



# Applied Artificial Intelligence

## An International Journal

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/uaai20>

# SAR Image Target Recognition Method by Global and Local Dictionary Sparse Representation

Hongliang Han, Wei Lu & Fan Feng

To cite this article: Hongliang Han, Wei Lu & Fan Feng (2023) SAR Image Target Recognition Method by Global and Local Dictionary Sparse Representation, Applied Artificial Intelligence, 37:1, 2189674, DOI: [10.1080/08839514.2023.2189674](https://doi.org/10.1080/08839514.2023.2189674)

To link to this article: <https://doi.org/10.1080/08839514.2023.2189674>



© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 16 Mar 2023.



Submit your article to this journal [↗](#)



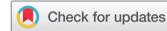
Article views: 363



View related articles [↗](#)



View Crossmark data [↗](#)



# SAR Image Target Recognition Method by Global and Local Dictionary Sparse Representation

Hongliang Han<sup>a</sup>, Wei Lu<sup>b</sup>, and Fan Feng<sup>a</sup>

<sup>a</sup>Department of Health Information Technology Teaching and Research, NCO School, Army Medical University, Shijiazhuang, PR China; <sup>b</sup>Teaching and Research Department of Computer Applied Technology, College of Nationalities, Hebei Normal University, Shijiazhuang, PR China

## ABSTRACT

Target variant detection has been challenging for Synthetic Aperture Radar (SAR), and the performance of the target variant recognition needs to be further enhanced. SAR images are widely used in the field of target recognition and are important to support reconnaissance operations. This paper proposes a novel research approach to improve the efficiency of target recognition. First, the theory surrounding sparse representation and dictionary learning is presented. Secondly, a SAR image target recognition model is constructed according to the theory. Meanwhile, a dynamic joint sparse representation model is proposed based on multi-information and applied to SAR image target recognition when sparse representation is under consideration. Finally, experiments are set up to validate the proposed model. The results are presented as follows: 1. the recognition rate range of the sample SAR image target is 0.811–0.995 and 0.867–0.990, respectively, when the two cases, without registration processing and with registration processing, utilize the recognition method of dictionary learning and sparse representation. 2. with the increase in dictionary size, the average recognition rate of SAR images based on multi-information dynamic joint sparse representation also increases under the conditions of no logarithm transformation and median filtering and after logarithm transformation and median filtering are run. Then, the average recognition rate range is 0.63–0.9 and 0.65–0.96, respectively. Thus, the recognition rate is improved by 5%–10%. 3. the recognition approach based on the sparse representation of multi-information dynamic joint has distinct sparse degrees of SAR image recognition in the case of logarithmic transformation and median filtering when the sample image has been registered and not been registered. The relative average recognition rates were 0.950–0.970 and 0.955–0.977, respectively. The key contribution of the research is to offer a workable solution to the issues plaguing SAR target variant recognition and to improve the significant limitations in the state-of-the-art literature thoroughly.

## ARTICLE HISTORY

Received 29 December 2022

Revised 2 March 2023

Accepted 3 March 2023

**CONTACT** Hongliang Han  [Shzwy2120@163.com](mailto:Shzwy2120@163.com)  Department of Health Information Technology Teaching and Research, NCO School, Army Medical University, Shijiazhuang 050081, PR China

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

## Introduction

The continuous operation of synthetic aperture radar (SAR) under all-weather situations makes it a useful intelligence reconnaissance instrument. SAR automated target recognition (ATR) is a crucial aspect of SAR picture interpretation that has been intensively researched over the past two decades. Particular SAR target recognition approaches concentrate mostly on two technologies: feature extraction and classifier development. The objective of feature extraction is to minimize the dimensionality of the original SAR data to facilitate future decision-making. The target area, target shape, scattering centers, principal component analysis (PCA) features, linear discriminant analysis (LDA) features, image decompositions, etc., are popular SAR target recognition characteristics. The objective of the classifier design is to select the target category based on the original SAR picture or the derived data. Many classifiers, such as K-nearest neighbor (KNN), support vector machine (SVM), sparse representation-based classification (SRC), and convolutional neural network (CNN), have been used for SAR target detection as pattern recognition technology has advanced. Also, other implementations employed PCA and LDA to extract SAR image features and categorize them using KNN. In addition, the elliptical Fourier descriptor was used for the SAR target contour, and then used the SVM classifier to recognize the target. The classification performance and application breadth of various classifiers are often distinct. Consequently, to increase the performance of recognition, it is required to choose a more efficient and robust classifier.

On the other hand, sparse representations have attracted a lot of interest recently in several disciplines, particularly pattern recognition. Sparse representations try to represent signals with the smallest number of significant coefficients possible. This is crucial for a variety of applications, including compression. It is often possible to achieve a high compression rate with practically undetectable data loss using wavelets. Using SAR pictures for automated target detection and decision-making is gaining popularity. The success of such tasks is contingent on the degree to which the reconstructed SAR pictures display particular characteristics of the underlying scene using sparse representation techniques. We create an image generation approach that formulates the SAR imaging issue as a sparse signal representation problem based on the fact that typical underlying sceneries display sparsity in terms of such characteristics. Sparse signal representation, mostly utilized in real-valued issues, provides several capabilities for reconstruction and recognition tasks, including superresolution and feature improvement. For complex-valued issues, such as SAR, selecting the dictionary and representation method for effective sparse representation is a significant difficulty. As we are often interested in characteristics of the SAR reflectivity field's magnitude, our novel method is intended to express the magnitude of the complex-valued

dispersed field sparsely. This transforms the picture reconstruction task into a joint optimization problem, including the size and phase of the underlying field reflectivities. We construct the mathematical underpinning for this strategy and suggest an iterative solution for the associated joint optimization issue. In terms of creating high-quality SAR pictures and displaying robustness in the face of ambiguous or restricted data, our experimental findings illustrate the superiority of our technology over earlier approaches.

Many scholars have employed sparse representation to improve pattern recognition, face recognition, image reduction, etc. For example, a face recognition technique based on Wright's sparse representation is the standard for all applications. This method calculates the reconstruction error for face recognition based on the sparse representation result utilizing the training sample data to build a dictionary. Then the training stage achieves a relatively robust recognition suited to recognize picture targets with local similarity even when the face is partially obscured by noise, pollution, etc. Some of the most significant methodologies of the literature review are presented in the next section of the paper. Specifically, [Section 2](#) is the literature review section. [Section 3](#) describes the methods and the design of the model. Exemptions for applying the proposed method are outlined in [Section 3](#). [Section 4](#) includes details about the experimental design, and finally [section 5](#) concludes by summarizing the findings and drafting the following potential directions for the work.

## Literature Review

Numerous researchers have successfully incorporated sparse representation into the field of Synthetic Aperture Radar (SAR) to detect target images (Ma, Jia, and Hu 2020). The Method of Optimal Directions (MOD) for learning dictionaries was proposed in 1999. To train the vocabulary, the technique takes training samples from real-world images. The generated dictionary atoms exhibit Gabor-like behaviors and have receptive field characteristics that are extremely comparable to those of basic cells (Zhou and Xu 2020). However, the MOD method's high level of complexity entails a significant matrix inversion operation when updating dictionary rules (Guo and Feng 2020). The joint orthogonal dictionary approach was put up in 2005. This technique creates a redundant dictionary by concatenating many orthogonal dictionaries (Wang et al. 2021). Due to the unique structure of the dictionary, the Block Coordinate Relaxation (BCR) algorithm can be employed to update the sparse representation coefficients in units of blocks during sparse coding to speed up the solution process of sparse coding.

Additionally, the strong structural attributes of the dictionary also make the updated rules of the dictionary have a stronger theoretical basis (Shi et al. 2021). In 2006, the K-Singular Value Decomposition (K-SVD)

dictionary learning algorithm was proposed. The singular value decomposition technique is employed to resolve the rank approximation problem, and each atom is updated one by one, which simplifies the computational process (Xue et al. 2020). Basis functions like the receptive fields of cells in the primary visual cortex were obtained by utilizing the sparse coding neural gas algorithm (Xu 2021). It was later proposed in the idea of sparse orthogonal transformation. So, images are clustered according to certain characteristics. Then, an orthogonal dictionary is trained separately for each class. The algorithm utilizes an orthogonal dictionary, making forward and inverse transformations easy.

However, this method requires clustering the entire image. This undoubtedly increases the time to process and space complexity of the algorithm (Yan et al. 2020). The translation invariance dictionary method is designed since similar structural features appear at different positions in dictionary atoms (Yi and Zhao 2020). In 2010, under the assumption that dictionary atoms themselves also have redundant structural information, a double-sparse dictionary learning algorithm was proposed (Yang, Tang, and Tang 2021). In the later proposed hierarchical dictionary idea, the atoms in the dictionary that contribute the most energy to the signal representation are distributed in the first layer. Atoms with smaller contributions are distributed among the remaining dictionary levels. The experimental results show that the convergence and efficiency of the algorithm are better than the K-SVD algorithm (Xing et al. 2021).

It is practically impossible to gather training samples in all states or configurations of the target due to the complexity of the target environment, target configuration, and target structure. Thus, it cannot represent all situations in the real world. Moreover, training has difficulties in SAR target variant recognition, and it is necessary to improve the recognition performance of target variants. Therefore, the research employs dictionary learning techniques and sparse representation models to investigate target recognition in SAR images.

The novelty is presented that a multi-information dynamic joint sparse representation model based on the sparse representation model is proposed in addition to the study of the recognition effect of the sparse representation model in SAR images. The performance of the model will be compared by implementing experiments. The effect of recognition is confirmed. The objective of the research is to offer a workable concept to recognize targets effectively in SAR images.

The rest of the manuscript is constructed as follows: [Section 2](#) defines the methods and model design. [Section 3](#) introduces the experimental design. [Section 4](#) presents the results and analysis of the conducted research with interesting outcomes. [Section 5](#) is allocated to discussion. [Section 6](#) concludes the research.

## Methods and the Design of the Model

### *Dictionary Learning and Sparse Representation: Method of Optimal Directions*

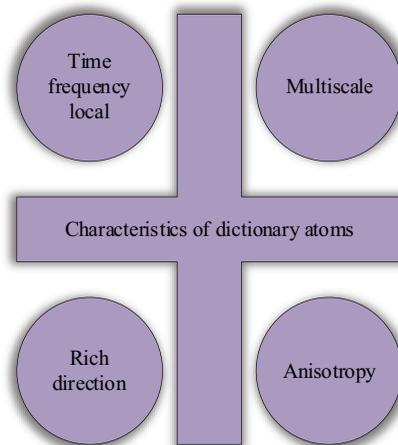
#### *The Method for the Construction of the Dictionary*

The overcomplete (or redundant) dictionary is the basic premise of sparse representation of signals. The degree to which the dictionary atoms accurately describe the signal determines whether the representation coefficients of the signal in the dictionary are sparse (Huang, Xiao, and Yin 2020). The representation coefficients of the signal are sparser and closer than the dictionary atoms and the signal agrees. Therefore, obtaining a dictionary that matches the image signal is one of the core contents of the sparse representation theory. Researchers pointed out that the atoms of an ideal dictionary in which the image signal is sparsely represented should have the characteristics shown in Figure 1 (Zhang and Li 2021).

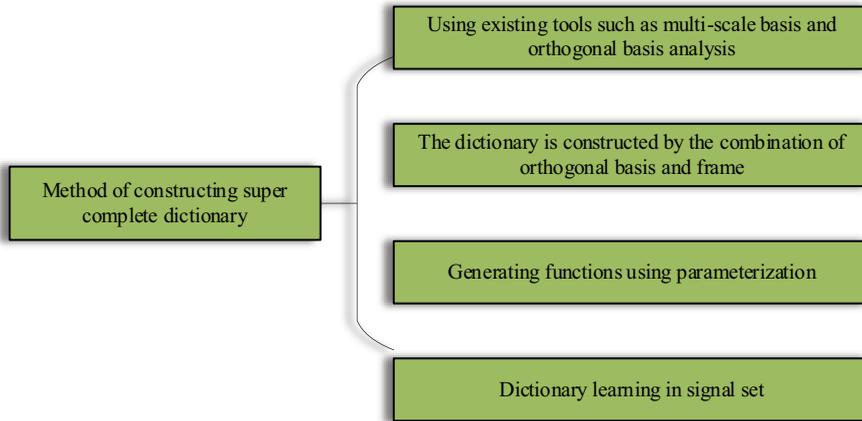
Figure 1 depicts that to be able to represent various local structural features in the image sparsely, the types and numbers of atoms need to be increased. Therefore, overcomplete dictionaries were proposed (Luo et al. 2021). So far, the construction of the method classification of the overcomplete dictionary has been shown in Figure 2.

#### *The Model of Sparse Representation*

Suppose that there are C-type training samples, and the c-th type of training samples form a small dictionary  $A_c$ , then the K-type training samples can form a dictionary  $A = [A_1, A_2, \dots, A_c]$ ,  $c = 1, 2, \dots, C$ . Hypothesis is constructed as follows: any sample of class c can be linearly represented by the training samples of that class. Then, a test sample can be linearly represented by the class of training samples to which it belongs, i.e.,  $y = Ax_0$ . If the test sample



**Figure 1.** The features of dictionary atoms.



**Figure 2.** The method to construct an overcomplete dictionary.

belongs to class  $c$ , then  $x_0 = [0, \dots, 0, u_{c,1}, u_{c,2}, \dots, 0, \dots, 0]^T$ . Its expression is generalized as  $y = Ax$ . Then, according to dictionary  $A$ , the test sample  $y$  is sparsely resolved, as shown in Eq. (1).

$$\hat{x}_2 = \operatorname{argmin} \|x\|_2, s.t. Ax = y \quad (1)$$

The resulting solution is not unique, but the coefficient  $x_2$ , which is not necessarily sparse, is obtained. According to the data, numerous training target classes linearly represent test samples. Ideally, the test target is only relevant to the training samples of the class to which it belongs. Therefore,  $l_2$  is transformed to the  $l_0$  norm to guarantee the sparsity and uniqueness of the coefficients, as shown in Eq. (2).

$$\hat{x}_0 = \operatorname{argmin} \|x\|_0, s.t. Ax = y \quad (2)$$

Orthogonal matching pursuit and other greedy algorithms can obtain approximate resolutions. Meanwhile, Eq. (2) is a Non-deterministic Polynomial-hard (NP-hard) problem, and it is difficult to obtain a more accurate solution. But if  $x_0$  is sparse enough, Eq. (2) can be equivalent to Eq. (3).

$$\hat{x}_1 = \operatorname{argmin} \|x\|_1, s.t. Ax = y \quad (3)$$

The resolution for  $l_0$  is transformed into the resolution for the norm  $l_1$ . In this way, many methods can obtain more accurate coefficients (Damotharasamy 2020).

### **The Classification of a Sparse Representation**

The test samples are categorized based on the reconstruction error after the sparse coefficients are produced according to the sparse solution. In Eq. (4), the categorization criteria are displayed.

$$\min r_c(y) = \|y - A\delta_c(\hat{x}_0)\|_2 \quad (4)$$

Among them,  $rc(y)$  represents the reconstruction error;  $\delta_c(\hat{x}_0)$  means taking the value at the position corresponding to the  $c$ -th target in  $\hat{x}_0$ , and the remaining position values equal to 0 (Jabs, Acharya, and Denniston 2021).

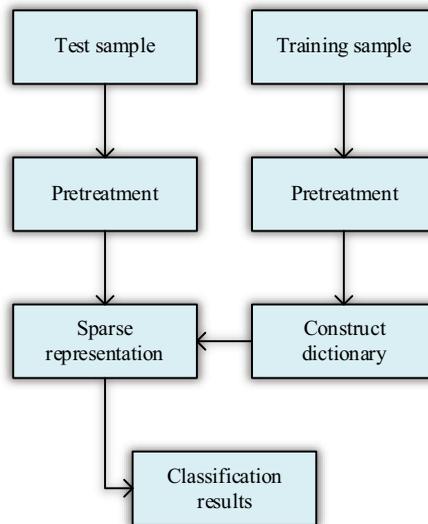
### ***SAR Target Recognition Based on Sparse Representation and Dictionary Learning***

According to the proposed sparse representation model, the given SAR image recognition process based on sparse representation and dictionary learning is shown in Figure 3.

Pre-processing is defined as follows: the training and test samples are pre-processed by segmentation, registration, and interception. The experiment examines the effect of registration on the experimental outcomes by comparing the recognition results in unregistered and registered cases, respectively.

Constructing a dictionary is defined as follows: the intercepted training sample images are column vectorized and spliced into a dictionary. The eigenvectors are subjected to dimensionality reduction. Random matrices are employed to randomly project vectors to reduce computational complexity.

Sparse representation is defined as follows: the truncated test sample images are column vectorized. The test sample vector is sparsely represented



**Figure 3.** The recognition process of SAR target based on sparse representation and dictionary learning.

according to the sparse representation model, and the sparse representation coefficients are obtained.

Classification is described as follows: referring to Eq. (4), sparse coefficients are used to reconstruct the vector of the test sample. According to the size of the reconstruction error, the test objects are identified (Hajipour, Namin, and Shirazi 2021).

### ***SAR Target Recognition Based on Multi-Information Joint Dynamic Sparse Representation and Dictionary Learning***

#### ***The Joint Sparse Representation Model***

The sparse joint model refers to the linear representation of multiple input signals utilizing the same dictionary atoms on the same dictionary. Like the sparse representation model of a single input signal,  $M$  images of the same target are incorporated into the sparse representation model, as shown in Eq. (5).

$$\{\hat{x}_i\}_{i=1}^M = \operatorname{argmin} \sum_{i=1}^M \|y_i - Ax_i\|_2^2, s.t. \|x_i\|_0 \leq S, \quad 1 \leq i \leq M \quad (5)$$

The elements in Eq. (5) are expressed in the form of a matrix, as shown in Eq. (6).

$$X = \operatorname{argmin} \|Y - AX\|_F^2, s.t. \|X\|_{l_0 \setminus l_2} \leq S \quad (6)$$

Among them,  $X = [x_1, x_2, \dots, x_i, x_M]$  is the sparse coefficient matrix;  $Y = [y_1, y_2, \dots, y_i, y_M]$  is the input signal matrix, which refers to the input information formed by  $M$  images of the same test target matrix;  $x_i$  is the sparse coefficient vector of the input signal on the dictionary  $A$ ;  $S$  is the sparsity;  $\|X\|_{l_0 \setminus l_2}$  means that after each row of the matrix is resolved by the  $l_2$  norm, the obtained vector is resolved by norm  $l_0$ .

Again, the classification of the final test target is like the single-signal sparse representation. The magnitude of the comparison reconstruction error is utilized to achieve classification, as shown in Eq. (7).

$$\hat{c} = \operatorname{argmin} \|Y - \hat{Y}_C\|_F = \operatorname{argmin} \|Y - A\delta_C(\hat{X})\|_F \quad (7)$$

Then,

$$\hat{Y}_C = A\delta_C(\hat{X}) \quad (8)$$

$\hat{Y}_C$  represents the reconstruction of the test samples utilizing the  $c$ -th training samples.

When compared with the single-signal sparse representation model, the joint sparse representation model restricts multiple input signals to have the same sparsity pattern. The model comprehensively utilizes the similarity and

correlation between multiple input signals to enhance the final recognition performance of the test target (Peng 2020).

### ***The Joint Dynamic Sparse Representation Model***

The joint dynamic sparse representation model has two types of applications. The first method divides the training sample image into multiple different sub-regions, constructs corresponding dictionaries for the images of different sub-regions, and utilizes the joint dynamic sparse representation model to sparsely resolve the images of different sub-regions of the test sample. The error obtains the final recognition result (Wu et al. 2021). The second method implements all the training samples to form a dictionary and utilizes the joint dynamic sparse representation model to resolve the multiple images of the test samples sparsely. This approach is like the joint sparse representation model, but it emphasizes that multiple test sample images have similar sparse representation patterns in the same dictionary (Wu, Yong, and Fan 2020).

The first joint dynamic sparse representation model is shown in Eqs. (9)-(11).

$$\hat{X} = \underset{X}{\operatorname{argmin}} \sum_{k=1}^K \|y_k - A_k x_k\|_2^2, \text{ s.t. } \|X\|_G \leq S \quad (9)$$

$$\|X\|_G = \left\| \left[ \|x_{g_1}\|_2, \|x_{g_2}\|_2, \dots \right] \right\|_0 \quad (10)$$

$$x_{g_s} = X(g_s) = [X(g_s(1), 1), \dots, X(g_s(K), K)]^T \in R^K \quad (11)$$

$X$  is the obtained sparse coefficient matrix;  $y_k$  is the signal of the  $k$ -th subregion of the input sample;  $x_k$  is the sparse coefficient vector corresponding to the input signal  $y_k$ ;  $A_k$  is the dictionary formed by the  $k$ -th subregion of the training sample;  $S$  is the sparsity;  $x_{g_s}$  represents an index vector of non-zero elements in the sparse coefficient matrix.

The second joint dynamic sparse representation model is shown in Eqs. (12)-(13).

$$\hat{X} = \underset{X}{\operatorname{argmin}} \sum_{m=1}^M \|y_m - Ax_m\|_2^2, \text{ s.t. } \|X\|_G \leq S \quad (12)$$

$$x_{g_s} = X(g_s) = [X(g_s(1), 1), \dots, X(g_s(M), M)]^T \in R^M \quad (13)$$

Among them,  $y_m$  is the  $m$ -th image of the input sample;  $x_m$  is the sparse coefficient vector corresponding to the input signal  $y_m$ ;  $A$  is a dictionary composed of training samples.

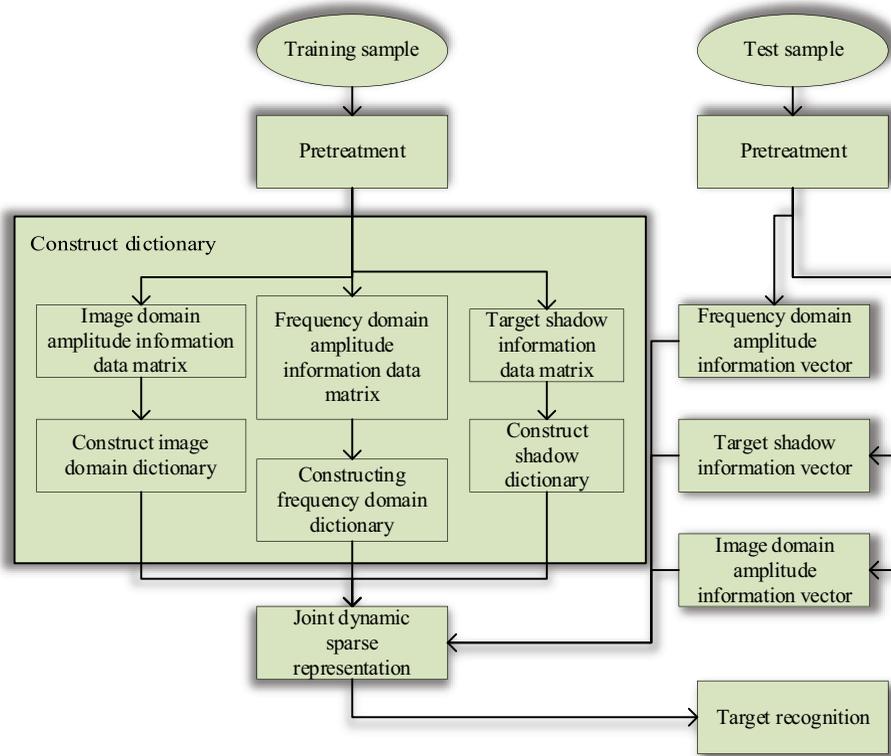
## ***SAR Target Recognition Based on Multi-Information Joint Dynamic Sparse Representation and Dictionary Learning***

### ***The Identification of the Process Design***

The flowchart for the SAR target recognition method follows the associated theory of the joint sparse representation model and the joint dynamic sparse representation model based on the multi-information joint dynamic sparse representation is shown in [Figure 4](#).

Image pre-processing is described as follows: SAR images, in contrast to optical images, are very azimuth sensitive. Coherence speckles are also easily able to degrade the image quality. Furthermore, the target location information is more sensitive to the target image domain information. As a result, the pre-processing of a target picture, shadow segmentation, target registration, and interception are required before target detection in SAR images is realized. The image is pre-processed, such as segmentation, registration, and an interception, and the target image used for final recognition is obtained.

Constructing the dictionary is described as follows: through the pre-processing method, all the original training samples are processed. The



**Figure 4.** The flow chart of SAR target recognition method based on multi-information joint dynamic sparse representation.

training samples for recognition of size  $63 \times 63$  are obtained. On the one hand, the image-domain magnitude information of all training samples used for recognition is column-vectorized. The data matrix of the formed image domain amplitude information is represented by  $S_1 = [t_1, t_2, \dots, t_{Num}] \in R^{d \times Num}$ . Among them,  $t_i \in R^d$  represents the image domain amplitude information vector of the  $i$ -th training sample;  $i$  represents the dimension of the information vector; Num is the number of training samples. On the other hand, the images of all training samples used for recognition are transformed to the frequency domain by a two-dimensional (2D) Fourier transformation. The frequency-domain amplitude information is quantized by a column vector and the frequency domain amplitude information data matrix is denoted by  $S_2 = [p_1, p_2, \dots, p_{Num}] \in R^{d \times Num}$ . Among them,  $p_i \in R^d$  represents the frequency domain amplitude information vector of the  $i$ -th training sample.

Additionally, the original training images are segmented. The binary image of the shadow area of the object is obtained by utilizing segmentation processes. Among them, the amplitude of the shadow area is 1, the amplitude of the other areas is 0, and the amplitude information of the binary image is converted into a column vector to form the target shadow information matrix  $S_3 = [s_1, s_2, \dots, s_{Num}] \in R^{d \times Num}$ . Among them,  $s_i \in R^d$  represents the shadow information vector of the  $i$ -th training sample. The three kinds of information of the training samples directly construct the dictionary, namely  $A1=S1$ ,  $A2=S2$ , and  $A3=S3$ .

Pre-processing of test samples is defined as follows: the test samples are processed utilizing the same steps as the training phase is conducted. To determine the information vector  $y1$  of the test sample image domain amplitude, the pre-processed test picture is employed, and the frequency domain amplitude information vector  $y2$ , and the target shadow information vector  $y3$  are found.

The joint dynamic sparse representation of test samples is described as follows: when combined with the dictionaries  $A1$ ,  $A2$ , and  $A3$  constructed in the training phase, the joint dynamic sparse representation model sparsely resolves the image domain amplitude vector  $y1$ , frequency domain amplitude vector  $y2$ , and shadow information vector  $y3$  of the test sample. The sparse representation coefficients  $x1$ ,  $x2$ , and  $x3$  of the two kinds of information are based on the test samples when their respective dictionaries are obtained.

Identification of test samples is described as follows: the minimum reconstruction error criterion is employed to identify the test target, as shown in Eq. (14).

$$\hat{c} = \underset{k=1}{\operatorname{argmin}} \sum_{k=1}^K \omega_k \|y_k - A_k \delta_c(x_k)\|_2^2 \quad (14)$$

$\hat{c}$  represents the category of the test target;  $\omega_k$  represents the weight of the information;  $\|y_k - A_k \delta_c(x_k)\|_2^2$  represents the reconstruction error of the k-th information vector of the test sample;  $\delta_c(x_k)$  represents the coefficient value of the position corresponding to the c-th category in  $x_k$ ; the value of  $\delta_c(x_k)$  corresponding to the position of other categories is 0.

## Experimental Design

The experimental data comes from the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset. The dataset contains target images at multiple elevation angles and multiple azimuth angles. The resolution of the target image is 0.3 m  $\times$  0.3 m. The training samples are BMP2 (BMP2SN9563, BMP2SN9566, and BMP2SNC21), armored vehicles BTR70SNC71, and main battle tanks T72 (T72SN132, T72SN812, and T72SNS7) when the radar pitch angle is 17°. The test samples are images of 3 categories and seven models when the radar elevation angle is 15°, including armored vehicles BMP2 (BMP2SN9563, BMP2SN9566, BMP2SNC21), armored vehicles BTR70SNC71, and main battle tanks T72 (T72SN132, T72SN812, T72SNS7). The specific content of the parameters taken in the experiment is shown in Table 1.

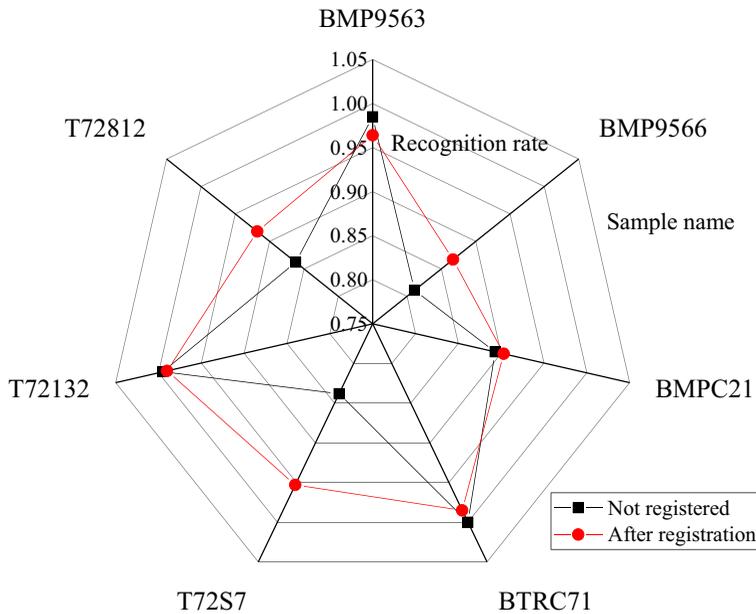
### Target Recognition Effect Based on Sparse Representation and Dictionary Learning

Figure 5 shows the results obtained by the recognition method based on sparse representation and dictionary learning to identify the target of the sample SAR image.

Figure 5 depicts that in the case of no registration process, the recognition rate of this recognition method is 0.985, 0.811, 0.893, 1, 0.838, 0.995, and 0.862, respectively when recognizing the target of the sample SAR image. On the other hand, in the case of registration processing, when the recognition model detects the target of the sample SAR image, the recognition rates are 0.964, 0.867, 0.903, 0.985, 0.953, 0.990, and 0.918, respectively. The data show that without registration processing, the recognition results of non-target variants are better, but the recognition results of target variants are poor.

**Table 1.** Settings of MSTAR experimental parameters.

Data Category		The Training Samples (Pitch Angle =17°)	The Test Samples (Pitch Angle =15°)
BMP2	SN9563	233	195
	SN9566	232	196
	SNC21	233	
BTR70	C71		
T72	SN132		
	SNS7	228	191
	SN812	231	195



**Figure 5.** Model recognition effects based on sparse representation and dictionary learning in different situations.

After the registration process is conducted, the recognition performance of non-target variants remains stable, and the recognition results of target variants are improved to a certain extent. For the two variants of the T72 tank, the recognition rate has increased by 5%-10%. Thus, the average recognition rate has also increased by 2.8%. The data shows that the recognition method based on sparse representation has higher requirements on the location registration of the target in the SAR image.

When the registration process is not performed, the recognition results are not very robust. But after registering all the targets, the method is feasible and effective, and the operation is simple. It directly utilizes the amplitude information of the target area and surrounding of the image without extracting other features of the target and does not consider the azimuth angle of the target, etc., information; no need to design other classifiers and employs the reconstruction error for identification.

### ***The Effect of Image Recognition Based on Multi-Information Joint Dynamic Sparse Representation and Dictionary Learning***

#### ***Image Recognition Results Without Registration***

The sample SAR picture is recognized using the multi-information joint dynamic sparse representation and dictionary learning identification method in the unregistered scenario, as shown in [Figure 6](#).

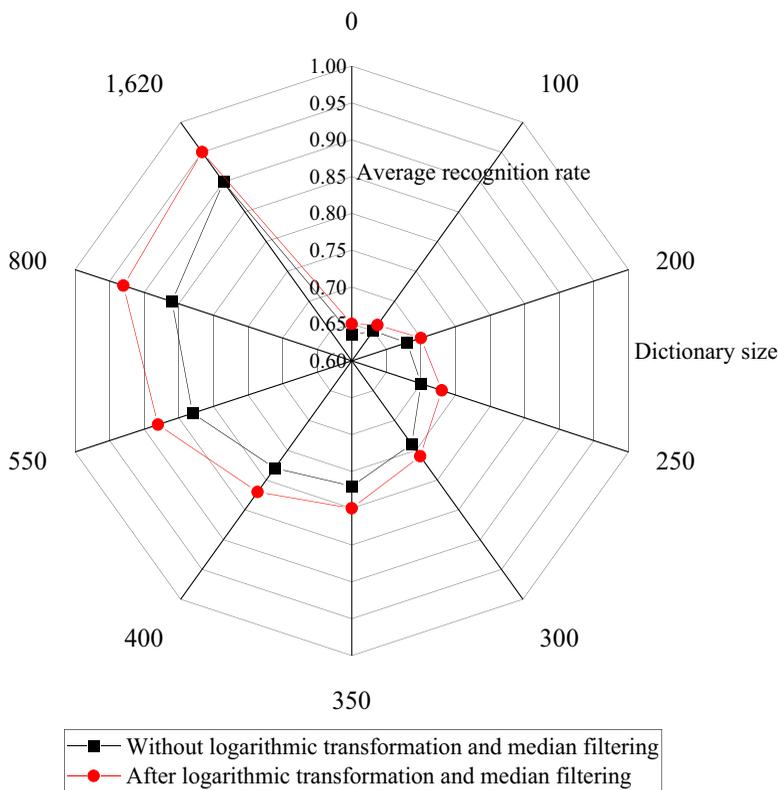
Without logarithmic transformation and median filtering, as shown in [Figure 6](#), the average identification rate of this technique for SAR images similarly increases as the vocabulary size increases. The recognition rate ranges from 0.63 to 0.99 on average. The average recognition rate of the method for SAR images with logarithmic transformation and median filtering rises as the dictionary size does.

The average recognition rate is between 0.65–0.96. Overall, the average recognition rate of this method is improved by 5%-10%. The data show that the number change and median filter processing can improve the target recognition performance in this experimental scenario.

### *Image Recognition Results After Logarithmic Transformation and Median Filtering*

In the case of logarithmic transformation and median filtering, the recognition technique based on multi-information joint dynamic sparse representation and dictionary learning is utilized to detect the sample SAR image, as shown in [Figure 7](#).

When the sample image is not registered in [Figure 7](#), following logarithmic transformation and median filtering, the average identification rate of the recognition method for SAR images under varying sparsity is between 0.955



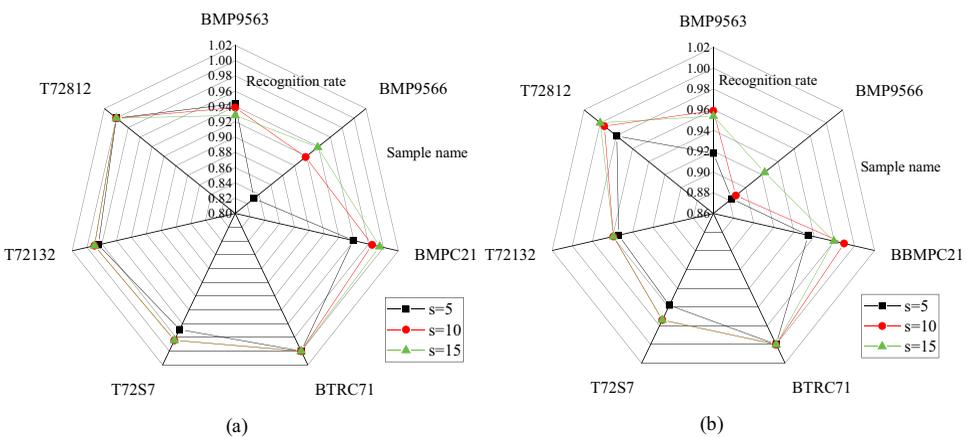
**Figure 6.** Results of image recognition without registration.

and 0.977. The average recognition rate of the SAR image recognition approach under various sparsity conditions ranges from 0.950 to 0.970 when the sample images are registered, and this rate increases as sparsity increases. This recognition technique’s typical recognition rate has increased.

The average recognition rate of this method after registration processing is slightly lower than the average recognition rate without registration processing. In this experimental scenario, both training and target are assumed to be seven classes, but there are three classes of targets. Among the three models of BMP2, the three models of T72 are relatively close. After the registration, the influence of the difference in the position of the target in the image is removed, and the target and the target variant would become more similar. Therefore, it is easy to misclassify each other when the seven categories are divided.

After the registration process is conducted, the recognition performance of non-target variants remains stable, and the recognition results of target variants are improved to a certain extent. For the two variants of the T72 tank, the recognition rate has increased by 5%-10%. Thus, the average recognition rate has also increased by 2.8%. The data shows that the recognition method based on sparse representation has higher requirements on the location registration of the target in the SAR image.

When the registration process is not performed, the recognition results are not very robust. But after registering all the targets, the method is feasible and effective, and the operation is simple. So, it directly utilizes the amplitude information of the target area and surrounding of the image without extracting other features of the target and does not consider the azimuth angle of the target, etc., information; no need to design other classifiers and employs the reconstruction error for identification.



**Figure 7.** Image recognition results after logarithmic transformation and median filtering (a) Model recognition effect without registration processing; (b) Model recognition effect after registration processing.

When logarithmic transformation and median filtering are conducted with the increment of the dictionary size, the average detection ratio of the proposed approach is improved by 5%-10%. The data show that the number change and median filter processing can improve the target recognition performance in this experimental scenario.

## Conclusion

This investigation primarily examines sparse representation-based approaches for SAR image target detection, with a particular emphasis on sparse joint representation, joint dynamic sparse representation, and dictionary learning.

The benefits of the two models of discriminative dictionary learning and joint dynamic sparse representation are combined in the SAR target recognition approach based on these two techniques. Still, it maintains good recognition performance when the dictionary size is small. Additionally, the sparse joint representation exploits the local similarity between the target and target variants and improves the recognition of target variants. The conclusions are as follows:

- (1) after all targets are registered, the recognition method based on sparse representation and dictionary learning is feasible and effective, and the operation is simple. The identification method directly utilizes the amplitude information of the target area and its surroundings in the image, without extracting other features of the target, nor considering the azimuth angle and other information of the target, and does not need to design other classifiers. It uses reconstruction error for identification.
- (2) Whether to perform the logarithmic transformation, median filtering, and registration processing on the image data has a certain influence on the recognition effect of the dynamic joint sparse representation model based on multi-information. The drawback is that all the investigated SAR image target recognition techniques are experimentally tested using slices that were found and retrieved from the MSTAR data. In actuality, the geography of the target geography and environment is complicated.

Therefore, the proposed method will be further validated when combined with SAR target detection in complex environments. The purpose is to construct a detection model employing theories of sparse representation and dictionary learning. The research aims at playing a role in the detection of SAR image targets in a certain manner.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This study did not receive any funding

## Data availability statement

Data will be provided upon request.

## References

- Damotharasamy, S. 2020. Approach to model human appearance based on sparse representation for human tracking in surveillance. *IET Image Processing* 14 (11):2383–94. doi:10.1049/iet-ipr.2018.5961.
- Guo, Y., and X. Feng. 2020. Analysis on balance improvement of cross grate cooler assembly production line based on MOD method. *IOP Conference Series: Earth and Environmental Science*, United kingdom, 440 (2).
- Hajipour, S., F. Namin, and R. Shirazi. 2021. A novel method for GPR imaging based on neural networks and dictionary learning. *Waves in Random and Complex Media* 33 (2):1–21. doi:10.1080/17455030.2021.1880667.
- Huang, G., Y. Xiao, and Z. Yin. 2020. Denoising method for underwater acoustic signals based on sparse decomposition. *Journal of Physics Conference Series*, United Kingdom, 1550 (032139).
- Jabs, D., N. Acharya, and A. K. Denniston. 2021. Classification Criteria for Sarcoidosis-associated Uveitis. *American Journal of Ophthalmology* 228 (2):11–13.
- Luo, K., X. Liu, J. Li, Y. Ma, Q. Ye, J. Bai, C. Liang, and F. Zou. 2021. Redundant Gaussian dictionary in compressed sensing for ambulatory photoplethysmography monitoring. *Biomedical Signal Processing and Control* 66 (5). doi:10.1016/j.bspc.2021.102479.
- Ma, Y., X. Jia, and Q. Hu. 2020. A new state recognition and prognosis method based on a sparse representation feature and the Hidden Semi-Markov model. *IEEE Access* 99:1. doi:10.1109/ACCESS.2020.3019810.
- Peng, G. 2020. Joint and direct optimization for dictionary learning in convolutional sparse representation. *IEEE Transactions on Neural Networks and Learning Systems* 31 (2):559–73. doi:10.1109/TNNLS.2019.2906074.
- Shi, D., B. Lam, W. Gan, and S. Wen. 2021. Block coordinate descent based algorithm for computational complexity reduction in multichannel active noise control system. *Mechanical Systems and Signal Processing* 151 (4):107346. doi:10.1016/j.ymsp.2020.107346.
- Wang, Y., Y. Liu, B. She, G. Hu, and S. Jin. 2021. Data-driven pre-stack AVO inversion method based on fast Orthogonal dictionary. *Journal of Petroleum Science and Engineering* 201 (11):108362. doi:10.1016/j.petrol.2021.108362.
- Wu, M., M. Yong, and F. Fan. 2020. Infrared and visible image fusion via joint convolutional sparse representation. *Journal of the Optical Society of America A* 37 (7):34–35. doi:10.1364/JOSAA.388447.

- Wu, X., X. Zhang, J. Mustard, J. Tarnas, H. Lin, and Y. Liu. 2021. Joint Hapke model and spatial adaptive sparse representation with iterative background purification for Martian serpentine detection. *Remote Sensing* 13 (3):3. doi:10.3390/rs13030500.
- Xing, Z., C. Yi, J. Lin, and Q. Zhou. 2021. Multi-component fault diagnosis of wheelset-bearing using shift-invariant impulsive dictionary matching pursuit and sparrow search algorithm. *Measurement* 178 (4):109375. doi:10.1016/j.measurement.2021.109375.
- Xu, Z. 2021. Research on software credibility algorithm based on deep convolutional sparse coding. *MATEC Web of Conferences*, China, 336 (6).
- Xue, S., C. Yin, Y. Su, Y. -H. Liu, Y. Wang, C. -H. Liu, B. Xiong, and H. -F. Sun. 2020. Airborne electromagnetic data denoising based on dictionary learning. *Applied Geophysics* 17 (2):306–13. doi:10.1007/s11770-020-0810-1.
- Yang, T., L. Tang, and Q. Tang. 2021. Sparse angle CT reconstruction with weighted dictionary learning algorithm based on adaptive group-sparsity regularization. *Journal of X-Ray Science and Technology* 29 (3):1–18. doi:10.3233/XST-200735.
- Yan, H., Y. Wang, Y. Wang, and Y. G. Zhou. 2020. Electrical Capacitance Tomography image reconstruction by improved Orthogonal matching pursuit algorithm. *IET Science, Measurement & Technology* 14 (3):367–75. doi:10.1049/iet-smt.2019.0255.
- Yi, T., and X. Zhao. 2020. Propagation dynamics for monotone evolution systems without spatial translation invariance. *Journal of Functional Analysis* 279 (10):108722. doi:10.1016/j.jfa.2020.108722.
- Zhang, S., and H. Li. 2021. The Energetic characteristics of surface atoms in Cu Clusters: size and site consideration by first-principle calculation. *Vacuum* 184 (109971):109971. doi:10.1016/j.vacuum.2020.109971.
- Zhou, Y., and L. Xu. 2020. Convergence of an alternating direction and projection method for sparse dictionary learning. *Journal of Physics Conference Series*, China, 1592 (012066).