IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 20, 2019, accepted February 23, 2019, date of publication February 28, 2019, date of current version March 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2902218

An Intelligent Sample Selection Method for Space-Time Adaptive Processing in Heterogeneous Environment

CHONGDI DUAN[1,2,3], YU LI[3], AND WEIWEI WANG[3]
[1]School of Information and Electronics, Beijing Institute of Technology, Beijing, China
[2]National Key Laboratory of Science and Technology on Space Microwave, China
[3]China Academy of Space Technology, Xi'an, China

Corresponding author: Yu Li (405416718@qq.com)

This work was supported by the National Natural Science Foundation of China, Xi'an Institute of Space Radio Technology, under Grant 61701395.

ABSTRACT In the heterogeneous interference and target-rich environment, the clutter-plus-noise covariance matrix (CCM) estimate with a good match for the cell under test is significant for space–time adaptive processing. In this paper, an intelligent sample selection method is proposed to estimate the CCM and, thus, to suppress the clutter in the heterogeneous background. First, the proposed method applies the sample amplitude and the spatial relationship between the sample and the antenna beam center to describe the characteristics of samples, which preliminarily selects the crucial homogeneous and heterogeneous samples based on the joint histogram property. Then, the generalized inner product value and the spatial included angle are used to work as two representative features for the training samples as well as the testing samples. Based on the feature vector, the multi-dimensional feature projection based on the kernel function is carried out to achieve the offline training classifier, and the testing samples are classified with intelligence in the multi-dimensional feature space by utilizing the trained support vector machine. Due to the utilization of space–time property in representative feature extraction and intelligent strategy for the decision-making process, the proposed method acquires more reliable and efficient sample selection results compared with most of conventional sample selection methods. The experimental results based on the simulated and measured data show the effectiveness of the proposed method.

INDEX TERMS Space-time adaptive processing (STAP), sample selection, clutter-plus-noise covariance matrix (CCM), heterogeneous environment, support vector machine (SVM).

I. INTRODUCTION

For space-based early warning radar working under down-looking mode, clutter suppression is important to meet the requirements of moving target detection. It is well-known that space-time adaptive processing (STAP) is a crucial tool to suppress the background clutter. However, the detection background [1], [2] is usually non-homogeneous and non-stationary because of the type-different and altitude-fluctuated backscattering scatterers [3], isolated interferences [4], and the spatial-dependent irradiation property of platform-moving radar systems, such as the range-dependent changes of the clutter space-time spectrum caused by the antenna gain and the yaw angle of the platform during the data collection [5]. The Clutter-plus-noise Covariance Matrix (CCM) estimated by the sample covariance matrix would exhibit mismatched clutter statistical properties compared with the cell under test (CUT), which results in sever performance degradation in clutter suppression. Thus, it is necessary to select the independent identical distribution (IID) training samples for the CUT to well suppress the heterogeneous background clutter.

In recent years, non-homogeneity detection (NHD) method is proposed to find the homogeneous and target-free samples for STAP in heterogeneous environment. In the literature [6], [7], generalized inner product (GIP) method is developed to improve the CCM estimation accuracy by eliminating the nonhomogeneous samples. However, this method

The associate editor coordinating the review of this manuscript and approving it for publication was Wenchao Meng.

cannot work properly for the extremely heterogeneous environment because the initial CCM estimated by using all the samples severely deviates from the true one due to the high ratio of the nonhomogeneous samples. To handle the problem, the sample iteration weighting method [8] calculates the sample weighting coefficient iteratively to achieve the matched CCM estimation for the GIP processing. Since the iterative times for the method depends on the initial CCM value in a large degree, sometimes the greatly unmatched initial CCM would lead to a high computation complexity and limit its application. Knowledge-aided (KA) methods, i.e. Bayesian method [9] and KA-GIP NHD method [10], require prior radar parameters and environment information to estimate a better CCM and thus to enhance clutter suppression performance. However, the prior information may become inaccurate or even unavailable when the environment is changed. Under the condition, their clutter suppression performance deteriorates severely. Intelligent sample selection methods [11], [12], in which the classifier is achieved offline by employing a large number of training samples based on unsupervised training strategy, are called as learning-based STAP (LBSTAP). This method utilizes multiple features to improve the estimation accuracy of CCM. However, the high dimension of the feature space inevitably increases the computational complexity and this case may not be suitable for real-time STAP processing.

To deal with the issue, an intelligent sample selection method for STAP in heterogeneous environment is proposed in this paper. Firstly, the training samples, including homogeneous samples and heterogeneous samples, are marked by employing the joint histogram property of amplitude and spatial included angle. Then, GIP value and the spatial included angle are extracted from these marked samples as the representative feature vector. Based on this, high-dimensional feature projection using the kernel function is carried out to achieve the offline training classifier, and the testing samples are classified in high-dimensional space by utilizing the trained support vector machine (SVM). Compared with current sample selection methods, the proposed method possesses the following characteristics. On one hand, the samples are classified with the trained SVM based on the probability property, which avoids the unreliability of artificial decision strategy employed in the conventional sample selection weighting methods. On the other hand, supervised learning method is adopted in view of the point that only representative features are required in the feature extraction process. Accordingly, poor real-time processing ability of unsupervised learning method due to high feature space dimension can be much relieved. The results on simulated and measured data show that the proposed method achieves a significant performance improvement for STAP in heterogeneous environment.

The rest of this paper is outlined as follows. The signal model of STAP is introduced in Section II. In Section III, the proposed sample selection method is depicted in detail, where clutter suppression performance and the computation complexity are analyzed. The experimental results based on simulated and measured data are shown in Section IV to verify the effectiveness of the proposed method. Finally, a brief conclusion is drawn in Section V.

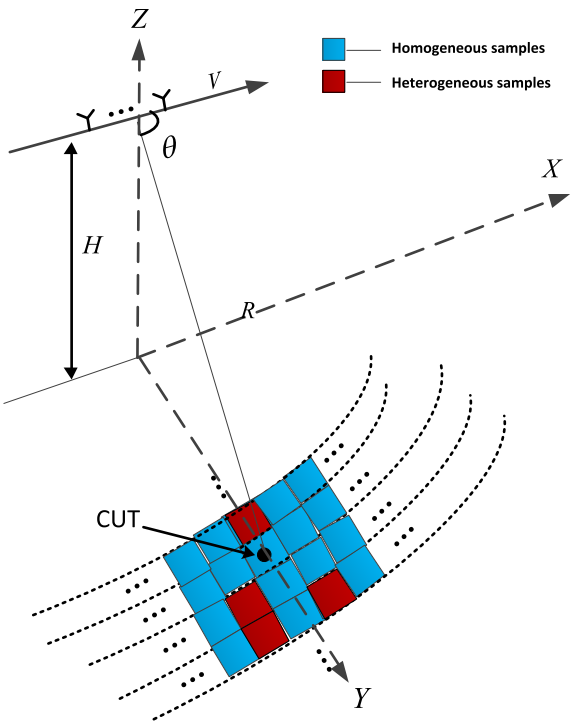


FIGURE 1. Geometry of side-looking array antenna.

II. STAP MODEL

The geometry of side-looking antenna array is illustrated in Figure 1. Assuming that the radar antenna is composed of N elements with an interval of d, the number of pulses during one coherent processing interval (CPI) is K. C_l denotes the scatter number in the l th range bin, where $1 \le l \le L$. Thus, the return data of the l th range bin can be expressed as

$$\begin{aligned} H_0 : \boldsymbol{x}_l &= \sum_{i=1}^{C_l} G_i \left(\boldsymbol{s}_t^i \otimes \boldsymbol{s}_s^i \right) + N_l \\ H_1 : \boldsymbol{x}_l &= \sum_{i=1}^{C_l} G_i \left(\boldsymbol{s}_t^i \otimes \boldsymbol{s}_s^i \right) + G_{tar} \left(\boldsymbol{s}_t^{tar} \otimes \boldsymbol{s}_s^{tar} \right) + N_l \end{aligned} \quad (1)$$

where H_0 and H_1 stand for the target-free and target-present cases, respectively, N_l represents the Gaussian white noise, G_i and G_{tar} denote the complex amplitude of the ith scatter and the target, respectively, symbol $\otimes$ denotes the Kronecker product operation, and $j = sqrt(-1)$. Let the normalized temporal steering vector and spatial steering vector for clutter are expressed as $\boldsymbol{s}_t^i$ and $\boldsymbol{s}_s^i$, respectively, while the normalized temporal steering vector and spatial steering vector of target are expressed as $\boldsymbol{s}_t^{tar}$ and $\boldsymbol{s}_s^{tar}$, respectively. The above-mentioned

vectors for targets are given by:

$$\begin{cases} \boldsymbol{s}_t^{tar} = \left[1, e^{j2\pi(\frac{2V}{\lambda \cdot f_r}\cos\theta_{tar}+\frac{2v}{\lambda \cdot f_r})}, \ldots, \right. \\ \quad \left. e^{j2\pi(\frac{2V}{\lambda \cdot f_r}\cos\theta_{tar}+\frac{2v}{\lambda \cdot f_r})(K-1)}\right]^T \\ \boldsymbol{s}_s^{tar} = \left[1, e^{j2\pi\frac{d}{\lambda}\cos\theta_{tar}}, \ldots, e^{j2\pi\frac{d}{\lambda}\cos\theta_{tar}(N-1)}\right]^T \end{cases} \tag{2}$$

where f_r indicates the pulse repetition frequency (PRF), V is the platform velocity, λ is the wave length of radar signal, θ_{tar} is the cone angle, v is the target velocity, and the subscript T denotes the transpose operation. The velocity of ground clutter is assumed to be 0 m/s, and then one can know that $\boldsymbol{s}_t^i$ and $\boldsymbol{s}_s^i$ of clutter are special cases of equation (2) with $v = 0$ m/s. When a target presents in the CUT, the target signal differences in the spatial steering vector and the temporal steering vector from clutter are the key factor for STAP processing to suppress clutter and remain the target.

Consider that the full-dimensional STAP method has a high computational complexity in the order of $O(NK)^3$ and requires $2NK$ homogeneous samples at least to achieve a maximum of 3dB performance deterioration theoretically. Hence, it is usually not suitable for real-time processing for most cases in the real world.

The dimension-reduced method is widely used to reduce the computational complexity of STAP [2]. Let $\boldsymbol{T} \in \mathbb{C}^{NM\times PQ}$ denotes the reduced-dimension matrix. P and Q stand for the degrees of freedom (DOFs) in the spatial domain and the temporal domain, respectively. The dimension-reduced data of the l th range bin can be rewritten as

$$\boldsymbol{x}_{Tl} = \boldsymbol{T}^H \boldsymbol{x}_l \tag{3}$$

where $(\cdot)^H$ is the conjugate transpose operation.

The dimension-reduced steering vector of the target can be re-represented as

$$\boldsymbol{s}_T = \boldsymbol{T}^H(\boldsymbol{s}_t^{tar} \otimes \boldsymbol{s}_s^{tar}) \tag{4}$$

By employing the linearly constrained minimum variance criterion, the optimal weight vector for dimension-reduced STAP is generally defined as:

$$\boldsymbol{w}_T = \frac{\boldsymbol{R}_T^{-1}\boldsymbol{s}_T}{\boldsymbol{s}_T^H \boldsymbol{R}_T^{-1}\boldsymbol{s}_T} \tag{5}$$

where $\boldsymbol{R}_T = \sum_l \boldsymbol{x}_{Tl}\boldsymbol{x}_{Tl}^H$ represents the dimension-reduced CCM.

III. INTELLIGENT SAMPLE SELECTION METHOD

Based on the dimension-reduced STAP model, an intelligent sample selection method is proposed in this section.

A. FEATURE EXTRACTION

1) DATA PREPROCESSING

During radar observation on a moving aircraft, the platform sometimes yaws in practice and the non-side-looking trajectory leads to the clutter temporal and spatial spectrum being range-dependent. In the point, the necessary compensation process is performed to ensure the clutter samples in different range bins with the IID property.

a: TEMPORAL COMPENSATION

The temporal domain compensation function of the ith $(1 \le i \le N)$ channel can be expressed as

$$\boldsymbol{x}_{l,i}^{'} = \boldsymbol{x}_{l,i} \cdot exp\left(-j2\pi\, DopplerCenter_{l,i} \cdot \boldsymbol{T}_{k,i}\right) \tag{6}$$

where $\boldsymbol{x}_{l,i}$ stands for the range-pulse domain signal of the ith channel in the l th range bin, the variable $\boldsymbol{T}_{k,i}$ is given by $\boldsymbol{T}_{k,i} = [0 : K-1]/f_r$, and the Doppler center of the l th range bin is expressed as:

$$DopplerCenter_{l,i} = \frac{\boldsymbol{angle}(\boldsymbol{mean}(\boldsymbol{x}_{l,i}(i,2{:}K)\cdot(\boldsymbol{x}_{l,i}(i,1{:}K\text{-}1))))}{2\pi f_r} \tag{7}$$

where ***angle*** $(\cdot)$ and ***mean*** $(\cdot)$ represent the angle calculation and averaging operation, respectively.

b: SPATIAL COMPENSATION

The spatial domain compensation function of the ith $(1 \le i \le N)$ channel can be indicated as:

$$\boldsymbol{x}_{l,i}^{''} = \boldsymbol{x}_{l,i}^{'} \cdot exp\left(-j2\pi\, W_{l,i}\right) \tag{8}$$

$$W_{l,i} = \frac{(i-1)d}{\lambda} \cdot cos(\theta_{azimuth}) sin\,(\theta_{down}) \tag{9}$$

where $\theta_{azimuth}$ is the azimuth angle, and θ_{down} is the downwards angle.

2) SAMPLE MARKING

Different types of samples can be trained with classifier on condition that the category information of the training samples has been obtained in advance. That is to say, the marking operation for training samples should be completed beforehand. Since a limited CPI is used for a quick detection of interested targets and different objects cannot be directly distinguished in the range-azimuth-compressed image with a low resolution, the sample marking methods in high-resolution images based on image gray level and image structure information becomes unsuitable. To address the issue, two relevant features, i.e., the amplitude distribution and the spatial included angle distribution, are created to describe the samples characteristics and then the sample marking operation is proposed.

a: AMPLITUDE DISTRIBUTION

For the pre-processed data in the range-Doppler domain, the amplitude probability of the l th $(1 \le l \le L)$ sample p_{amp}^l is computed by counting for the amplitude probability of L samples, and then normalize p_{amp}^l using $P_{amp}^l = \frac{p_{amp}^l}{\max(p_{amp}^l)}$. The samples account for a high proportion in the sample amplitude probability histogram is more possible to belong to the dominant scatter type in the observation scene. Thus, The large proportion is usually to be regarded as the homogeneous samples, while the samples in a small proportion are considered as the heterogeneous ones.

b: SPATIAL INCLUDED ANGLE DISTRIBUTION

With $\boldsymbol{x}_{l\_fft}$ indicating the l th pre-processed sample in range-Doppler domain, the spatial included angle between the CUT steering vector $\boldsymbol{s}_{T\_ff}$ and the l th sample steering vector $\boldsymbol{x}_{l\_fft}$ can be expressed as

$$\beta_l = acosd\left(\frac{\boldsymbol{x}_{l\_fft}^H \boldsymbol{s}_{T\_ff}}{\left|\boldsymbol{x}_{l\_fft}\right| \cdot \left|\boldsymbol{s}_{T\_ff}\right|}\right) \tag{10}$$

Then, the spatial included angle probability distribution is estimated with L samples, i.e., p_{angle}^l, and its normalized form for the l th sample is defined as $P_{angle}^l = \frac{p_{angle}^l}{\max(p_{angle}^l)}$.

For the L samples in the Doppler channel of f_{d0} in the range-Doppler domain, the temporal steering vector of the CUT is the same as that of the l th sample. Therefore, the spatial included angle of space-time steering vector are equivalent to that of the spatial steering vector. With data pre-processing operation, the steering vector of the l th clutter cell is given by

$$\boldsymbol{x}_{l\_fft}(f_{d0}) = \left[1, \exp\left(\mathrm{j}\frac{\pi d}{V} f_{d0}\right), \ldots, \exp\left(\mathrm{j}\frac{\pi\,(N-1)\,d}{V} f_{d0}\right)\right]^T \tag{11}$$

One can see that the interference phase between the CUT and the clutter cell in Doppler channel f_{d0} is irrelevant to the cell position and thus the spatial included angle β_l is identical for each clutter cell. In fact, the actual sample cell may also contain the extended target components of the CUT and other interference targets. Consequently, there will be a spatial included angle difference $\Delta\beta_l$ between homogeneous clutter samples and these heterogeneous interference samples. Based on this, different types of samples can be distinguished reasonably. Here, $\Delta\beta_l$ is related to the interference-to-clutter ratio. Considering that the interference phase would be impacted by the array error and signal to clutter plus noise ratio (SCNR) of the received echo, a slight expansion may occur in the spatial included angle distribution for L samples. The above sample selection method does not depend on the characteristics of the CUT, so it is applicable for different cases without regard to the existence of targets in the CUT.

c: MARKING SAMPLES WITH DIFFERENT TYPES

Based on the characteristics of the amplitude distribution and the spatial included angle distribution for the training samples, homogeneous samples and heterogeneous samples are marked in this section.

Let $P_{total}^l = P_{amp}^l + P_{angle}^l$, and P_{total}^l is sorted in descending order. In heterogeneous environment, the sample number L is subject to the condition $L \gg 2PQ$ in STAP processing. Thus the first $\min\{4PQ, L/2\}$ samples are regarded as the homogeneous samples, which are marked as $\omega = 1$. Meanwhile, the samples whose amplitude probability or spatial included angle probability belong to the minor peak lobes are regarded as heterogeneous samples, which are marked as $\omega = 2$. The remaining samples without distinctive characteristics are marked as $\omega = 3$. A graphic expression for the marking results is indicated in Figure 2.

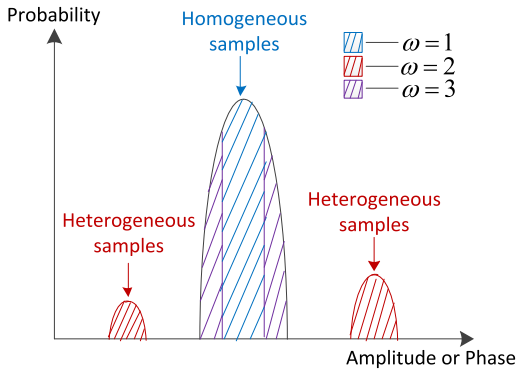


FIGURE 2. A graphic expression for the marking results.

3) FEATURE EXTRACTION

a: FEATURE EXTRACTION FOR THE TRAINING SAMPLES

The modified CCM is obtained using the homogeneous samples in the set $\{\omega = 1\}$ as

$$\boldsymbol{R}_{ref} = \sum_i \boldsymbol{x}_{Ti}\boldsymbol{x}_{Ti}^H,\ i \in \{\omega = 1\} \tag{12}$$

Then the GIP value of the l th sample is expressed as

$$y_l = E[\boldsymbol{x}_{Tl}^H \boldsymbol{R}_{ref}^{-1} \boldsymbol{x}_{Tl}] \tag{13}$$

Here, let y_l be the first feature and the spatial included angle β_l be the second feature. The feature space of the homogeneous samples is given by $\left\{\boldsymbol{y}_{training}^{\omega=1}, \boldsymbol{\beta}_{training}^{\omega=1}\right\}$. Similarly, the feature space for the heterogeneous samples is given by $\left\{\boldsymbol{y}_{training}^{\omega=2}, \boldsymbol{\beta}_{training}^{\omega=2}\right\}$.

b: FEATURE EXTRACTION FOR THE TESTING SAMPLES

Analogous to the above feature extraction method for the training samples, the feature space of the testing samples is written as $\left\{\boldsymbol{y}_{testing}, \boldsymbol{\beta}_{testing}\right\}$.

B. CLASSIFIER TRAINING

Assume the feature vector of the ith sample is represented as $\boldsymbol{d}_i (i \in \{\omega = 1 || \omega = 2\})$, Symbol || denotes the OR operation. Then a set of weight vector $\boldsymbol{w}$ and bias terms b of the samples which possess a minimum distance from the classification hyperplane are required to satisfy $\left|\boldsymbol{w}^T\boldsymbol{d}_i + b\right| = 1$. Herein, the distance is $2/\left\|\boldsymbol{w}\right\|_F^2$, where $\|\cdot\|_F$ stand for F norm operation. Thus, the max classification distance between the homogeneous samples and the heterogeneous samples is obtained based on classification hyperplane. The selection of the optimal classification hyperplane is equivalent to maximize the following objective function

$$\max_{\boldsymbol{w}}(\|\boldsymbol{w}\|_F^2/2 - \sum_{i=1}^{L} \alpha_i[r_i(\boldsymbol{w}^T\boldsymbol{d}_i + b) - 1]) \tag{14}$$

where α_i stands for the Lagrange coefficient, and $\alpha_i \geq 0$, and $r_i \in \{1, 2\}$ stand for the category of the ith sample.

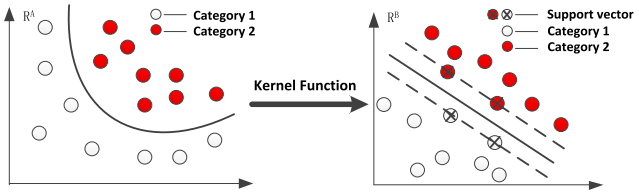


FIGURE 3. Non-linear mapping based on kernel function.

From equation (14), the non-linear classification problem is transformed into a linear classification problem by using kernel function $\boldsymbol{K}(\boldsymbol{d}_i, \boldsymbol{d})$ [13]. Figure 3 illustrates the above process. Samples with different types in the original feature space $\boldsymbol{R}^{\boldsymbol{A}}$ cannot be divided directly, while these two types of samples can be separated linearly after mapping the original feature space into high-dimensional feature space $\boldsymbol{R}^{\boldsymbol{B}}$. Thus, the homogeneous and heterogeneous samples are trained offline based on SVM by means of the mapped feature space, respectively.

The computational complexity of non-linear mapping depends on the number of samples which are closest to the classification hyperplane, namely support vector (SV), which are irrelevant to the sample number and the dimension of the feature. In this sense, the dimension disaster can be avoided skillfully for the convenience of real-time processing in the proposed method.

The standard kernel functions include four types, i.e. linear kernel function, polynomial kernel function, radial basic function (Rbf) and sigmoid kernel function. Here, Rbf is selected to handle non-linear classification problems as a result of its simple parameter setting mode and robust classification performance. Rbf is described as

$$\boldsymbol{K}(\boldsymbol{d}_i, \boldsymbol{d}) = \exp\left\{-\frac{|\boldsymbol{d}_i - \boldsymbol{d}|^2}{\eta}\right\} \tag{15}$$

where η stands for the curve variance of the kernel function.

Equation (14) can be rewritten as

$$\max_{\boldsymbol{w}}(\|\boldsymbol{w}\|_F^2/2 - \sum_{i=1}^{L}\alpha_i[r_i(\boldsymbol{w}\boldsymbol{K}(\boldsymbol{d}_i, \boldsymbol{d}) + b) - 1]) \tag{16}$$

where $|\boldsymbol{w}\boldsymbol{K}(\boldsymbol{d}_i, \boldsymbol{d}) + b| = 1$.

C. INTELLIGENT SAMPLE SELECTION

1) CLASSIFICATION OF THE TESTING SAMPLES

Assume the feature vector of the testing sample is $\boldsymbol{g}$, and the corresponding classification function is given by

$$Class(\boldsymbol{g}) = Sgn\left\{\sum_{i\in SV}\alpha_i r_i \boldsymbol{K}(\boldsymbol{g}_i, \boldsymbol{g}) + b\right\} \tag{17}$$

where α_i stands for Lagrange multiplier, b is the sample classification threshold, and Sgn represents the sign function. α_i and b are determined by the employed SVM.

Then, the testing samples are classified based on equation (17), and marked as $\omega = 1$ or $\omega = 2$ according to the sample types.

2) SPACE-TIME ADAPTIVE PROCESSING

The CCM is calculated with the selected homogeneous samples

$$\boldsymbol{R}_{ref}^{*} = \sum_{i}\boldsymbol{x}_{Ti}\boldsymbol{x}_{Ti}^{H},\ i \in \{\omega = 1\} \tag{18}$$

Based on the CCM, the adaptive weight $\boldsymbol{w}_T$ for STAP can be expressed as

$$\boldsymbol{w}_T = \frac{\boldsymbol{R}_{ref}^{*} - 1\boldsymbol{s}_T}{\boldsymbol{s}_T^H\boldsymbol{R}_{ref}^{*} - 1\boldsymbol{s}_T} \tag{19}$$

and clutter suppression is conducted to the testing samples.

Figure 4 shows the flowchart of the proposed method.

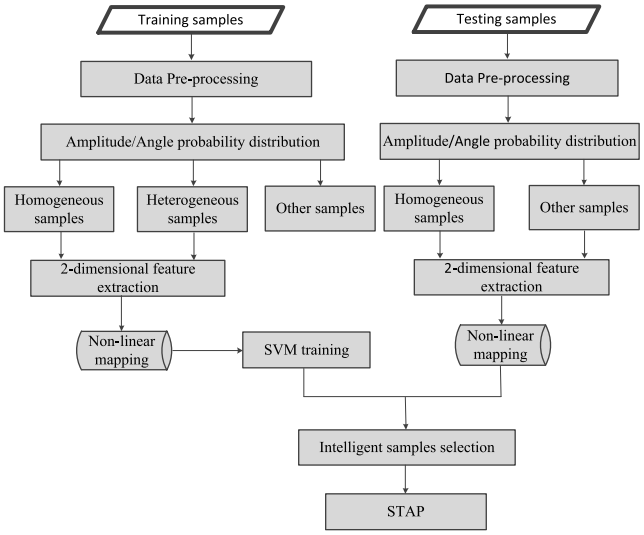


FIGURE 4. Flowchart of the proposed method.

D. COMPUTATIONAL COMPLEXITY ANALYSIS

The computational complexity of the proposed method is mainly composed of three parts, i.e. feature extraction of the training samples, feature extraction of the testing samples and SVM classification. For the first part, the computational complexity of GIP feature extraction is $O\left(\left[L_{select\_train} + PQ\right][PQ]^2\right)$, while that of the spatial include angle is negligible. Analogously, the computational complexity of the second part is $O\left(\left[L_{select\_test} + PQ\right][PQ]^2\right)$. As the third part, the computational complexity of SVM classification is $O\left(\left[Z^2 \cdot L \cdot N_{svm}\right]\right)$, where L_{select_train} and L_{select_test} stand for the training sample number and testing sample number in feature extraction procedure, respectively. Here, Z indicates the dimension of feature space, and it is set to 2 in this paper. L is the total number of training samples. N_{svm} is the SV number. Considering that the first part can be achieved offline, and the computational complexity of the

third part $Z^2 \cdot L \cdot N_{svm} \ll [L_{select\_test} + PQ][PQ]^2$. Thus, the total computational complexity can be approximated as $O([L_{select\_test} + PQ][PQ]^2)$.

Let L_{select} be the homogeneous samples selected from the testing samples. Then the computational complexity of GIP method, Spectral Similarity (SS) method [14], KA method [10], and iteration weighting (IW) method [15] can be expressed as $O([L + PQ][PQ]^2 + L_{select}[PQ]^2)$, $O([L + L_{select}][PQ]^2)$, $O(L[PQ]^2)$ and $O(I_{Iterative}[L+PQ][PQ]^2 + L[PQ]^2)$, respectively, where $I_{Iterative}$ represents the iteration time.

TABLE 1. Computational complexity of different methods.

Method	Computational complexity
GIP method	$O\left([L+PQ][PQ]^2 + L_{select}[PQ]^2\right)$
SS method	$O\left([L+L_{select}][PQ]^2\right)$
KA method	$O\left(L[PQ]^2\right)$
IW method	$O\left(I_{Iterative}[L+PQ][PQ]^2 + L[PQ]^2\right)$
The proposed method	$O\left([L_{select}+PQ][PQ]^2 + L_{select}[PQ]^2\right)$

Table 1 shows the computational complexity of different methods. Among the above five methods, the computational complexity of the proposed method is slightly less than that of GIP method. IW method has the highest computational complexity as a result of its iteration process, while SS method and KA method possess the least computational complexity.

TABLE 2. Simulation parameters.

Parameters	Values	Parameters	Values
Velocity of platform	100m·s^{-1}	Number of spatial channels	8
Bandwidth	20MHz	Number of range gates	201
Height of platform	3000m	Number of pulses per CPI	250
PRF	2500Hz		

IV. EXPERIMENTAL RESULTS

A. SIMULATION RESULTS

In this section, the effectiveness of the proposed method is verified based on simulation data. A L-band side-looking airborne radar array is considered. Simulation parameters are shown in Table 2. 11 targets are distributed in 201 different range gates, where the corresponding SNRs are 8 dB, 16 dB, 11 dB, 9 dB, 10 dB, 6 dB, 12 dB, 13 dB, 7 dB, 14 dB and 15 dB, respectively. Moreover, the normalized Doppler frequencies of these targets are uniformly distributed between 0.34 and 0.44. Here, the training samples and the testing samples employed in the proposed method are generated independently based on the above simulation parameters. STAP outputs with different methods are given in Figure 5 to 9, where a 60 dB Chebyshev temporal taper is applied to the Doppler domain. All range gates are employed in the following STAP processing. Comprehensively, the output SCNR of the proposed method with regard to 11 targets can be improved significantly compared with other methods. This is due to the heterogeneous samples can be effectively eliminated based on intelligent sample selection method.

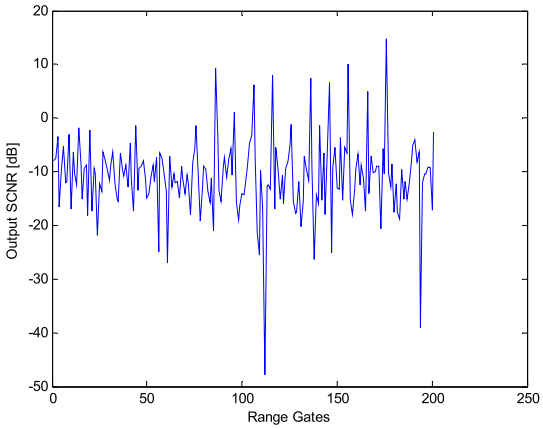


FIGURE 5. STAP output using GIP method.

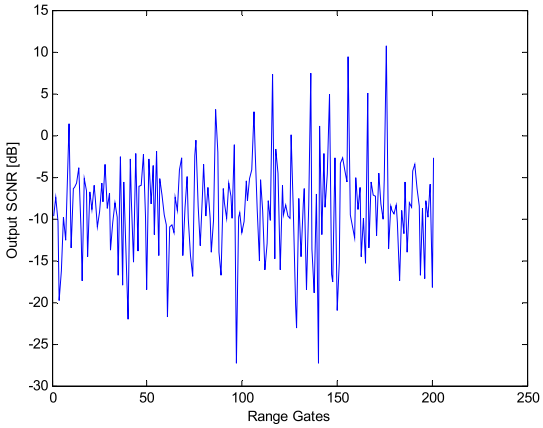


FIGURE 6. STAP output with SS method.

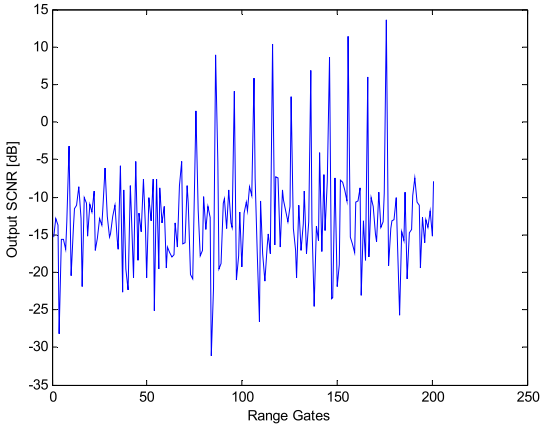


FIGURE 7. STAP output with KA method.

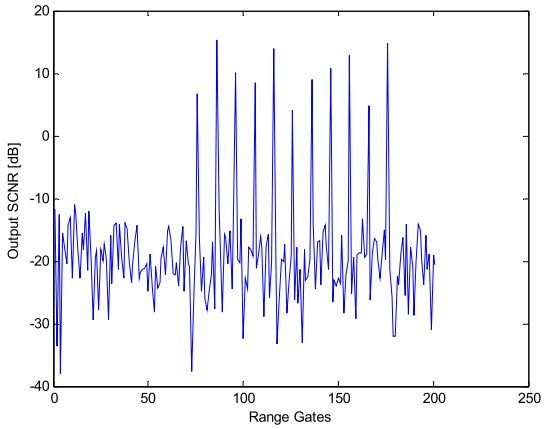


FIGURE 8. STAP output with IW method.

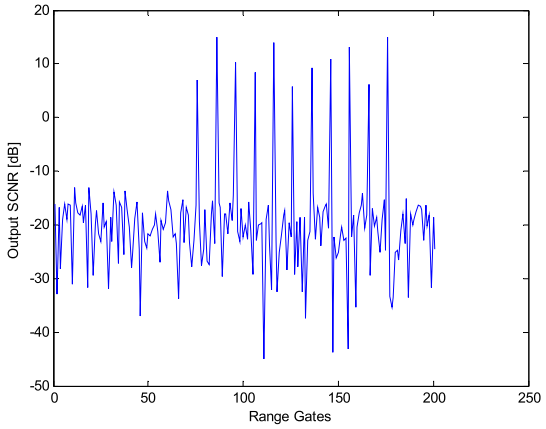


FIGURE 9. STAP output with the proposed method.

Thus, the problem of detection loss caused by CCM estimation deviation is alleviated by retaining the samples matched with the CUT.

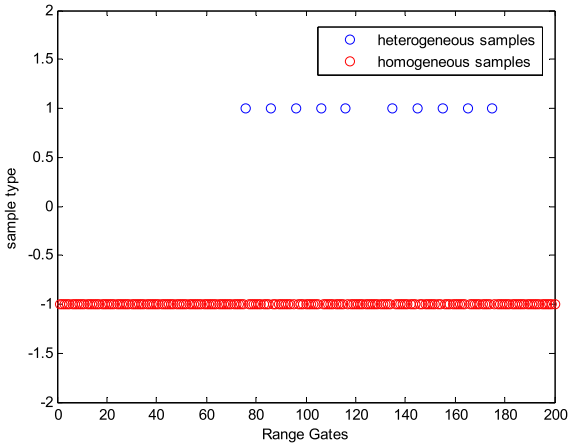


FIGURE 10. Sample types of different range gates.

Samples types of different range gates are given in Figure 10, one can see that the testing samples are classified exactly by means of the intelligent sample selection strategy.

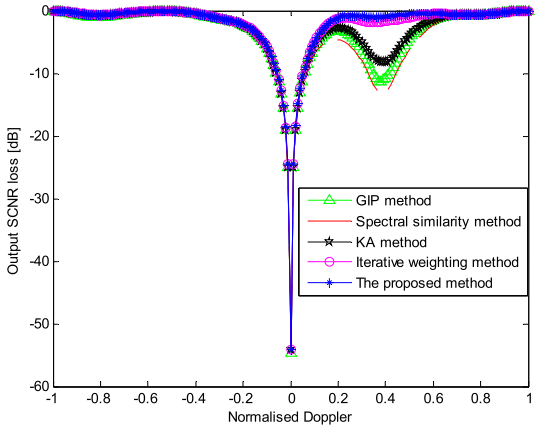


FIGURE 11. SCNR loss of target 6 with different methods.

Figure 11 takes target 6 as an example to further analyze the performance of the proposed method quantitatively, where the output SNR loss with the normalized Doppler frequency using different methods are compared. The output SNR loss of the proposed method and IW method are more than 8 dB higher than other methods, where the performance of the proposed method has a slightly improvement to IW method.

TABLE 3. Radar system parameters.

Parameters	Values	Parameters	**Values**
Number of spatial channels	8	Velocity of platform	100m·s^{-1}
Number of pulses per CPI	250	Bandwidth	20MHz
Number of range gates	1400	PRF	2500Hz
Height of platform	3000m		

B. REAL MEASURED DATA ANALYSIS

In this section, the effectiveness of the proposed method is validated by utilizing the L-band 8-channel real data, measured in eastern China with an airborne radar system. Radar system parameters are shown in Table 3. The received data include interferences, complex terrains and other non-ideal factors. In order to simulate the AMTI mode more realistically, a drone aircraft is designed to fly within the antenna's main lobe of carrier aircraft. Thus, the measured data is composed of real target echo, real clutter echo and real interference echo, with which the performance of different methods can be compared reliably. Based on multigroup collected data with different velocity cone angles relative to the along-track direction, the performance of the proposed method has been validated several times. Here, we take a group of data as an example to analyze the effectiveness of different methods.

Figure 12 shows the Range-Doppler data of channel 1 with 250 pulses. Figure 13 is the interferogram of channel 1 and channel 2. One can see that the interferences and discrete terrain clutters exist in most of the Range-Doppler domain. As a

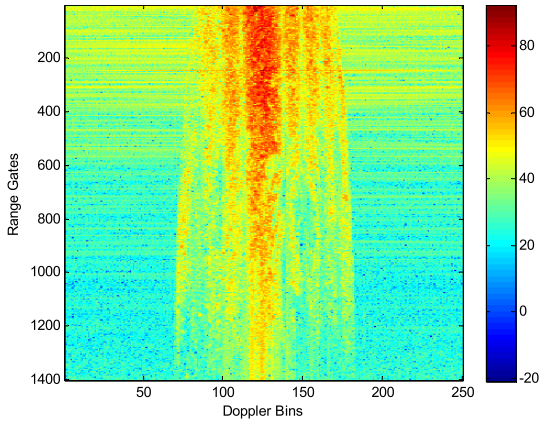


FIGURE 12. Range-Doppler data of channel 1 with 250 pulses.

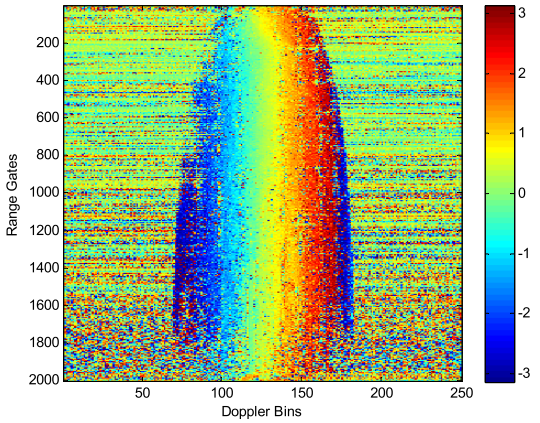


FIGURE 13. Interferogram of channel 1 and channel 2.

result, the subsequent processing is based on heterogeneous clutters and complicated electromagnetic factors.

The three-dimensional images of the GIP method, SS method, KA method, IW method and the proposed method after clutter suppression are given in Figure 14 to 18, where FSA [16] is selected as the dimension reduction technology. The temporal taper is consistent with that in Section 4.1.

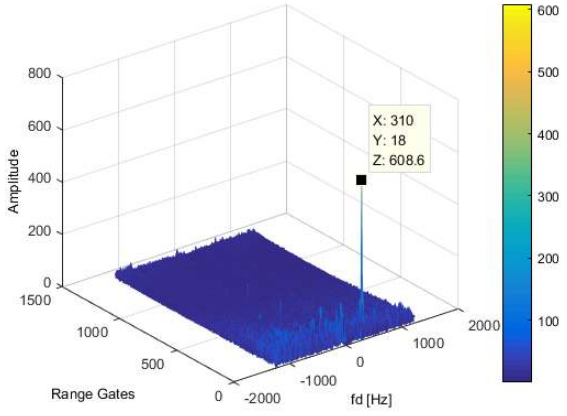


FIGURE 14. STAP output with GIP method.

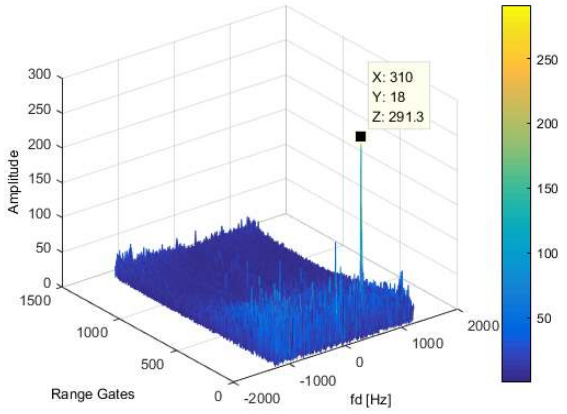


FIGURE 15. STAP output with SS method.

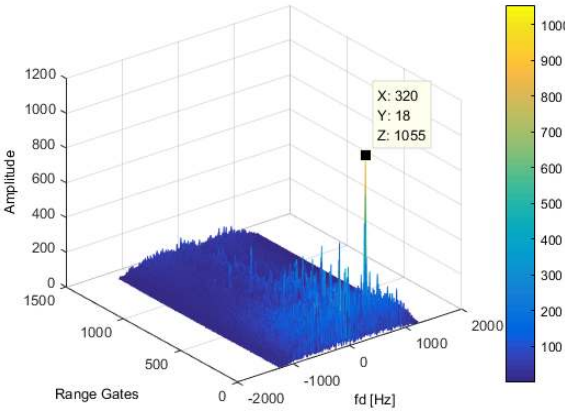


FIGURE 16. STAP output with KA method.

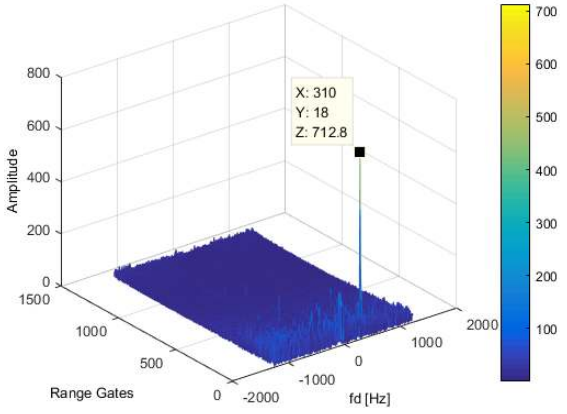


FIGURE 17. STAP output with IW method.

For GIP methods, high proportion of heterogeneous samples in the real measured data would cause the deviation of GIP value, which deteriorates the clutter suppression performance. SS method chooses the screening criterion for different types of samples based on the spectrum similarity between the testing samples and the CUT, which is suitable for targets with high SCNR. Nevertheless, the RCS of the drone aircraft is less than $1m^2$. Therefore, the target spectrum is submerged by the clutter spectrum because of low SNR, which results in the detection loss inevitably. For KA method,

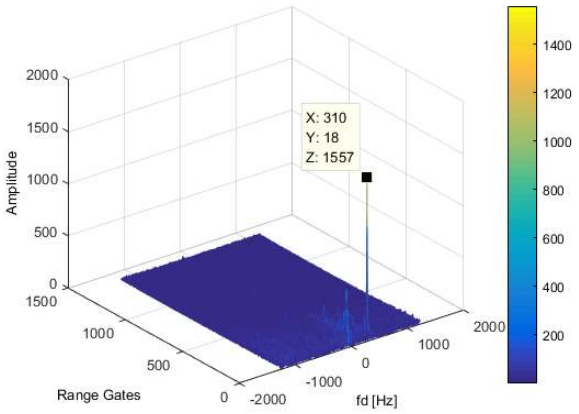


FIGURE 18. STAP output with the proposed method.

the clutter suppression performance depends on the accuracy of prior CCM. However, inaccurate prior information may deteriorate the performance, so it is generally applicable to complex radar systems in which robust prior information can be provided. In this section, the prior information is achieved by employing the previous frame data. IW method improves the output SCNR by eliminating the heterogeneous samples involved in the calculation of CCM iteratively based on the continually updated GIP value. The application of this method is mainly restricted by the high computational complexity caused by multiple iterations.

The experimental results based on real measured data show that the CCM obtained by the proposed method is better matched with the CUT. The clutters and interferences in the whole range Doppler plane can be suppressed to the noise level approximately (except for some discrete sidelobe clutter). Meanwhile, the output SNR of the drone aircraft located within the main clutter lobe is obviously improved based on our intelligent sample selection strategy, compared with other methods.

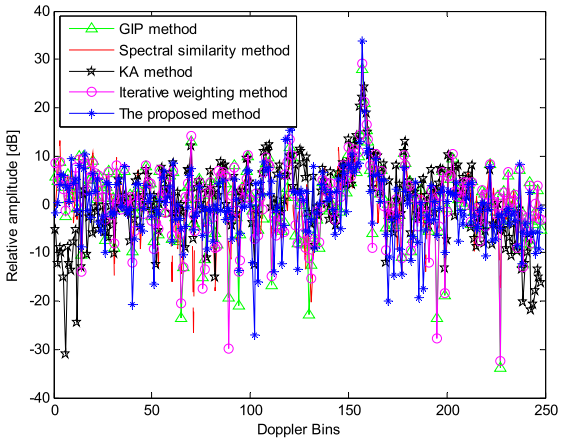


FIGURE 19. STAP output of the target drone with different methods.

To evaluate the performance of these methods in detail, slices of the drone aircraft along Doppler bins are compared in Figure 19. One can see that the relative amplitude curve of the proposed method decreases fastest than other methods. Moreover, the relative amplitude after clutter suppression is about 5-15dB more than that of other methods. Thus, the proposed method is more conducive for STAP.

V. CONCLUSION

An intelligent learning-based STAP method is proposed for heterogeneous environment in this paper. The proposed method has the following advantages over conventional sample selection methods. Firstly, the sample selection threshold of the proposed method is obtained through off-line training with a large amount of training samples and can be more robust in the varied environment compared with the methods adjusting the threshold manually. Secondly, the amplitude histogram and the spatial included angle histogram are introduced to describe sample characteristics on account of the space-time property, and the two features properly reflects the difference between the homogeneous samples and the heterogeneous samples in space-time domain. Thus, the STAP detection performance can be improved by eliminating the heterogeneous samples caused by interferences, non-ideal terrains and target expansion components. Moreover, the proposed method applies supervised learning rather than unsupervised learning to extract the representative sample features and reduce the redundant features, which alleviates the computational complexity problem in real-time STAP processing. Experimental results indicate that the proposed method improves both the accuracy of CCM estimation and the processing efficiency. Our future work will concentrate on exploring other representative feature space based on our existing research to obtain better sample classification performance in complex scenes.

ACKNOWLEDGMENT

The authors would like to thank the associate editor and anonymous reviewers for their insightful comments and suggestions.

REFERENCES

[1] J.-F. Degurse, S. Marcos, and L. Savy, "Subspace-based and single dataset methods for STAP in heterogeneous environments," in *Proc. IET Int. Conf. Radar Syst.*, 2013.

[2] W. L. Melvin, "A STAP overview," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 19, no. 1, pp. 19–35, Jan. 2004.

[3] O. Besson, S. Bidon, and J. Y. Tourneret, "Covariance matrix estimation with heterogeneous samples," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 909–920, Mar. 2008.

[4] T. K. Sarkar *et al.*, "A deterministic least-squares approach to space-time adaptive processing (STAP)," *IEEE Trans. Antennas Propag.*, vol. 49, no. 1, pp. 91–103, Jan. 2009.

[5] S. Bidon, O. Besson, and J. Y. Tourneret, "A Bayesian approach to adaptive detection in nonhomogeneous environments," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 205–217, Jan. 2008.

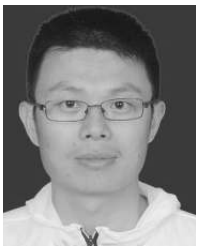
[6] X. Yang, Y. Liu, X. Hu, and T. Long, "Robust generalized inner products algorithm using prolate spheroidal wave functions," in *Proc. IEEE Radar Conf.*, May 2012, pp. 0581–0584.

[7] X. Yang, Y. Liu, and T. Long, "Robust non-homogeneity detection algorithm based on prolate spheroidal wave functions for space-time adaptive processing," *IET Radar, Sonar Navigat.*, vol. 7, no. 1, pp. 47–54, Jan. 2013.

[8] P. Stoica, J. Li, and J. Ling, "Missing data recovery via a nonparametric iterative adaptive approach," *IEEE Signal Process. Lett.*, vol. 16, no. 4, pp. 241–244, Apr. 2009.
[9] M. Barbary and P. Zong, "Novel anti-stealth on sub-Nyquist scattering wave deception jammer with stratospheric balloon-borne bistatic radar using KA-STAP-FTRAB algorithm," *IEEE Sensors J.*, vol. 15, no. 11, pp. 6437–6453, Nov. 2015.
[10] X. Zhu, J. Li, and P. Stoica, "Knowledge-aided space-time adaptive processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 1325–1336, Apr. 2011.
[11] A. El Khatib, K. Assaleh, and H. Mir, "Learning-based space-time adaptive processing," in *Proc. IEEE Int. Conf. Commun.*, 2013, pp. 1–4.
[12] A. E. Khatib, K. Assaleh, and H. Mir, "Space-time adaptive processing using pattern classification," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 766–779, Feb. 2015.
[13] Y. Liu, K. Wen, Q. Gao, X. Gao, and F. Nie, "SVM based multi-label learning with missing labels for image annotation," *Pattern Recognit.*, vol. 78, pp. 307–317, Jun. 2018.
[14] Y. F. Wu, W. Tong, W. Jianxin, and D. Jia, "Robust training samples selection algorithm based on spectral similarity for space–time adaptive processing in heterogeneous interference environments," *IET Radar Sonar, Navigat.*, vol. 9, no. 7, pp. 778–782, 2015.
[15] J. J. Guo *et al.*, "Iterative weighted covariance matrix estimation method for STAP based on generalized inner products," *J. Electron. Inf. Technol.*, vol. 36, no. 2, pp. 422–427, 2014.
[16] L. E. Brennan, D. J. Piwinski, and F. M. Staudaher, "Comparison of space-time processing approaches using experimental airborne radar data," in *Proc. IEEE Nat. Rec. Radar Conf.*, Apr. 1993, pp. 176–181.

CHONGDI DUAN received the B.S. degree from the Harbin Institute of Technology, in 1994, and the M.S. degree from the School of Electronic Engineering, Xidian University, in 2005. He is currently pursuing the Ph.D. degree with the School of Beijing Institute of Technology. He is currently a Professor with the National Key Laboratory of Science and Technology on Space Microwave, China Academy of Space Technology. His research interests include radar signal waveform design, adaptive and array signal processing, mobile target detection, and high-speed real-time signal processors.

YU LI received the B.Sc. degree from Northwest University, in 2011, and the M.Sc. and Ph.D. degrees from the Xi'an Institute of Space Radio Technology, in 2014 and 2018, respectively. He is currently with the Xi'an Institute of Space Radio Technology. His main research interests include moving target detection and tracking, and radar imaging.



WEIWEI WANG received the B.Sc., M.Sc., and Ph.D. degrees from Xidian University, in 2005, 2008, and 2013, respectively. He is currently a Senior Engineer with the Xi'an Institute of Space Radio Technology. His main research interest includes ground moving target indicator.