e_{\text{lower}}(x, y)\right]/2 \tag{1}$$

Step 4: Subtract the local mean from the original image as equation (2).

$$h_{lk}(x, y) = h_{l(k-1)}(x, y) - e_{\text{mean}}(x, y) \tag{2}$$

Step 5: Repeat Step 1-4 until reaching the stop criterion and $h_{lk}(x, y)$ is output as a BIMF. The stop criterion is designed as the normalized standard deviation (SD), between $h_{lk}(x, y)$ and $h_{l(k-1)}(x, y)$, defined as equation (3).

$$SD = \sum_{x=0}^{X}\sum_{y=0}^{Y}\left[\frac{\left|h_{l(k-1)}(x, y) - h_{lk}(x, y)\right|^2}{h^2_{l(k-1)}(x, y)}\right] \tag{3}$$

where (x, y) represents the 2D coordinate; X and Y correspond to the total numbers of rows and columns in the input image; l is the index of the lth BIMF; and k denotes the iteration number.

Step 6: When reaching the stop criterion, the BIMF $c_1(x, y)$ is defined as the final result from Step 4:

$$c_1(x, y) = h_{lk}(x, y) \tag{4}$$

After obtaining the BIMF, residue $r_1(x, y)$ is defined as equation (5).

$$r_1(x, y) = h_{l0}(x, y) - c_l(x, y) \tag{5}$$

Step 7: The residue is used as the input image to generate the next BIMF from Step 1:

$$h_{(l-1)0}(x, y) = r_1(x, y) \tag{6}$$

By repeating Steps 1-7, several BIMFs can be generated from the original image. When the extrema points in the residue image are not enough, the sifting process in BEMD is completed. Then, the original image $I(x, y)$ can be represented by a series of BIMFs and a residue term as equation (7).

$$I(x, y) = \sum_{j=1}^{L} c_j(x, y) + r_L(x, y) \tag{7}$$

where $c_j(x, y)$ corresponds to the jth BIMF and $r_L(x, y)$ denotes the residue. It is assumed that the generated BIMFs using the 2D sifting process can better show the textural information at different frequencies (scales) than the sole use of original image. Thus, more information is available for image interpretation by exploiting the BIMFs.

Inspired by the merits of BEMD, this work applies it to SAR image feature extraction in order to enhance SAR ATR performance. We set the stop criterion SD to be 0.24 according to the experiential guidance and results from repetitive tests [47]–[49]. The triangle-based cubic spline interpolation is used for the 2D interpolation in Step 2 and the boundary extension is employed to relieve the boundary effect [48]. Fig. 1 illustrates the generated BIMFs of three SAR images from the MSTAR dataset. As shown, the first two BIMFs could reflect the dominant points in the original image. In addition, some details, which are not significant in the original image, can be embodied in the BIMFs. However, the third BIMF can hardly describe the intuitive information in the original image. Therefore, in this study, original the first two BIMFs are selected to complement the original image for the following target recognition.

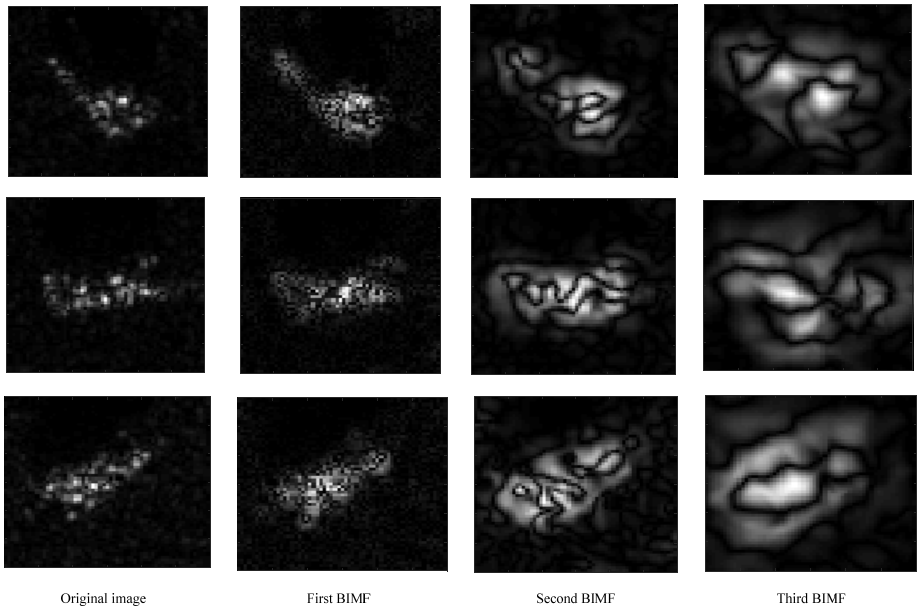


FIGURE 1. Illustration of BIMFs with some MSTAR images.

III. JOINT SPARSE REPRESENTATION FOR TARGET CLASSIFICATION

SRC is a typical application of compressive sensing theory in pattern recognition field, which was first employed by Wright et al. for face recognition [57]. Owing to its good classification performance, it was introduced into SAR target recognition and the reported results validated its effectiveness [23]–[26]. As a generalization of SRC to multi-task problems, JSR was used to classify multiple views [56], monogenic components [15], etc. for SAR ATR. It is assumed that JSR not only represents each task properly but also exploit the inner correlations between different tasks thus improving the overall representation precision. In this study, the original SAR image and its BIMFs actually describe the same target so they are related to some extent. Then, it is preferred to classify them in a unified framework and JSR can be used for the classification problem.

Denote the training samples from C classes as $X_k = \left[x_{k,1}, x_{k,2}, \ldots, x_{k,n_k}\right], k = 1, \ldots, C$. For each training sample, there are several observations, which share some correlations. Specifically, in this study, there are three observations of a single SAR image, i.e., original image, 1st BIMF, and 2nd BIMF. Let $X^{(i)}$ denotes the dictionary formed by individual components, where $i = 1, 2, 3$ correspond to the features extracted from the original image, 1st BIMF, and 2nd BIMF, respectively. For the test sample y to be classified, the corresponding features are denoted as $\left[y^{(1)}, \ldots, y^{(2)}\right]$. Then, the JSR of the three components can be preliminarily formulated as follow:

$$\min_{A} \left\{ g(A) = \sum_{l=1}^{3} \left\| y^{(l)} - X^{(l)} \alpha^{(l)} \right\| \right\} \tag{8}$$

In equation (8), $A = \left[\alpha^{(1)}, \ldots, \alpha^{(2)}\right]$ is the sparse coefficient matrix comprising the coefficient vectors from the three components. However, it should be noted that the objective in equation (8) does not consider the possible correlations between different components. Consequently, there are inevitable discrimination loss from the aspect of target recognition. Hence, the ℓ_1/ℓ_2 mixed-norm can be used to restrain the structure of the sparse coefficient matrix as equation (9).

$$\min_{A} g(A) + \lambda \left\| A \right\|_{1,2} \tag{9}$$

where $\|A\|_{1,2}$ represents the ℓ_1/ℓ_2 mixed-norm of A, which is calculated by first, imposing the ℓ_2 norm on each row of A and then performing the ℓ_1 norm on the resulting vector. Such a constraint impels the coefficient vectors from different component share approaching sparsity patterns, which reflect their inner correlations.

To solve the refined JSR problem in equation (9), several preceding works can be employed such as the simultaneous orthogonal matching pursuit (SOMP) [52] and multi-task compressive sensing learning (MTCS) [53]. Moreover, Dong et al. deduced a numerical solution to the tri-task JSR problem in [15]. After estimating the optimal coefficient matrix $\hat{A}$, the total reconstruction error of each training class is obtained to determine the target label of the test sample as follow:

$$\text{indentity}(y) = \min_{k=1,\cdots,C} \sum_{l=1}^{3} \left\| y^{(l)} - X_k^{(l)} \hat{\alpha}_k^{(l)} \right\|_2 \tag{10}$$

In equation (10), $X_k^{(l)}$ represents the lth component from the kth class; and $\hat{\alpha}_k^{(l)}$ denotes the coefficient vector from the kth class in the lth task.

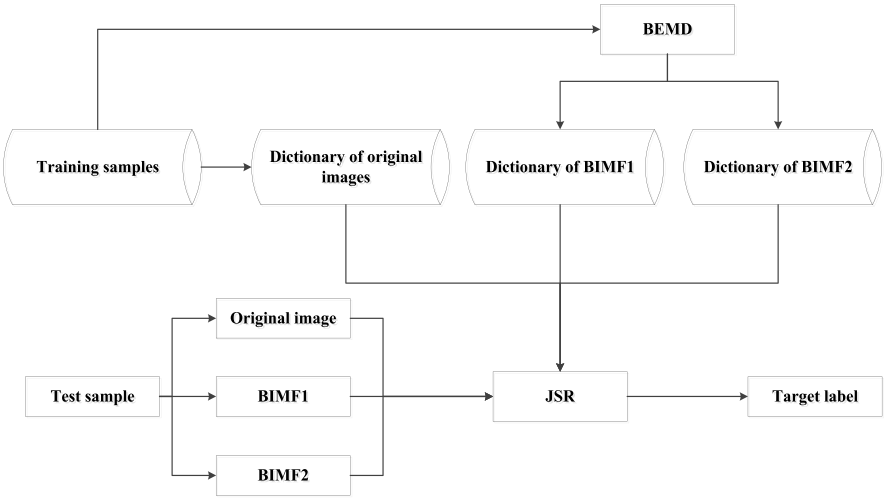


FIGURE 2. Basic procedure of JSR of the original image and BIMFs for target recognition.

FIGURE 3. Optical images of the ten targets.

This study applies JSR to the classification of the original image and its BIMFs for SAR target recognition as show in Fig. 2. The original image together with its BIMFs are used to describe the target because they could provide complementary discrimination for each other. BEMD is performed in the way as the procedure in Fig. 1. To reduce the high dimensionality, the random projection [23], [58] is employed to project the original image and BIMFs as 1024-dimension vectors. Random projection was a recently proposed dimension reduction algorithm based on the compressive sensing theory, which was demonstrated to be a good partner of the sparse representation classifiers like SRC and JSR [55]. The SOMP algorithm is employed to solve the multi-task learning problem. The detailed steps of the proposed target recognition method are as follows:

Step 1: The original training samples and their first two BIMFs are generated to establish the corresponding dictionaries via random projection;

Step 2: The first two BIMFs of the test sample are generated and random projection is performed on them and the original image;

Step 3: The three random projection vectors from the test sample are jointly represented using JSR;

Step 4: The target label is decided to the training class, which represents the feature vectors form the test sample with the minimum reconstruction error.

IV. EXPERIMENTS

A. DATA SET AND BASELINE ALGORITHMS

The MSTAR data set is employed to numerically evaluate the proposed method, which is widely taken as the testbed for SAR ATR algorithms. The data set includes SAR images of ten vehicle targets, whose optical images are shown in Fig. 3. For each target, the collected SAR images cover the full azimuths of $0 \sim 360°$ with a step of about $1 \sim 2°$. Several depression angles are available, e.g., 15°, 17°, 30°, and 45°. The SAR images are measured by the X-band airborne sensors at the resolution of about $0.3m \times 0.3m$.

Based on the MSTAR data set, several different operating conditions can be designed to test the proposed method including the standard operating condition (SOC) and EOCs like configuration variations, depression angle variations [3].

TABLE 1. Image samples for training and testing under SOC.

	BMP2	BTR70	T72	T62	BDRM2	BTR60	ZSU23/4	D7	ZIL131	2S1
Training (17°)	233(SN_9563)	233	232(SN_132)	299	298	256	299	299	299	299
Test (15°)	195(SN_9563) 196(SN_9566) 196(SN_C21)	196	196(SN_132) 195(SN_812) 191(SN_S7)	273	274	195	274	274	274	274

In addition, some simulations can be performed on the original MSTAR images to generated extra EOCs like noise corruption and partial occlusion.

Some baseline SAR ATR algorithms are used for comparison given as follows, which involves classifier-dependent and feature-dependent ones but focus on some classical and recently proposed deep learning-based methods. All these methods are implemented on the same platform (a PC with Intel i7 CPU) for fair comparison.

- SVM: The multi-class SVM in LIBSVM package [59] is employed as the classifier. The radial basis function (RBF) kernel is used, whose parameters are obtained via grid-searching in the package. PCA is used to obtain 80-dimension feature vectors.
- SRC: The SparseLab package [60] is employed to perform SRC. The sparse coefficients are solve by OMP algorithm with a sparsity level of 13 and tolerance error of 10^{-4}. The random projection is used to reduce the original images as 1024-dimension feature vectors in consistency with the proposed method.
- A-ConvNet: the all-convolutional networks designed in [30].
- ESENet: the enhanced squeeze and excitation network used in [33].
- Multi-stream CNN: the multi-stream CNN designed in [34].
- DL+SC: the method based on deep learning and sparse coding proposed in [38].
- Augmented CNN: The original training samples are augmented by noise addition, multi-resolution representation, and partial occlusion. Afterwards, the enriched training samples are used to train the CNN in [36].
- Region Matching: The region matching method proposed in [6] is used. The binary target region from the test sample is compared with its corresponding template samples. And target recognition is performed based on the region residuals.
- ASC Matching: The SAR target recognition method via the matching of attributed scattering centers proposed in [19] is used. The two scattering center sets are matched to reach a similarity measure for target recognition.

In the remainder, the proposed method is first tested under SOC to classify all the ten targets in MSTAR dataset. Afterwards, some typical EOCs are used for the evaluation of the recognition performance including the variations in configurations, depression angles, noise corruption, and partial occlusion.

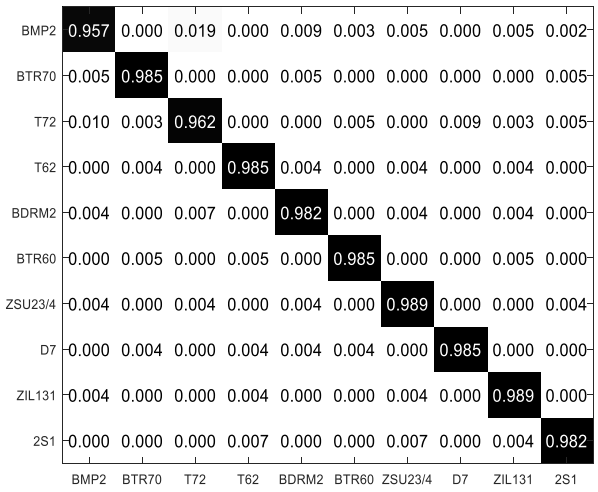


FIGURE 4. Confusion matrix of the proposed method under SOC.

B. SOC

Experiment is first conducted under SOC to preliminarily validate the classification capability of the proposed method. Under SOC, the operating conditions of the test samples are close to those of the training ones. Table 1 tabulates the training and test samples for the recognition problem under SOC. Images acquired at 17° depression angle are used as the training samples, while those at 15° depression angle are tested. It should be noted that there are some configuration variances between the training and test samples of BMP2 and T72, which are denoted by different serial numbers (SN). Here, only SN_9563 from BMP2 and SN_132 from T72 are contained in the training set but three configurations are used in the training set for each target.

The detailed results of the proposed method under SOC are presented as the confusion matrix in Fig. 4. From the diagonal elements of the confusion matrix, we can see that the recognition rate of each target is over 95% and the overall recognition rate of all the ten targets is calculated to be 98.08%. Table 2 compares the overall recognition rates of different methods. It shows that the proposed method performs better than the baseline algorithms except for deep-learning-based ones. The Augmented CNN ranks first among all the methods owing to the good classification ability of the deep learning technique and large amount of training samples. However, due to configurations variants existed in BMP2 and T72, the performance of the networks trained by only a part of the test configurations is impaired to some extent for A-ConvNets, ESENet, DL+SC, and Augmented CNN.

TABLE 2. The proposed method versus baseline algorithms under SOC.

Method	Overall recognition rate (%)
Proposed	98.08
SVM	95.24
SRC	94.66
A-ConvNet	98.12
ESENet	98.34
Multi-stream CNN	98.04
DL+SC	98.17
Augmented CNN	98.67
Region Matching	94.28
ASC Matching	95.07

TABLE 3. Image samples for training and testing under configuration variance.

	BMP2	BDRM2	BTR70	T72
Training (17°)	233 (SN_9563)	298	233	232 (SN_132)
Test set (15°, 17°)	428 (SN_9566) 429 (SN_C21)	0	0	426 (SN_812) 573 (SN_A04) 573 (SN_A05) 573 (SN_A07) 567 (SN_A10)

As a result, they achieve only a slight superiority over the proposed method. In comparison with Region Matching and ASC Matching methods, the better performance of the proposed method validates the high discrimination of the BEMD features, i.e., BIMFs. In addition, JSR is a suitable classification strategy to jointly classify the original image and its BIMFs. Therefore, the proposed method can achieve good recognition performance under SOC.

C. EOC1: CONFIGURATION VARIANCE

Due to the structural modifications, a special class of vehicle target may have different configurations denoted as different serial numbers. In the present experiment, the proposed method is tested under configuration variance using the training and test samples tabulated in Table 3. The test samples of BMP2 and T72 are from different configurations with the training ones. In addition, two outlier targets, i.e., BDRM2 and BTR70, are included in the training set. The results of the proposed method under configuration variance are given in Table 4. We can see each configuration from BMP2 or T72 can be classified with accuracy over 96% and the overall recognition rate reaches 96.86%. The results show that the proposed method could keep its high effectiveness under configuration variance. The proposed method is compared with the baseline algorithms in Table 5. Accordingly, the proposed method achieves the best robustness under configuration variance. Compared with SOC, the configuration variances in this scenario are much more severe. As a result, the overall recognition rate of deep learning-based methods including A-ConvNets, ESENet, Multi-stream CNN, DL+SC, and Augmented CNN falls below the proposed method. As reported in previous literatures, the classification performance of deep learning models is highly related to the

TABLE 4. Results of the proposed method under configuration variance.

Class	Configuration	BMP2	BRDM2	BTR70	T72	Recognition rate (%)
BMP2	SN_9566	413	8	3	4	96.50
	SN_C21	415	8	2	4	96.74
T72	SN_812	11	5	2	408	95.77
	SN_A04	6	5	8	554	96.68
	SN_A05	2	7	4	560	97.73
	SN_A07	4	7	4	558	97.38
	SN_A10	7	1	3	549	96.83
Overall						96.86

TABLE 5. The proposed method versus baseline algorithms under configuration variance.

Method	Overall recognition rate (%)
Proposed	96.86
SVM	94.72
SRC	94.26
A-ConvNet	96.74
ESENet	95.84
Multi-stream CNN	95.16
DL+SC	96.02
Augmented CNN	96.65
Region Matching	95.84
ASC Matching	96.12

TABLE 6. Image samples for training and testing under depression angle variance.

	Depr.	2S1	BDRM2	ZSU23/4	T72 (SN_A64)
Training set	17°	299	298	299	299
Test set	30°	288	287	288	288
	45°	303	303	303	303

amount and coverage of training samples. In this case, the test samples have many differences with the training ones with regard to the target configurations. As a result, the trained networks lose some robustness. The Region Matching and ASC Matching algorithms perform better than SVM and SRC owing to the benefits of their features, i.e., binary target region and attributed scattering centers, respectively. The targets with different configurations still share very close geometrical shape. Then, the binary target regions of different configurations can be well matched to make correct decisions. The attributed scattering centers are local keypoints with rich physical descriptions. So, they can sense the structural modifications caused by configuration variances thus maintaining good robustness. In our method, the generated BIMFs can effectively reflect the details of the target and capture broader spectral information of the original image. Then, the joint classification of the original image and BIMFs can help improve the recognition performance under configuration variance.

D. EOC2: DEPRESSION ANGLE VARIANCE

SAR images are much more sensitive to the azimuth and depression angle. Consequently, SAR images from different depression angles tend to have many differences [61], which results in more obstacles to correct target recognition. The training and test samples in this experiment are presented in Table 6, which include SAR images of four targets

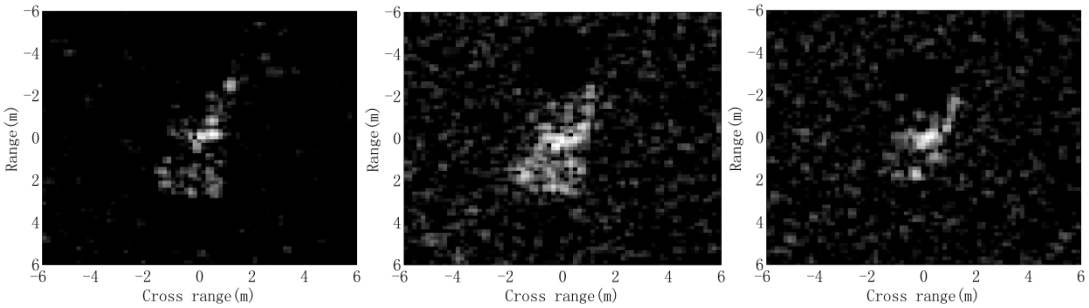


FIGURE 5. Illustration of the differences between SAR images of different depression angles with the 2S1 target. (a) 17° (b) 30° (c) 45°.

TABLE 7. Results of the proposed method under different depression angles.

Depr.	Class	Results				Recognition rate (%)	Overall (%)
		2S1	BDRM2	ZSU23/4	T72		
30°	2S1	281	2	4	1	97.57	97.83
	BDRM2	1	282	1	3	98.26	
	ZSU23/4	2	1	284	1	98.61	
	T72	1	5	3	279	96.88	
45°	2S1	210	52	19	22	69.31	72.28
	BDRM2	34	238	13	18	78.55	
	ZSU23/4	27	41	209	26	69.98	
	T72	19	37	28	219	72.28	

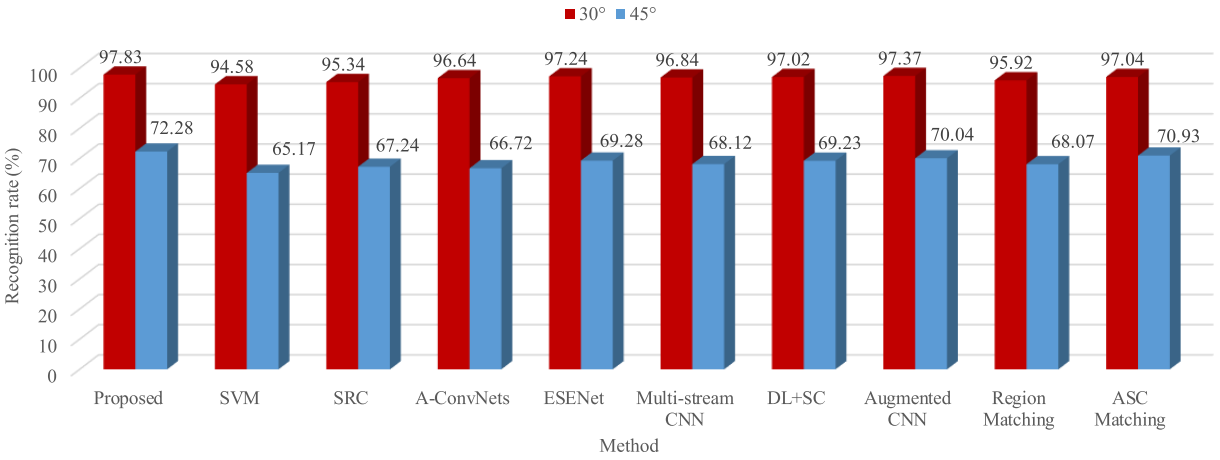


FIGURE 6. The proposed method versus baseline algorithms under depression angle variance.

(2S1, BDRM2, ZSU23/4, and T72(SN_A64)) from three depression angles (17°, 30°, and 45°). Fig. 5 shows the differences between SAR images acquired at different depression angles with the 2S1 target. Taking the image at 17° depression angle as the reference, Fig. 5 shows that the depression angle variance deforms both the target region and radar shadow. In this experiment, images at 17° are trained to classify those at 30° and 45°. Table 7 gives the recognition results of the proposed method at 30° and 45°, respectively, whose overall recognition rates are calculated to be 98.34% and 73.15% correspondingly. The large depression angle change from 17° to 45° causes a severe degradation to the classification accuracy because the test samples have much more differences with the training ones as shown in Fig. 5. Fig. 6 lists the recognition rates of the proposed method and baseline algorithms at different depression angles. All the methods share a similar trend with the change of the depression angle. In comparison, the proposed method maintain the best performance under depression angle variance. Similar to the condition of configuration variance, the deep learning-based methods could hardly compete the proposed one because of the notable differences existed between the training and test samples. It is assumed that the original image together with the generated BIMFs could provide more detailed information about the

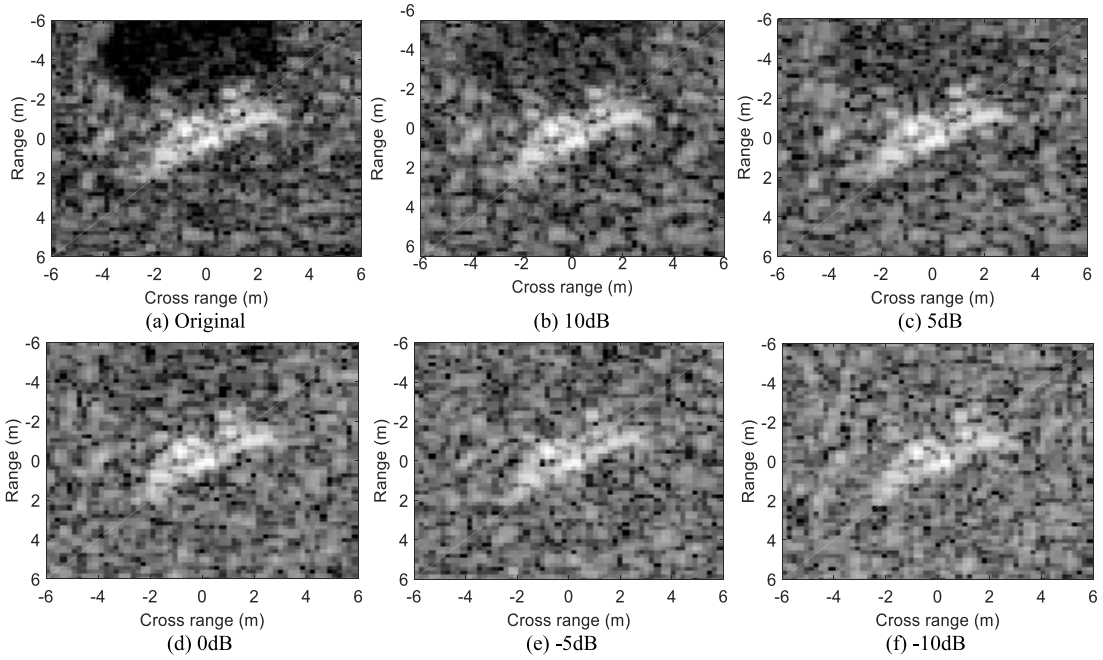


FIGURE 7. Illustration of influence of noise corruption with SAR images from different SNRs.

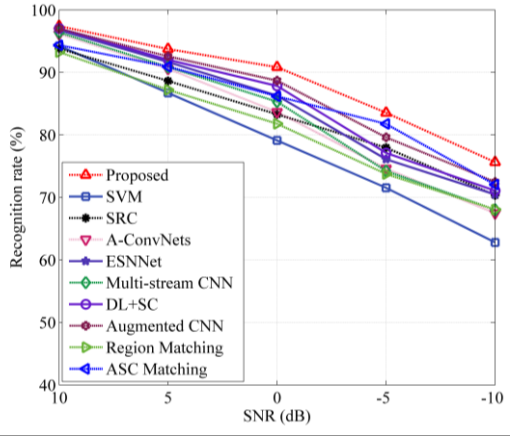
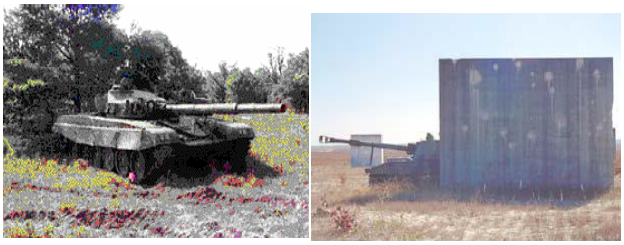


FIGURE 8. The proposed method versus baseline algorithms under noise corruption.



(a) Occlusion by trees (b) Occlusion by man-made walls.

FIGURE 9. Examples of partial occlusion in the real-world scenarios.

target thus help handling the obstacles caused by large depression angle variance. In addition, the ASC Matching algorithm achieves relatively superior performance at 45° depression angle because some of the attribute scattering centers could keep stable under depression angle variances.

E. EOC3: NOISE CORRUPTION

Noise corruption is also a common situation in practical situation because there are inevitable clutters or noises during data acquisition [62]. The original MSTAR data set at collected under cooperative conditions and the noises in the raw data are suppressed before target recognition. So, they are actually with high signal-to-noise ratios (SNR), which relieve the difficulty of target recognition. To examine the robustness of the proposed method to possible noise corruption, several different levels of additive noises are added to the test samples in Table 1. Some exemplar noisy SAR images are illustrated in Fig. 7, which are at different SNRs. Fig. 8 shows the performance of different methods under noise corruption. Noticeably, the proposed method outperforms the baseline algorithms significantly and the predominance becomes more remarkable with the deterioration of noise levels. In comparison with SVM, A-ConvNets, ESENet, and Multi-stream CNN, SRC performs better under low SNRs owing to the merit of sparse representation as validated in [57]. Similarly, the DL+SC method gets better performance because of the advantages of sparse coding. As the simulation of noisy samples in Augmented CNN, its noise robustness is improved to some extent. For the proposed method, its excellent robustness benefits form two factors. First, the generated BIMFs could provide more discrimination to handle the possible noise corruption. Second, JSR is actually a generalization of SRC so the robustness of sparse representation to noise corruption can be inherited. The ASC Matching method also achieve very good performance in this situation mainly

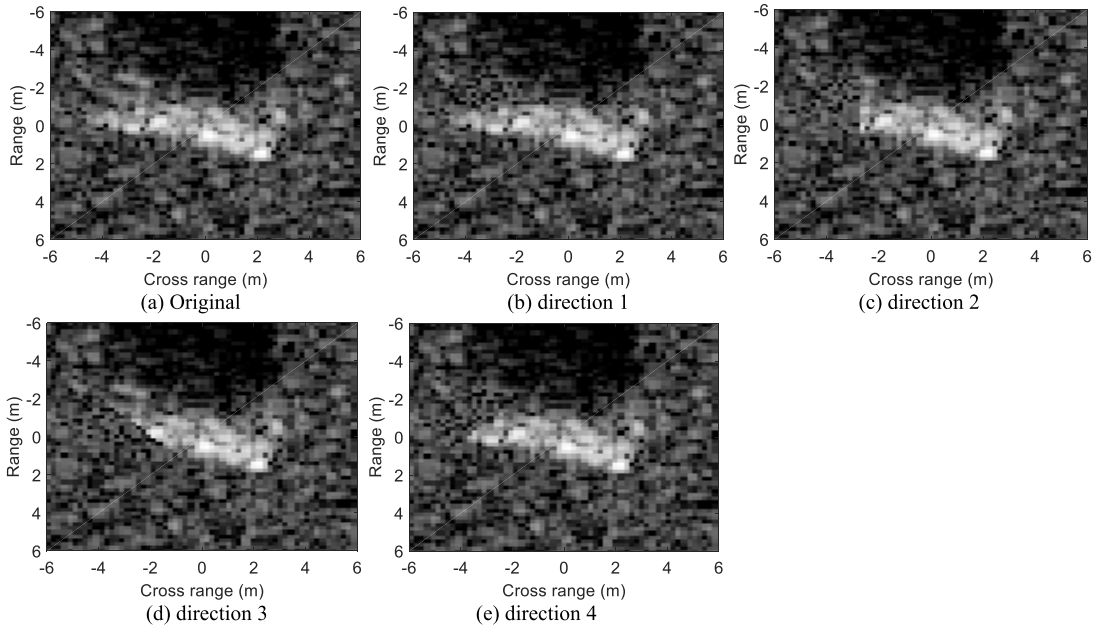


FIGURE 10. Illustration of influence of resolution variance with 20% occluded SAR images from different directions.

because the attribute scattering centers are noise-robust features.

F. EOC4: PARTIAL OCCLUSION

The ground vehicle targets may be occluded by the adjacent objects likes trees or walls as shown in Fig. 9. Consequently, some of the target's characteristics can not be measured in the collected SAR images, which makes it harder to correctly classify the test samples. In this EOC scenario, the occluded images are generated from test samples in Table 1 based on the occlusion model in [19], [63]. A certain proportion of the target's pixels are replaced by the background pixels from a certain direction. The detailed procedure can refer to [19]. Fig. 10 illustrates some partially occluded SAR images from different directions, in which 20% percent of the target region are occluded. The performance of all the methods is shown in Fig. 11. At each occlusion level, the proposed method achieves the highest recognition rate, validating its better robustness to partial occlusion than the baseline algorithms. Similar to the situation of noise corruption, SRC and DL+SC methods perform better than SVM, A-ConvNets, ESENet, and Multi-stream CNN at high occlusion levels because of the robustness of sparse representation (coding). The Augmented CNN achieves better robustness to occlusion owing to the simulation of partially occluded training samples. Therefore, the good robustness of the proposed method comes from the advantages of both the generated BIMFs and sparse representation. The ASC Matching method keeps relatively robust under partial occlusion because the local descriptors, i.e., attributed scattering centers, is able the sense the local variations caused by occlusions. The Region Matching method degrades significantly in this case because the basic feature in it, i.e., binary target region, is severely corrupted by partial occlusion.

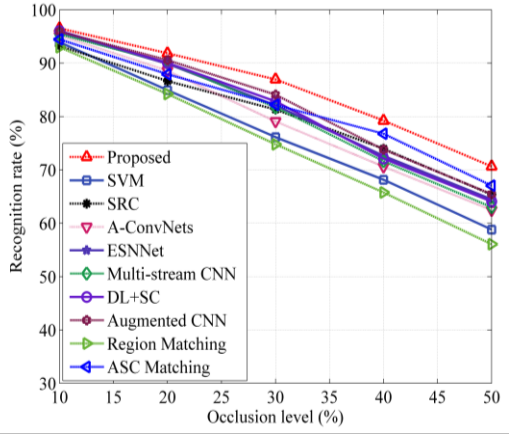


FIGURE 11. The proposed method versus baseline algorithms under partial occlusion.

V. DISCUSSION

In this section, some key messages from the extensive experiments in Section IV are summarized. First, we undertake the experimental investigations under SOC scenario against several baseline algorithms. And the overall recognition rates of different methods on the ten classes of ground vehicles are compared. In this case, the proposed method achieves good performance with a slightly lower recognition rate than deep learning-based methods. In comparison with features only from the original image such as PCA feature vectors, binary target region, and attributed scattering centers, the joint use of the original image and its BIMFs provide more effective information to discriminative different targets. In addition, the cooperative use of the BIMFs and JSR further enhances the recognition performance.

Second, the proposed method is evaluated under several typical EOCs including configuration variance, depression angle variance, noise corruption, and partial occlusion, which are common to see in the real-world environment. Some of the EOCs are directly drawn from the MSTAR data such as configuration variance and depression angle variance; and some are simulated according to some empirical models such as noise corruption and partial occlusion. Overall, the proposed method outperforms the baseline algorithms under different types of EOCs. BEMD is capable of analyzing the details of the target and capturing broader spectral information of the target. Therefore, the combination of the original image and its BIMFs can help reflect the divergences between different targets although the test samples are partly corrupted by EOCs. In addition, JSR inherits the merits of sparse representation such as robustness to noise corruption and occlusion. Therefore, it further strengthens the robustness of the proposed method to EOCs.

Although the performance of the proposed method is not the best under SOC, it still achieves a very high recognition rate of 98.02%. Actually, the SOC recognition task is not a hard problem in SAR ATR nowadays. Differently, various kinds of EOCs including those tested in this study are the real obstacles to the practical application of SAR ATR methods. Comprehensively considering the performance under both SOC and EOCs, the overall effectiveness and practicability of the proposed method has superiority over the baseline methods.

VI. CONCLUSION

In this study, we propose a SAR target recognition method using the features extracted by BEMD. BEMD provides an adaptive and empirical way to process the images and the generated BIMFs is able to capture the details of the original image. Thus, the joint classification of the original image and generated BIMFs could provide more discrimination for robust target recognition. The feasibility of the proposed method is quantitively evaluated on the MSTAR data set under both SOC and different types of EOCs. According to the experimental reports, some conclusions are as follows: (1) BEMD could effectively capture the characteristics of SAR image thus providing complementary information to the original image. (1) The classification via JSR could combine the advantages of the original image and generated BIMFs to enhance the classification accuracy. (2) The proposed method could achieve a high recognition rate under SOC and maintain good robustness under several types of EOCs. As a potential future work, more classification strategies can be designed to exploit the BIMFs to further the recognition accuracy.

REFERENCES

[1] K. El-Darymli, E. W. Gill, P. McGuire, D. Power, and C. Moloney, "Automatic target recognition in synthetic aperture radar imagery: A state-of-the-art review," *IEEE Access*, vol. 4, pp. 6014–6058, 2016.

[2] L. M. Novak, G. J. Owirka, and W. S. Brower, "Performance of 10- and 20-target MSE classifiers," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 4, pp. 1279–1289, Oct. 2000.

[3] J. R. Diemunsch and J. Wissinger, "Moving and stationary target acquisition and recognition (MSTAR) model-based automatic target recognition: Search technology for a robust ATR," *Proc. SPIE*, vol. 3370, pp. 481–492, Sep. 1998.

[4] G. C. Anagnostopoulos, "SVM-based target recognition from synthetic aperture radar images using target region outline descriptors," *Nonlinear Anal., Theory, Methods Appl.*, vol. 71, no. 12, pp. e2934–e2939, 2009.

[5] M. Amoon and G.-A. Rezai-Rad, "Automatic target recognition of synthetic aperture radar (SAR) images based on optimal selection of Zernike moments features," *IET Comput. Vis.*, vol. 8, no. 2, pp. 77–85, Apr. 2014.

[6] B. Ding, G. Wen, C. Ma, and X. Yang, "Target recognition in synthetic aperture radar images using binary morphological operations," *J. Appl. Remote Sens.*, vol. 10, no. 4, 2016, Art. no. 046006.

[7] S. Papson and R. M. Narayanan, "Classification via the shadow region in SAR imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 969–980, Apr. 2012.

[8] J.-I. Park and K.-T. Kim, "Modified polar mapping classifier for SAR automatic target recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 1092–1107, Apr. 2014.

[9] Z. He, J. Lu, and G. Kuang, "A fast SAR target recognition approach using PCA features," in *Proc. 4th Int. Conf. Images Graph.*, Aug. 2007, pp. 1–4.

[10] A. K. Mishra, "Validation of PCA and LDA for SAR ATR," in *Proc. IEEE Region 10 Conf.*, Nov. 2008, pp. 1–6.

[11] Z. Cui, Z. Cao, J. Yang, J. Feng, and H. Ren, "Target recognition in synthetic aperture radar images via non-negative matrix factorisation," *IET Radar Sonar Navigat.*, vol. 9, no. 9, pp. 1376–1385, Dec. 2015.

[12] Y. Huang, J. Peia, J. Yanga, B. Wang, and X. Liu, "Neighborhood geometric center scaling embedding for SAR ATR," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 180–192, Jan. 2014.

[13] M. Yu, L. Zhao, S. Zhang, B. Xiong, and G. Kuang, "SAR target recognition using parametric supervised t-stochastic neighbor embedding," *Remote Sens. Lett.*, vol. 8, no. 9, pp. 849–858, 2017.

[14] H. Wang, S. Li, Y. Zhou, and S. Chen, "SAR automatic target recognition using a roto-translational invariant wavelet-scattering convolution network," *Remote Sens.*, vol. 10, no. 4, p. 501, 2018.

[15] G. Dong, G. Kuang, N. Wang, L. Zhao, and J. Lu, "SAR target recognition via joint sparse representation of monogenic signal," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 7, pp. 3316–3328, Jul. 2015.

[16] L. C. Potter and R. L. Moses, "Attributed scattering centers for SAR ATR," *IEEE Trans. Image Process.*, vol. 6, no. 1, pp. 79–91, Jan. 1997.

[17] B. Ding, G. Wen, J. Zhong, C. Ma, and X. Yang, "Robust method for the matching of attributed scattering centers with application to synthetic aperture radar automatic target recognition," *J. Appl. Remote Sens.*, vol. 10, no. 1, 2016, Art. no. 016010.

[18] B. Ding, G. Wen, C. Ma, and X. Yang, "Decision fusion based on physically relevant features for SAR ATR," *IET Radar, Sonar Navigat.*, vol. 11, no. 4, pp. 682–690, 2017.

[19] B. Ding, G. Wen, X. Huang, C. Ma, and X. Yang, "Target recognition in synthetic aperture radar images via matching of attributed scattering centers," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 7, pp. 3334–3347, Jul. 2017.

[20] Q. Zhao and J. C. Principe, "Support vector machines for SAR automatic target recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 2, pp. 643–654, Apr. 2001.

[21] H. Liu and S. Li, "Decision fusion of sparse representation and support vector machine for SAR image target recognition," *Neurocomputing*, vol. 113, pp. 97–104, Aug. 2013.

[22] S. A. Wagner, "SAR ATR by a combination of convolutional neural network and support vector machines," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 6, pp. 2861–2872, Dec. 2016.

[23] J. J. Thiagarajan, K. N. Ramamurthy, P. Knee, A. Spanias, and V. Berisha, "Sparse representations for automatic target classification in SAR images," in *Proc. 4th Int. Symp. Commun., Control Signal Process.*, Mar. 2010, pp. 1–4.

[24] H. Song, K. Ji, Y. Zhang, X. Xing, and H. Zou, "Sparse representation-based SAR image target classification on the 10-class MSTAR data set," *Appl. Sci.*, vol. 6, no. 1, p. 26, 2016.

[25] S. Song, B. Xu, and J. Yang, "SAR target recognition via supervised discriminative dictionary learning and sparse representation of the SAR-HOG feature," *Remote Sens.*, vol. 8, no. 8, p. 683, 2016.

[26] A. Karine, A. Toumi, A. Khenchaf, and M. El Hassouni, "Target recognition in radar images using weighted statistical dictionary-based sparse representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2403–2407, Dec. 2017.

[27] Y. Sun, Z. Liu, S. Todorovic, and J. Li, "Adaptive boosting for SAR automatic target recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 112–125, Jan. 2007.
[28] U. Srinivas, V. Monga, and R. G. Raj, "SAR automatic target recognition using discriminative graphical models," *IEEE Trans. Aerosp. Electron. Syst.*, no. 50, no. 1, pp. 591–606, Jan. 2014.
[29] D. A. E. Morgan, "Deep convolutional neural networks for ATR from SAR imagery," *Proc. SPIE*, vol. 9475, May 2015, Art. no. 94750F.
[30] S. Chen, H. Wang, F. Xu, and Y.-Q. Jin, "Target classification using the deep convolutional networks for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4806–4817, Aug. 2016.
[31] J. Zhao, Z. Zhang, W. Yu, and T.-K. Truong, "A cascade coupled convolutional neural network guided visual attention method for ship detection from SAR images," *IEEE Access*, vol. 6, pp. 50693–50708, 2018.
[32] R. Min, H. Lan, Z. Cao, and Z. Cui, "A gradually distilled CNN for SAR target recognition," *IEEE Access*, vol. 7, pp. 42190–42200, 2019.
[33] L. Wang, X. Bai, and F. Zhou, "SAR ATR of ground vehicles based on ESENet," *Remote Sens.*, vol. 11, no. 11, p. 1316, 2019.
[34] P. Zhao, K. Liu, H. Zou, and X. Zhen, "Multi-stream convolutional neural network for SAR automatic target recognition," *Remote Sens.*, vol. 10, no. 9, p. 1473, Sep. 2018.
[35] J. Ding, B. Chen, H. Liu, and M. Huang, "Convolutional neural network with data augmentation for SAR target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 364–368, Mar. 2016.
[36] Y. Yan, "Convolutional neural networks based on augmented training samples for synthetic aperture radar target recognition," *J. Electron. Imag.*, vol. 27, no. 2, 2018, Art. no. 023024.
[37] D. Malmgren-Hansen, A. Kusk, J. Dall, A. A. Nielsen, R. Engholm, and H. Skriver, "Improving SAR automatic target recognition models with transfer learning from simulated data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1484–1488, Sep. 2017.
[38] O. Kechagias-Stamatis and N. Aouf, "Fusing deep learning and sparse coding for SAR ATR," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 2, pp. 785–797, Apr. 2019.
[39] Z. Cui, C. Tang, Z. Cao, and S. Dang, "SAR unlabeled target recognition based on updating CNN with assistant decision," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 10, pp. 1585–1589, Oct. 2018.
[40] C. Jiang and Y. Zhou, "Hierarchical fusion of convolutional neural networks and attributed scattering centers with application to robust SAR ATR," *Remote Sens.*, vol. 10, no. 6, p. 819, 2018.
[41] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.
[42] J. H. Cho and C. G. Park, "Multiple feature aggregation using convolutional neural networks for SAR image-based automatic target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 12, pp. 1882–1886, Dec. 2018.
[43] M. Kang, K. Ji, X. Leng, X. Xing, and H. Zhou, "Synthetic aperture radar target recognition with feature fusion based on a stacked autoencoder," *Sensors*, vol. 17, no. 1, p. 192, 2017.
[44] S. Tian, C. Wang, H. Zhang, and B. Bhanu, "SAR object classification using the DAE with a modified triplet restriction," *IET Radar, Sonar Navigat.*, vol. 13, no. 7, pp. 1081–1091, Jul. 2019.
[45] Y. Kwak, W. Song, and S. Kim, "Speckle -noise-invariant convolutional neural network for SAR target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 549–553, 2019.
[46] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math. Phys. Eng. Sci.*, vol. 454, pp. 903–995, Mar. 1971.
[47] G. Rilling, P. Flandrin, P. Gonçalves, and J. M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 936–939, Dec. 2007.
[48] Y. Tian, Y. Huang, and Y. Li, "Image zooming method using 2D EMD technique," in *Proc. WCICA*, Jun. 2006, pp. 10036–10040.
[49] Y. Qin, L. Qiao, O. Wang, X. Ren, and C. Zhu, "Bidimensional empirical mode decomposition method for image processing in sensing system," *Comput. Elect. Eng.*, vol. 68, pp. 215–224, May 2018.
[50] W.-K. Chen, J.-C. Lee, W.-Y. Han, C.-K. Shih, and K.-C. Chang, "Iris recognition based on bidimensional empirical mode decomposition and fractal dimension," *Inf. Sci.*, vol. 221, pp. 439–451, Feb. 2013.
[51] J. C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and P. Bunel, "Image analysis by bidimensional empirical mode decomposition," *Image Vis. Comput.*, vol. 21, no. 12, pp. 1019–1026, 2003.
[52] S. Lahmiri, "Image denoising in bidimensional empirical mode decomposition domain: The role of student's probability distribution function," *Healthcare Technol. Lett.*, vol. 3, no. 1, pp. 67–71, Mar. 2016.
[53] J. Riffi, A. M. Mahraz, and H. Tairi, "Medical image registration based on fast and adaptive bidimensional empirical mode decomposition," *IET Image Process.*, vol. 7, no. 6, pp. 567–574, Aug. 2013.
[54] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Process.*, vol. 86, no. 3, pp. 589–602, 2006.
[55] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92–106, Jan. 2009.
[56] B. Ding and G. Wen, "Exploiting multi-view SAR images for robust target recognition," *Remote Sens.*, vol. 9, no. 11, p. 1150, 2017.
[57] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[58] A. Majumdar and R. K. Ward, "Robust classifiers for data reduced via random projections," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 5, pp. 1359–1371, Oct. 2010.
[59] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. no. 27.
[60] *SparseLab*. Accessed: May 3, 2017. [Online]. Available: http://sparselab.stanford.edu/
[61] B. Ravichandran, A. Gandhe, R. Smith, and R. Mehra, "Robust automatic target recognition using learning classifier systems," *Inf. Fusion*, vol. 8, no. 3, pp. 252–265, 2007.
[62] S. Doo, G. Smith, and C. Baker, "Target classification performance as a function of measurement uncertainty," in *Proc. 5th Asia–Pacific Conf. Synth. Aperture Radar*, Sep. 2015, pp. 1–4.
[63] G. Jones and B. Bhanu, "Recognition of articulated and occluded objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 7, pp. 603–613, Jul. 1999.

MING CHANG received the Ph.D. degree from Shaanxi Normal University, in 2013. He is currently an Associate Professor of the UAV and Artificial Intelligence Experimental Platform with the Shaanxi Provincial Key Laboratory of Cognition and Neuroscience, School of Psychology, Shaanxi Normal University. His current research interests include image processing, remote sensing applications, and artificial intelligence.

XUQUN YOU received the Ph.D. degree from East China Normal University, in 1998. He is currently a Professor of the artificial intelligence experimental platform with the Shaanxi Provincial Key Laboratory of Cognition and Neuroscience, Shaanxi Normal University. His current research interests include image processing and artificial intelligence.

ZHENGYANG CAO received the M.E. degree from Northwestern Polytechnical University, in 2011. He is currently a Senior Engineer and is in charge of the logistics support project for the UAV system of the National Engineering Research Center for UAS, Xi'an ASN UAV Technology Company Ltd.