Object recognition of real targets using modelled SAR images

D.A. Zherdev
Samara University, 34, Moskovskoye shosse, Samara, 443086, Russia Federation

Abstract. In this work the problem of recognition is studied using SAR images. The algorithm of recognition is based on the computation of conjugation indices with vectors of class. The support subspaces for each class are constructed by exception of the most and the less correlated vectors in a class. In the study we examine the ability of a significant feature vector size reduce that leads to recognition time decrease. The images of targets form the feature vectors that are transformed using pre-trained convolutional neural network (CNN).

1. Introduction
This research is a continuation of ideas and methods of work [1], in which the recognition method based on computation of conjugation indices was considered. The conjugation index is a function that uses a feature vector and matrix as the input values. The class matrix is formed using vectors of a class. In the output of this function there is a rate of similarity between inputs. In particular, we examine the support subspace formation algorithm. The support subspace presents the matrix that is a best in a case of recognition using the conjugation index. For this purpose, we used the algorithm of exhaustive search which requires many computations. The goal of this work is computational complexity reduce with the similar or higher recognition quality. It is obtained by a procedure of vector exception from classes and with previous feature extraction of 64x64 pixel images. We train the convolutional neural network on this dataset of images and in result of convolutions we receive the vector of features with 120 components. Therefore, the computational complexity on the training stage is raised, but on the test stage the algorithm executes faster besides without drop of recognition quality.

2. Modelling technology
The SAR images are more useful in the object recognition problem as a form of radar data in a case of a more detailed object representation. We developed the algorithm that based on the similar SAR modelling principles that described in [2-4]. The SAR image is formed in the process of phase history recording in the result of airborne radar movement on a specified trajectory with sequentially propagated and received radar impulses. Therefore, we improved the resolution along a track of synthesis, which is larger than a real antenna length with a less resolution. The SAR image is obtained using the received signals which stored in a phase and amplitude form. We can write the general expression for computation of a SAR image value in a pixel with the slant and cross range $x_s, y_s$ based on the presented structure of the general model of the radar observation:

---

1 E-mail: denis.zherdev.91@mail.ru.
\[ I(x_u, y_u) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v(x, y) \chi(x, y, x_u, y_u) \, dx \, dy. \]

The slant and cross range are coordinates of an obtained SAR image. Figure 1 shows the examples of modelled and real image. The model was formed in the similar condition of the real experiment, therefor the images can be matched.

![Figure 1. SAR images of (a) modelled and (b) real target.](image)

The program of a modelling has such parts as: map editor, 3D viewer with plug-in editor, relief generator, and common interface. There are two main stages to model: forming a relief that with some parameters and construction of a 3D object model which then is placed on this relief with specified properties. The 3D object is stored in a memory as a set of facets. The material properties are belonged to facets that reflect the radar impulses with different aspect angles.

The cost of real recognition experiments can be reduced by the using discussed modelling technology at the training stage. Moreover the results of recognition are improved by the fact that we can model specified object in the different conditions such as: ground properties, object reflection properties and etc.

3. Recognition and dimension reducing method

In this work we use the same method and algorithm that described in work [1]. There is a briefly present of the main part of method. The feature vector of size \( N \times 1 \) presents as:

\[ x_j = [x_{i1}, x_{i2}, ..., x_{iN}] \]

For a class can be constructed a matrix that consists of all vectors from this class:

\[ X_k = [x_1(k), x_2(k), ..., x_j(k), ..., x_M(k)], \quad k = 1, K \]

Then we can form a matrix \( Q_k \) of size \( N \times N \):

\[ Q_k = X_k [X_k^T X_k]^{-1} X_k^T, \quad k = 1, K, \]

for whole possible \( K \) classes.

Follow this notation we can make a decision and calculate the conjugation index:

\[ R_k(x_j) = \frac{x_j^T Q_k x_j}{x_j^T x_j}, \quad k = 1, K. \]

Then we searching the best subspaces that combined from class vectors and have higher recognition quality on the training set. Unfortunately, if the size or number of vectors are high then the algorithm will execute slowly. Thereby, in Section 4 we will present the technique of feature extraction and in this section we will discuss how to reduce the size of class matrix that constructed using training vectors.
We calculate correlations between vectors within a class in a goal to achieve the best separability of classes. The vectors with high and low correlation are excluded from the training set. The mean of this operation is that the new set without the most distant and typical vectors of class will be more effective in the recognition process. With the next step we train obtained set using algorithm of the support subspace construction.

4. CNN and feature vector size reducing

In the present time the CNN method is a fast grown method in the object recognition area. It finds many applications in the recognition of information in images [5,6]. One advantage of the CNN is in aggregating of most informative features in training set and representing it in more compact form. We are following the work [7] and explore the CNN advantage for achieving better recognition quality. Moreover, in the results of sequential convolutions the feature vector of less size is obtained. In a sense, it is similar to the fractal compression technique that authors used in [8] for reducing vector dimension and getting the less computational complexity of the support subspace method.

We implement CNN training to MSTAR database images. Figure 2 shows the general structure of the applied CNN. All images 128x128 were previously processed as described in [1] and finally have size of 64x64 pixel images. Our CNN implementation consists of two convolution layers, one max-pooling layer and one fully-connected layer.

The structure of layers is organized as follows. First layer is consists of twenty features with a kernel size of 13x13. The output of the layer contains the images of the size 52x52 pixels. Next layer is the downsampling with the factor of four. In the result on the third layer input twenty downsampled images of the size 13x13. The third layer proceed the convolution with 120 kernels size of 13x13 and in the output we obtain the vector 120x1. On the next stage of our work we use the fully-connected layer of 40 neurons. In the result on final output layer the class of current vector are estimated.

5. Results and discussion

In our experiments we used the MSTAR dataset with three objects: BMP2, BTR70, T72. The training and testing sets consists of 1618 and 1365 SAR images obtained under 17° and 15° of depression angle accordingly. The training stage include three main parts.

Firstly, we train the CNN that structure described in Section 4 using backpropagation algorithm. The weights and biases of network at initialisation generated randomly. The training dropout factor was 0.85. After training process we obtained the result 98.02% of correct recognitions on testing set. We saved the resulting weights and biases that necessary to reducing the size of feature vector. Then the new dataset of training and testing vectors with sizes of 120 components were formed using the convolutions of three first CNN layers described in Section 4.
Secondly, following the work [1] we clustered all classes using the conjugated index. The best result of recognition was obtained by clustering the BMP2 on 32, the BTR70 on 8 and T72 on 8 subclasses. We clustered the BMP2 class on the higher number because that leads to better separation of class on the recognition stage.

Finally, we trained our method of support subspaces on the datasets of new vectors. Moreover, we used the exclusion of high and low correlated vectors within a class that described in Section 3. Table 1 presents the confusion matrix of the recognition experiment. We reached the result of 98.15% that is higher than result was obtained by CNN. In addition experiment we obtained the result of 76.50% using modelled images on the training stage.

| Actual class | Estimated class | BMP2 | BTR70 | T72 |
|--------------|-----------------|------|-------|-----|
| BMP2         |                 | 572  | 5     | 5   |
| BTR70        |                 | 5    | 191   | 0   |
| T72          |                 | 10   | 0     | 577 |

In this work the recognition program made less errors in BMP2 class than that shown in [1]. It means that the clustering process of BMP2 on 32 subclasses have shown some influence. The obtained result of MSTAR database recognition using our method is a much close to result of the CNN in contrast of the work [7] where authors is kept close to 99% correct recognition rate. It can be explained by the fact that they used SVM with specified kernel function that has nonlinear transform of the feature space. In further works we would study the effect of nonlinear transformation of our subspaces. Moreover, it is a big interest in extremely extension of the training set with the modelled images with training using CNN.

Acknowledgments
The publication is realised with support of the Ministry of Education and Science of Russian Federation and RFBR project №17-29-03112.

References
[1] Fursov V., Zherdev D., Kazanskiy N. 2016 *International Journal of Advanced Robotic Systems*. **13**(5)
[2] Franceschetti G., Iodice A., Perna S., Riccio D. 2006 *IEEE Transactions on Geoscience and Remote Sensing* **44**(10) pp. 2851-2860
[3] Franceschetti G., Guida R., Iodice A., Riccio D., Ruello G. 2004 *IEEE Transactions on Geoscience and Remote Sensing* **42** (11) pp. 2385-2396
[4] Xie H., An D., Huang X., Zhou Z. 2016 *IEEE Transactions on Geoscience and Remote Sensing* **54**(6) pp. 3360-3377
[5] Spitsyn V.G., Bolotova Yu.A., Phan N.H., Bui T.T.T. 2016 *Computer optics* **40**(2) pp. 249-257 (in Russian)
[6] Vizilter Y.V., Gorbatevich V.S., Vorotnikov A.V., Kostromov N.A. 2017 *Computer optics* **41**(2) pp. 254-265 (in Russian)
[7] Wagner S.A. 2016 *IEEE Transactions on Aerospace and Electronic Systems* **52**(6) pp. 2861-2872
[8] Minaev Y., Fursov V. 2016 *Proceedings of the second International conference Information Technology and Nanotechnology* pp. 379-385