




Research Article

A novel reduced-dimensional dictionary grid screening strategy for SR-STAP



Ning Cui^{1,2}  · Zhongjun Yu^{1,2} · Kun Xing¹

Received: 21 August 2020 / Accepted: 19 December 2020 / Published online: 28 December 2020

© The Author(s) 2020 

Abstract

The traditional space-time adaptive processing (STAP) methods require large training samples to estimate the clutter covariance matrix. Sparse recovery STAP (SR-STAP) can get a highly accurate estimation in the case of insufficient samples and it alleviates the above problem of conventional STAP methods, however, the price of SR-STAP is that it involves a huge computational load in optimization, especially when the input data size is very big. This paper proposed a novel SR-STAP dictionary construction method that can significantly improve computational efficiency. We introduce the noise grids as the supplement of the spectrum-aid methods and design a second grids screening method based on the prior spectrum information. After the process of the proposed method, the dimension of the space-time dictionary has a sharp decrease compared to the over-complete dictionary, moreover, the sparsity of variable can be maintained better than the common spectrum-aid methods. The numerical experiments verify that the proposed method not only can significantly reduce the optimization time to far smaller than the SR-STAP but also more robust than spectrum-aid methods even in the small input size.

Keywords Airborne radar · Space-time adaptive processing · Sparse recovery · RD dictionary

1 Introduction

Space-time adaptive processing (STAP) is a mature technique which used in airborne radar for improving the signal-to-clutter-plus-noise ratio (SCNR) of slow-moving targets in severe clutter environments. It can adaptively construct two-dimensional space-time optimal filter to eliminate the ground clutter spreading [1]. However, it is unreal to achieve the optimal STAP in application due to the high computational load and unknown clutter-and-noise covariance matrix (CNCM) [2, 3].

To overcome the aforementioned challenges of optimal STAP, sample covariance matrix (SCM) was developed to replace practical CNCM. Reed et al. proved that the requirement of independent and identically distributed

(i.i.d.) training samples is at least twice degrees of freedom (DoFs) for the SCNR loss less than 3 dB [4]. Unfortunately, it is hard to obtain sufficient i.i.d. training samples during the short period of detection, especially in a heterogeneous environment [5]. Many algorithms have been designed to reduce sample limitation from different perspectives, which mainly consist of reduced-dimension (RD) and reduced-rank (RR). In RD-STAP, it reduces system DoFs by a linear transformation, which makes the number of training samples decrease to twice adaptive DoFs. Traditional RD-STAP algorithms include extended factored approach (EFA) [6], joint domain localized (JDL) [7], sum and difference beam-forming (SDBF) [8]. Among them, JDL-STAP and SDBF-STAP are sensitive to array error and they need to add additional channels to improve the robustness

✉ Ning Cui, cuning181@mails.ucas.ac.cn; Zhongjun Yu, speed_yzj@sohu.com; Kun Xing, xingkunmark@163.com | ¹Aerospace Information Research Institute, Chinese Academy of Science, Beijing, China. ²School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Science, Beijing, China.



which leads to the increasing of requirement for training samples, EFA is more robust in the presence of spatial errors but it suffers from performance loss in low pulse repetition frequency (LPRF). As for RR-STAP, it projects the echo onto a lower-dimensional subspace by data-dependent transformation, thus reducing the sample requirement to twice the clutter rank. Typical RR-STAP methods include principal component (PC) [9], cross-spectral metric (CSM) [10], and multistage Wiener filter (MSWF) [11]. Despite their success to reduce the sample requirement, the performance of these methods is still constrained by secondary data in practical application [12].

STAP combined with the sparse recovery (SR) technique has attracted great interest in recent years. SR-STAP methods utilize the intrinsic sparsity of the clutter in angle-Doppler space to recover the clutter with advanced SR algorithms, and it can achieve a perfect estimation of CNM under insufficient training samples condition, even in a single sample. Though the advantages of above SR-STAP methods have great potential, an unacceptable heavy computational burden is still required, which makes SR-STAP is unrealizable for application. Generally speaking, SR algorithms approximate the optimal solution by successive iterations, which causes a large time cost in repetitive operations, particularly in the region near to the ideal solution. Some studies have been devoted to reducing the computation complexity. Fast sparse Bayesian learning (SBL) turns sparse problem to maximum a posteriori estimation (MAP) and searches sparse coefficients according to Bayesian framework [13], it gets rid of the influence of regularization parameter. RD method transforms the global optimization problem to local optimization with RD techniques such as beam-space post-Doppler [14], which cuts down the dimension of variables and has a significant computational savings. Spectrum information can also be used to design a more suitable dictionary [15, 16], this way is easy to achieve and reduces iterating time significantly, however, a relatively small number of atoms may destroy the stability of algorithms [17, 18].

Consider the complexity limitation of hardware in practical application, the spectrum-aid method seems the most feasible way of the above fast methods, however, its performance and stability cannot be satisfied with the requirement of the application. This is because the sparsity of the residual grids after the spectrum-aid method is weaker than before. In this paper, the noise grids are introduced as additional information to improve the sparsity of the spectrum-aid methods and hence a novel RD dictionary is proposed, but different from the conventional overcomplete dictionary, only a few noise grids are used in the proposed method. It should be pointed out that our method is to focus on the improvement of the SR-STAP model but not on solving algorithms. Based on this core

idea, we specially design a second grid screening structure to make the best use of spectrum information and get rid of the useless grids as much as possible. In the proposed algorithm, the region segmentation criterion is redesigned and the whole angle-Doppler space can be separated into three different areas according to the contribution toward the clutter. Then three different selection strategies are independently performed in their areas to select suitable space-time grids. Numerical experiment results verify the performance of the proposed method.

The remainder of this paper is arranged as follows: Sect. 2 establishes the basic signal mathematical model, which mainly includes the system working environment and the signal components of radar echo. The conventional STAP is also introduced in this section. Section 3 describes the method of SR-STAP and Sect. 4 makes a analysis of the effect of clutter grids and noise grids for performance. Section 5 gives a detailed description of the proposed RD dictionary construction algorithm. In Sect. 6, the numerical experiments are executed to verify the performance of the RD dictionary. Finally, Sect. 7 is the conclusion of the paper.

2 Basic signal model and conventional STAP principle

Consider a phased-array pulsed Doppler (PD) radar is placed in a moving platform which keeps a constant velocity v flying along with x -axis in the height of H as shown in Fig. 1. PD radar transmits M pulse with fixed PRF f_r in each coherent processing interval (CPI), and the transmitted signal is a narrow-band modulating signal with the carrier

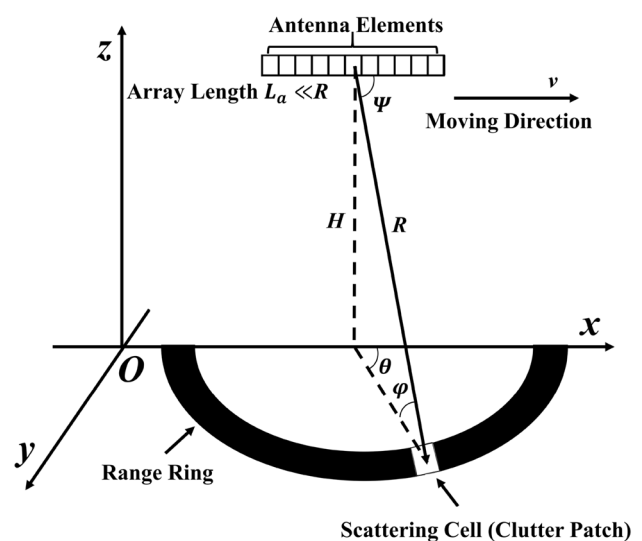


Fig. 1 The geometry of airborne radar illumination

frequency f_c . Assume that the structure of the antenna array is a uniform linear array (ULA) which includes N omnidirectional ideal elements, and these antenna units are uniformly distributed in the array with the interval d (usually it is equal to half wavelength). Here, our analysis focuses on side-looking mode which means crab angle $\theta_p = 0^\circ$, and the array length L_a is far smaller than the slant-range R which is the distance between the platform and scatter, hence all the elements can be viewed as having the same incidence direction with the reference element. Ψ , θ and ϕ are the cone angle, azimuth angle and elevation angle of the array relative to the ground unit, respectively.

Through the receiving process and matched filter, the multi-channel echo signal can be organized as a $N \times M \times L$ three-dimensional data matrix, where N , M and L are the number of spatial elements, slow beat (pulse number) and snapshot (range samples), respectively. For a given range cell, the target detection problem can be demonstrated as a binary hypothesis testing problem as follow

$$\begin{cases} H_1 : \mathbf{x} = \mathbf{x}_c + \mathbf{x}_t + \mathbf{n} \\ H_0 : \mathbf{x} = \mathbf{x}_c + \mathbf{n} \end{cases} \quad (1)$$

Hypothesis H_1 denotes the case of target presence and H_0 denotes the case of target absence. $\mathbf{x}_c \in \mathbb{C}^{NM \times 1}$ is the clutter component of the echo signal, and $\mathbf{x}_t \in \mathbb{C}^{NM \times 1}$ denotes the target signal component. \mathbf{n} is $NM \times 1$ white Gaussian noise vector. Note that the interference signal is not considered in Eq. (1) because it is not a key point in this paper. J.Ward clutter model [19] thought that the received signal could be expressed as the sum of reflection of small clutter patches when they are independent of each other. Based on this assumption, if no ambiguity appears in the current range cell, \mathbf{x}_c can be approximately written as

$$\mathbf{x}_c = \sum_{i=1}^{N_c} \gamma_i \mathbf{s}(f_i^s, f_i^d) \quad (2)$$

where N_c means the whole number of divided clutter patches. γ_i is the complex scatter coefficient of the i th clutter patch. $\mathbf{s}(f_i^s, f_i^d)$ represents the $NM \times 1$ complex space-time steering vector, where f_i^s and f_i^d are the spatial angle frequency and Doppler frequency, and can be denoted as $f_i^s = \frac{d}{\lambda} \cos \varphi \cos \theta_i$ and $f_i^d = \frac{2v}{\lambda f_r} \cos \varphi \cos(\theta_i + \theta_p)$, respectively. λ denotes the radar wavelength. Furthermore, $\mathbf{s}(f_i^s, f_i^d)$ can be written as

$$\begin{aligned} \mathbf{s}(f_i^s) &= [1, e^{2\pi f_i^s}, \dots, e^{2\pi n f_i^s}, \dots, e^{2\pi(N-1)f_i^s}]^T \\ \mathbf{s}(f_i^d) &= [1, e^{2\pi f_i^d}, \dots, e^{2\pi m f_i^d}, \dots, e^{2\pi(M-1)f_i^d}]^T \\ \mathbf{s}(f_i^s, f_i^d) &= \mathbf{s}(f_i^s) \otimes \mathbf{s}(f_i^d) \end{aligned} \quad (3)$$

where $[\cdot]^T$ is the matrix transposition and \otimes denotes the Kronecker product. n and m represent the index of

antenna element and pulse. The formation of \mathbf{x}_t is similar to the \mathbf{x}_c and can be written as

$$\mathbf{x}_t = \sum_{i=1}^{N_t} \xi_i \mathbf{s}(f_i^s, f_i^d) \quad (4)$$

where ξ_i means the complex amplitude for i th moving target in the current range cell. The only difference between (4) and (2) is that N_t and it means the moving targets count in CUT. Finally, the signal echo of H_1 is rewritten as

$$\mathbf{x} = \sum_{i=1}^{N_c} \gamma_i \mathbf{s}(f_i^s, f_i^d) + \sum_{i=1}^{N_t} \xi_i \mathbf{s}(f_i^s, f_i^d) + \mathbf{n} \quad (5)$$

and the $NM \times NM$ CNCM \mathbf{R} of the received signal in the current snapshot is

$$\mathbf{R} = \sum_{i=1}^{N_c} |\gamma_i|^2 \mathbf{s}(f_i^s, f_i^d) \mathbf{s}^H(f_i^s, f_i^d) + \sigma_n^2 \mathbf{I} \quad (6)$$

where $|\cdot|$ denotes the absolute value operation and $(\cdot)^H$ is the conjugate transpose operation. σ_n^2 represents the noise variance of the current test cell. \mathbf{I} is $NM \times NM$ identity matrix in which diagonal elements are one. It should be noted that the targets are not included in CNCM. For the conventional STAP methods, the adaptive filter weight can be designed according to

$$\mathbf{w} = \mu \mathbf{R}^{-1} \mathbf{s}_t \quad (7)$$

where $\mu = \frac{1}{\mathbf{s}_t^H \mathbf{R}^{-1} \mathbf{s}_t}$ represents the normalization factor of weight. \mathbf{R}^{-1} denotes the inverse of CNCM, which is mainly used to form the hollow of original clutter, $\mathbf{s}_t \in \mathbb{C}^{NM \times 1}$ represents the expected target direction in the angle-Doppler domain, here the spatial angle is the azimuth angle of illumination and the Doppler frequency is the tested Doppler channel. In practice, \mathbf{R} is unknown, and it is usually be replaced by its estimation $\hat{\mathbf{R}}$ which is calculated from the secondary samples near the CUT by

$$\hat{\mathbf{R}} = E\{\mathbf{X}\mathbf{X}^H\} \quad (8)$$

where $E\{\cdot\}$ is the expectation of range samples. Note that this estimation method supposes that training samples are satisfied with the condition of i.i.d, when in a non-ideal environment, it would suffer from performance loss. $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_l, \dots, \mathbf{x}_L]$ is sample data matrix, where \mathbf{x}_l is a complex vector of l -th range cell with NM elements.

3 SR-STAP method

Different from the traditional STAP method, in SR-STAP operations, the continuous angle-Doppler plane is discretized into $N_d N_s$ grid points (N_d points in Doppler axis, N_s points in angle axis), usually $N_s = \beta_s N$ and $N_d = \beta_d M$, where β_s and β_d are the coefficients of expansion in angle dimension and Doppler dimension. Note that grids number need to meet $N_d N_s \gg NM$ but it is not a strict condition. Then all the space-time steering vectors of grids can be used to construct the $N_d N_s \times L$ overcomplete dictionary set Φ and it is

$$\Phi = [\mathbf{s}_{1,1}, \dots, \mathbf{s}_{n_d, n_s}, \dots, \mathbf{s}_{N_d, N_s}] \tag{9}$$

where \mathbf{s}_{n_d, n_s} is space-time steering vector corresponding to the grid with k th spatial angle and j th Doppler frequency. Then echo Eq. (2) can be rewritten according to (9) as follow

$$\mathbf{x}_c = \Phi \boldsymbol{\gamma} \tag{10}$$

with

$$\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_{N_d N_s}]^T \tag{11}$$

where $\boldsymbol{\gamma}$ is the complex amplitude of grids in the angle-Doppler domain. The SR-STAP problem can be modeled as

$$\begin{aligned} \min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|_1 \\ \text{s.t. } \|\Phi \boldsymbol{\gamma} - \mathbf{x}\|_2 \leq \epsilon \end{aligned} \tag{12}$$

where $\min(\cdot)$ means to solve the minimum value of the object function. $\|\boldsymbol{\gamma}\|_1 = \sum_{i=1}^{N_d N_s} |\gamma_i|$ denotes the l_1 -norm which is used to describe the sparsity of recovery, and $\|\cdot\|_2$ is the l_2 -norm which measures the similarity between the recovery solution and original data. ϵ is noise threshold which is usually small value in optimization. One thing is worth pointing out that the original sparse recovery problem is l_0 -norm, here we take l_1 -norm relaxation [20]. Finally, the estimation $\hat{\mathbf{R}}_{sr}$ of SR-STAP can be easy calculated by

$$\hat{\mathbf{R}}_{sr} = \Phi^H \text{diag}(|\boldsymbol{\gamma}|^2) \Phi \tag{13}$$

where $\text{diag}(\cdot)$ is the diagonalization operation. The rest of the steps is the same as the conventional STAP filter design in Eq. (7).

4 The contribution of different type grids

The essence of (12) is to solve the underdetermined equation ($N_d N_s \gg NM$), and numerous solutions are existing in the SR-STAP problem. However, in the application, a

recessive restraint (clutter ridge is only occupying a little area relative to the whole space-time plane [21]) from the real environment will be added in raw data, which means that $\boldsymbol{\gamma}$ is high sparsity and can be recovered by SR theory. Overcomplete dictionary Φ plays a vital role in recovery, however observe Eq. (10), it should be known that the redundancy of Φ is significant increasing with the rise of β_s and β_d , which leads to the large computation. In fact, based on the prior knowledge of clutter ridge, we know that the most atoms of Φ has a little effect in performance, so whether there exists a way that reduce the number of atoms in Φ while keeping the sparsity of signal. Generally speaking, it needs add more grids in clutter area and less grids in noise area as in [16]. However when the noise grids reduce to zero [15], it will lead to a poor recovery due to the decreasing of sparsity. This illustrates that clutter atoms number represents the accuracy of recovery and noise atoms are used to stay the sparsity of signal. Therefore, a more concrete selecting principle can be given, to increase the accuracy of recovery, it needs to select more clutter grids, but it also needs to add more noise grids to keep the sparsity of signal. Following RD dictionary design is mainly based on the above principle.

5 RD dictionary construction strategy

As shown in Fig. 2, the proposed structure of RD dictionary consists of four parts including spectrum extracting, region segmentation, grids screening and RD dictionary construction. The prior knowledge which hides in the echo signal is extracted in the step one and then is used to exclude useless grids in step three. Finally, in step four, the RD transformation matrix which is formed by the saving grids. Following is the detailed description for each part.

5.1 Extracting spectrum information

Spectrum analysis offers a set of spectrum transformation methods to convert the time signal to frequency domain signal, such as Fourier transform (FT), Capon transform, Multiple Signal Classification (MUSIC) and Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [22]. In these algorithms, FT is the commonly used method in practical application due to the low-complexity. In the angle-Doppler domain, it is similar to the FT, the angle-Doppler transform can be achieved with the space-time steering vector basis as

$$P_k^j = |\mathbf{s}_{k,j}^H \tilde{\mathbf{R}} \mathbf{s}_{k,j}|, k = 1, \dots, N_s, j = 1, \dots, N_d \tag{14}$$

where $\tilde{\mathbf{R}} \in \mathbb{C}^{NM \times NM}$ is the estimation of the clutter covariance matrix with the current CUT and $\mathbf{s}_{k,j} \in \mathbb{C}^{NM \times 1}$ denotes the space-time steering vector basis which is

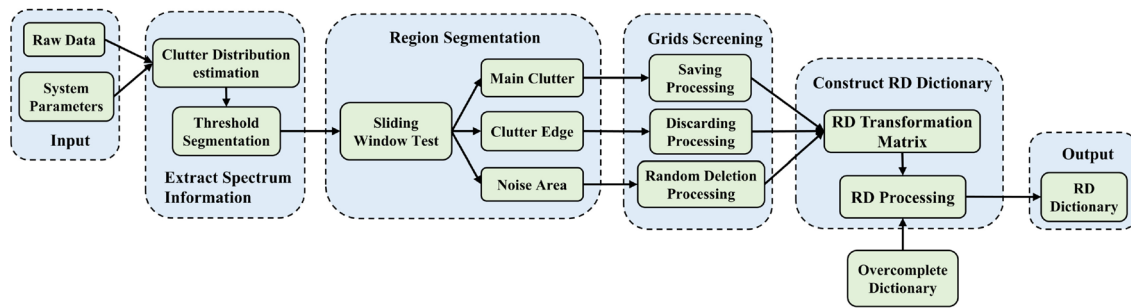


Fig. 2 The processing structure of screening strategy

corresponding to the k th and j th angle-Doppler grid. Different from the conventional FT method, Eq. (14) can calculate any grids in the angle-Doppler plane, however, its precision is low. Capon transform is more accurate than FT, which can be expressed as follow

$$P_k^j = \frac{1}{|\mathbf{s}_{k,j}^H \tilde{\mathbf{R}}^{-1} \mathbf{s}_{k,j}|}, k = 1, \dots, N_s, j = 1, \dots, N_d \quad (15)$$

The disadvantage of Capon method is that the big computation of the inverse operation, so it only suits to the situation which is small input data or weak real-time requirement. As for other high-accuracy algorithms, they all not suitable here due to the large running time.

Another alternative method is the least square (LS) algorithm which based on the optimal theory [23], and its spectrum estimate problem can be described as

$$\min \|\Phi \hat{\mathbf{P}} - \mathbf{x}\|_2 \quad (16)$$

where $\hat{\mathbf{P}}$ is the vectorization operation of power matrix \mathbf{P} . \mathbf{x} represents the raw echo data. It can be solved directly according to general expression without searching. Let the minimum value of object function (16) is equal to zero, and then the simplified solution is

$$\mathbf{P} = \text{mat}[(\Phi^T \Phi)^{-1} \Phi^T \mathbf{x}] \quad (17)$$

where $\text{mat}[\cdot]$ is the transformation operation, which means the $N_d N_s \times 1$ vector is converted to the $N_s \times N_d$ matrix. After that, the threshold segmentation method is used to extract spectrum clutter distribution [24]. Consider that the main aim in this step is only to acquire the rough area of clutter rather than the exact distribution, so the mean of angle-Doppler is enough to use, and it is written as

$$Th = \frac{1}{N_s N_d} \sum_{i=1}^{N_s} \sum_{j=1}^{N_d} P_{ij} \quad (18)$$

Some suggestions need to be mentioned. In a real application, it usually needs to process large data at the same

time. For accelerating the transformation, (14) can be solved by fast Fourier transform (FFT), the price of this is that it only can calculate the uniform spacing grids. The most time cost step of (15) and (17) is the inverse operation which can be calculated according to LU factorization. On the other hand, a more accurate threshold method only has a little improvement in final result but it would spend much time on computation, so a complex segmentation method is not be recommended here.

5.2 Region segmentation

In this section, the slide-window process structure is chosen to achieve the region segmentation. As shown in Fig. 3, four vertices of the window are selected as test grids which are used to detect the type of area. The angle-Doppler plane can be separated into three types according to the test grids. First is the main clutter area, it represents that all the grids in this window are clutter grids of spectrum distribution, on the contrary, the noise area denotes that all the grids are noise components. As for the clutter edge area, it lies in the boundary between clutter and noise, so this region both includes clutter and noise components. In processing, the test window slides successively in the whole plane, and check the type of current area by test grids. The second division aims to shrink the clutter and noise area, and isolate weak effect grids from above region.

5.3 Grids screening

For the main clutter region, it includes the most effective clutter grids, in fact, the real ground clutter points mainly locates in the clutter ridge (slash in side-looking mode, curve in non-side-looking mode) but it would spread into neighbor grids due to the grid mismatch problem. The slight change of these grids will cause a severe influence on recovery. Therefore it best to save all the space-time grids in this region.

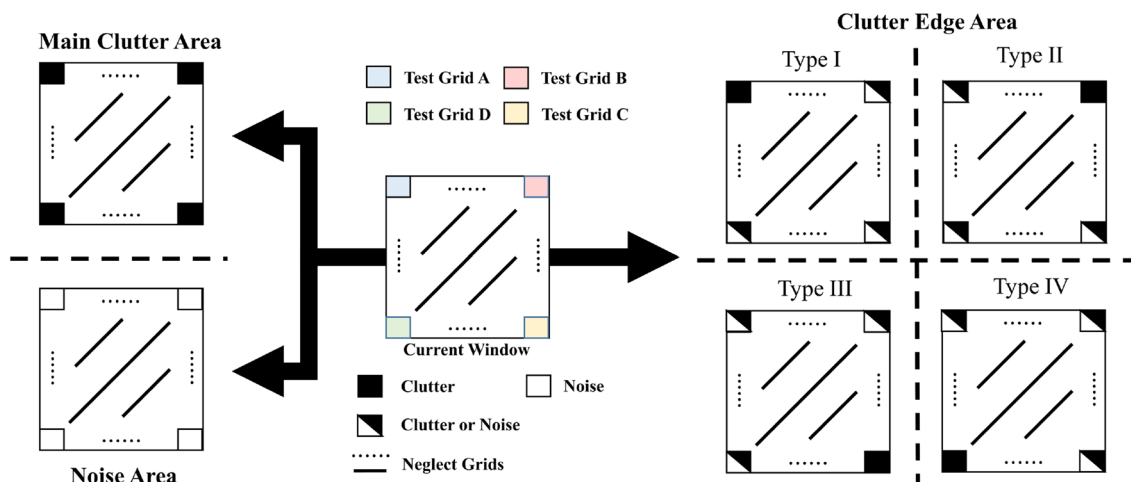


Fig. 3 The classification diagram of different areas

The clutter edge belongs to the junction of noise area and main clutter area, and it occupies a relatively small area to the whole angle-Doppler space. This means that only the marginal clutter grids are included in this area, i.e., its clutter power is lower than the main clutter power. Therefore the clutter edge has a weak performance improvement in the final accuracy of recovery. Above idea is established in the assumption that the number of edge grids is far smaller than the number of main clutter grids. However when this assumption is not satisfied in the application, the importance of edge grids would enhance with the increase of grids number and it cannot be ignored in this case. Fortunately, the number of the discretized grid is usually large in the most radar system. Therefore, to reduced the computation complexity, the clutter edge grids are directly discarded in the proposed algorithm, even this operation would have a little performance but it further improves the efficiency of the proposed method.

Noise area is the most important region in the screening strategy because its grids number is far bigger than other regions, this also means that the process method of it finally decides the size of the RD dictionary. Noise grids are employed as auxiliary grids to improve the sparsity of the main area but the nature of them is useless grids, hence it is reasonable to only utilize a few grids of noise area. Assume that the sliding-window is small enough compared to the angle-Doppler domain, and slow change occurs in neighbor grids, then the performance of the window can be approximately represented by part of grids in the window, it also can be known the locations of grids have no effect in final result according to the assumption. Here, we define the deletion ratio of grids, and it can be expressed as

$$\eta = \frac{n_0}{a^2} \tag{19}$$

where η is deletion ratio. n_0 represents the deletion grids in the noise area. When window length a is constant, the lower boundary of η is $\frac{1}{a^2}$ which means that only one grid is used in the window. Combined with the previous analysis, the random method is designed to delete grids. First is to generate an index of n_0 deletion grids with random number generation (RNG) algorithm, and then exclude grids by index. In application, it could take a fixed index to replace the random index, the advantage of it is no need additional algorithm complexity but it will require more storage memory.

5.4 RD transform

Through the screening process, the residual grids are the atoms of RD dictionary. It can directly construct RD dictionary by transformation matrix as follow

$$\tilde{\Phi} = \Phi T \tag{20}$$

where T is the $N_d N_s \times \hat{n}$ logical reduced-dimensional transform matrix and \hat{n} denotes the number of atoms of RD dictionary, for each column, the 1 only appears once in the index of saving grids (here each column only corresponds to one saving grids) and other locations are 0. However in practical application, the transformation matrix may be large due to the high-dimension of raw data. Here, we offer an easy way to solve this challenge, first is to prepare an additional matrix $\tilde{\Phi}$ with the size $N_d N_s \times \hat{n}$, then copy the column of the overcomplete dictionary to $\tilde{\Phi}$ one by one according to the index of remaining grids. The advantage

Table 1 Radar system parameters

Parameter	Value
f_0	450 MHz
λ	0.6662 m
f_r	300 Hz
d	0.3331 m
H	9 km
v	50 m/s
N	8
M	8
CNR	20 dB

of this method is that it no need to generate T and less computation than Eq. (20).

6 Numerical experiments

In this section, the numerical experiments are executed to examine the performance of the proposed method. The simulation data is generated by adapting the J.Ward clutter model [19] and some key radar system parameters are summarized in Table 1. In order to simulate the real illumination, we only consider the azimuth angle from -90° to 90° and set the number of clutter patch $N_c = 180$. In SR-STAP methods, the expansion factor β is 6 and the noise level ϵ is 10^{-6} . The number of i.i.d training samples is 2. For conciseness, the proposed method is distinguished from other SR-STAP methods by using the prefix RD, hence three methods are called RD-FT-SR-STAP, RD-LS-SR-STAP and RD-C-SR-STAP, respectively. Moreover, the deletion ratio η is fixed in $\frac{3}{4}$ and L_n is 2. CVX [25] is selected as the solving tool for SR-STAP in the following experiments. To evaluate the performance of SR-STAP, the signal-clutter-noise ratio (SCNR) loss is used as the criterion and it can be calculated by

$$SCNR_{Loss} = \frac{\sigma_n^2 |w^H s_t|^2}{MN(w^H R w)} \tag{21}$$

10 random trials for each algorithm are run to guarantee the accuracy of experiments.

6.1 The comparison of SCNR loss

In the first experiment, we verify the SCNR loss of the proposed methods. As shown in Fig. 4, the notch of C-SR-STAP and FT-SR-STAP are far wider than other methods. This is because the residual grids of the above methods are insufficient to recover the CNCM. As for LS-SR-STAP, it gets better performance than C-SR-STAP and FT-SR-STAP for the reason that it includes more grids due to the

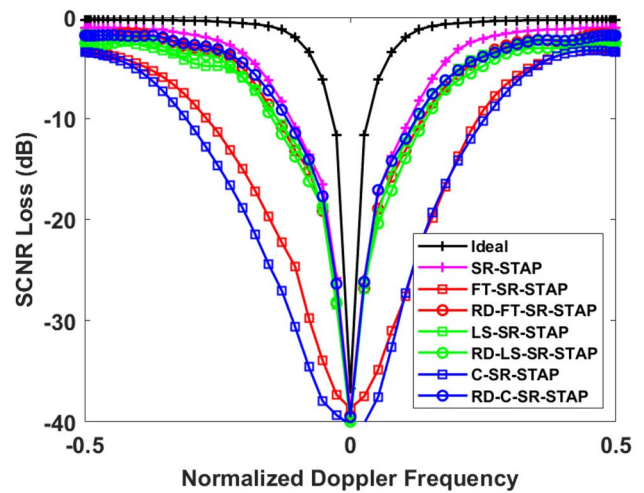


Fig. 4 The SCNR Loss of different dictionaries

low-resolution of the LS method. However compared with SR-STAP, it can be found that the above spectrum-aid SR-STAP methods have a severe degradation in performance. For the proposed methods, the SCNR loss of them only has a little bit down to the SR-STAP, but they still higher than the spectrum-aid methods. This result indicates the fact that the pure clutter grids cannot support the recovery of CNCM and utilizes the portion of noise grids as supplementary information can effectively improve the performance of spectrum-aid SR-STAP methods.

6.2 The effect of window size and detection ratio

In this experiment, we analyze the influence of window size and deletion ratio for the proposed method. As shown in Fig. 5, it is the SCNR loss of different window sizes. For three methods, their performance improve with the decreasing of the window size and they acquire the best result when L_n is equal to 2. However, when L_n increase to 5, they almost lost the ability due to the insufficient grids of the noise area. In these methods, the RD-LS-SR-STAP seems to have better performance than the other two methods but this is because the LS dictionary has already included more grids. As shown in Fig. 6, in order to evaluate the effect of the deletion ratio, we specially set $L_n = 3$ and take the deletion ratio $\eta = [\frac{5}{9}, \frac{6}{9}, \frac{7}{9}, \frac{8}{9}]$. There is almost no remarkable change for three RD dictionaries in small window size but it should be known that the effect of η will be intensified with the increase of L_n . Therefore, it can be found the window size is the main factor that decides the performance of the proposed methods. According to the above analysis, we recommend the combination of $L_n = 2$ and $\eta = \frac{3}{4}$ and this can cut down the most noise grids in the application.

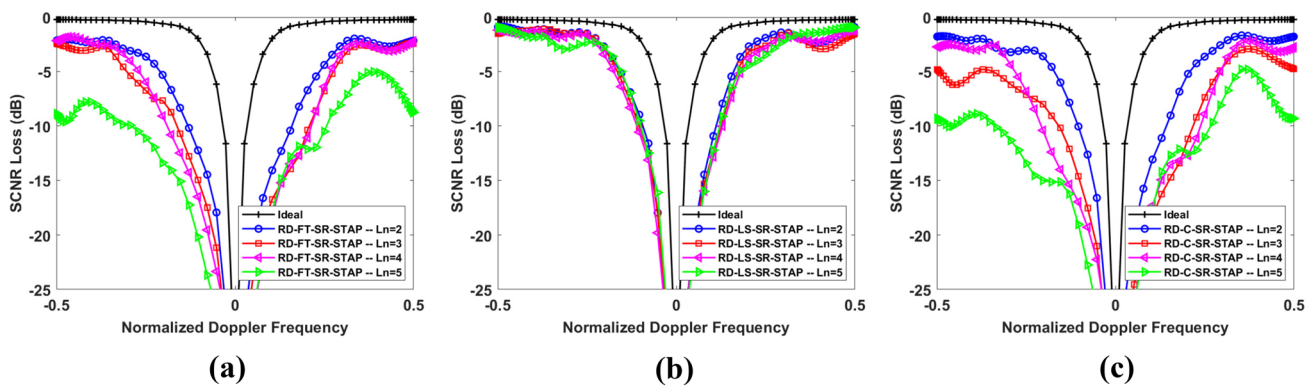


Fig. 5 The influence of different window size: **a** RD-FT-SR-STAP, **b** RD-LS-SR-STAP, **c** RD-C-SR-STAP

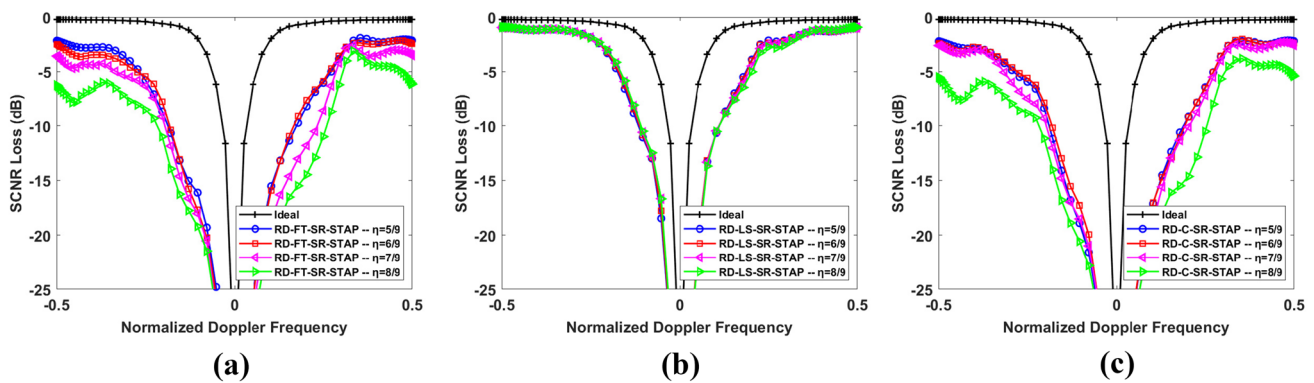


Fig. 6 The influence of different deletion ratio: **a** RD-FT-SR-STAP, **b** RD-LS-SR-STAP, **c** RD-C-SR-STAP

6.3 The effect of spatial error

In this section, we focus on the stability of the proposed methods in existing spatial error. The Guerci error model [3] is used to generate a spatial error matrix and it can be described as following

$$f_u(\delta_{\epsilon_{a,n}}) = \frac{1}{\Delta_{\epsilon_a}}, 0 \leq \delta_{\epsilon_{a,n}} \leq \Delta_{\epsilon_a}$$

$$f_u(\delta_{\epsilon_{p,n}}) = \frac{1}{\Delta_{\epsilon_p}}, -\frac{\Delta_{\epsilon_p}}{2} \leq \delta_{\epsilon_{p,n}} \leq \frac{\Delta_{\epsilon_p}}{2}$$
(22)

where $f_u(\delta_{\epsilon_{a,n}})$ and $f_u(\delta_{\epsilon_{p,n}})$ are the uniform distribution Probability Density Function (PDF) of amplitude error and phase error, respectively. In practical application, the gain error can be easily compensated by the preprocess, thereby this experiment only considers the case of phase error and the final spatial error can be expressed by

$$\Gamma = e^{j\delta_{\epsilon_{p,n}}} \tag{23}$$

As shown in Fig. 7, the phase error fluctuation $\Delta_{\epsilon_{p,n}}$ in here is set 5° . It is not surprise that all the SR-STAP methods

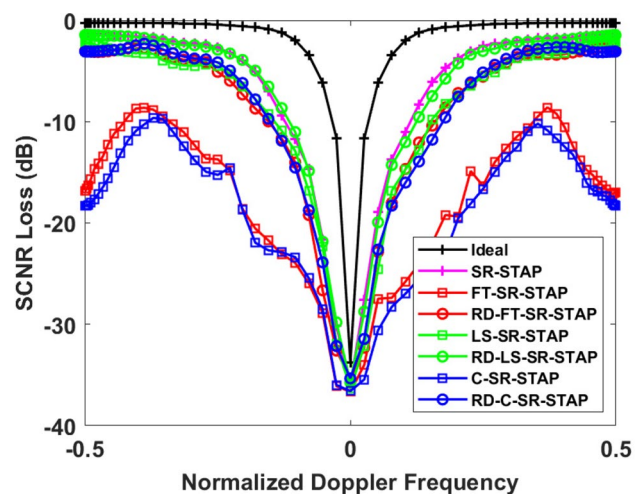


Fig. 7 The SCNR Loss of different SR-STAP methods in the case of spatial phase error

have a performance loss in the error case. In these methods, FT-SR-STAP and C-SR-STAP have a rapid decrease in performance, hence the spectrum-aid methods are instability to the error condition. In contrary to the above two

methods, the proposed methods only have a slight decrease compared to the no error case. This demonstrates that the proposed methods are more robust than spectrum-aid methods, moreover, it also suggests that the noise grids have a great influence in improving stability.

6.4 Real measurement data experiment

In order to verify the effectiveness of the RD dictionaries in real environment, we execute SR-STAP in the MCARM data [26]. In the 1990s, the Rome Lab and Northrop Grumman Company developed the MCARM plan, an L-band radar was installed in the BAC-111, and it recorded different ground types echo, these radar data can reveal the property of the real environment. The antenna array has 16 columns and each column has 8 elements, every four elements are organized into one channel. The index of using data is RL050575. The crab angle of the flight is $\psi = 7.28^\circ$ which could be regarded as work in the side-looking mode. The radar wavelength $\lambda = 0.2419$ m and PRF $f_r = 1984$ Hz. The distance of elements $d = 0.1092$ m. The speed of the plane $v \approx 100$ m/s. In the experiment, we select 8 channels to process and the pulses number $M = 12$. The index of the range is from 300 to 350. Consider the running time, the expansion factor $\beta = 6$ in here and the number of sample $L = 6$. The error limitation $\epsilon = 10^{-6}$. For comparing the result, we add one moving target in a strong clutter area by Moving Target Simulation (MTS), the range cells of it is 313. The normalized Doppler frequency of the target is 0.223. The power and spatial angle are -7 dB and 0° , respectively. The range-Doppler spectrum of

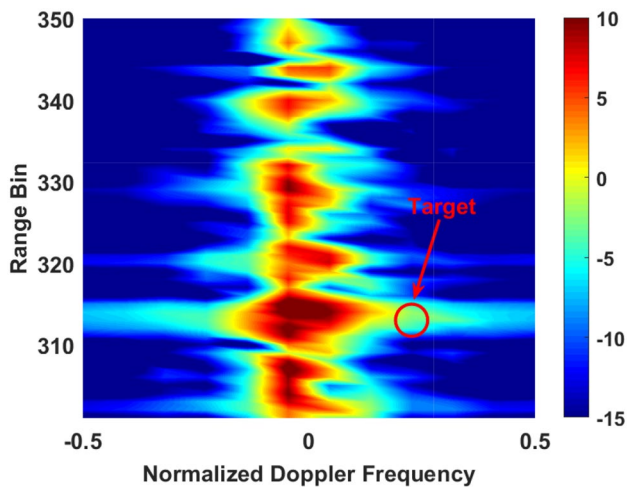
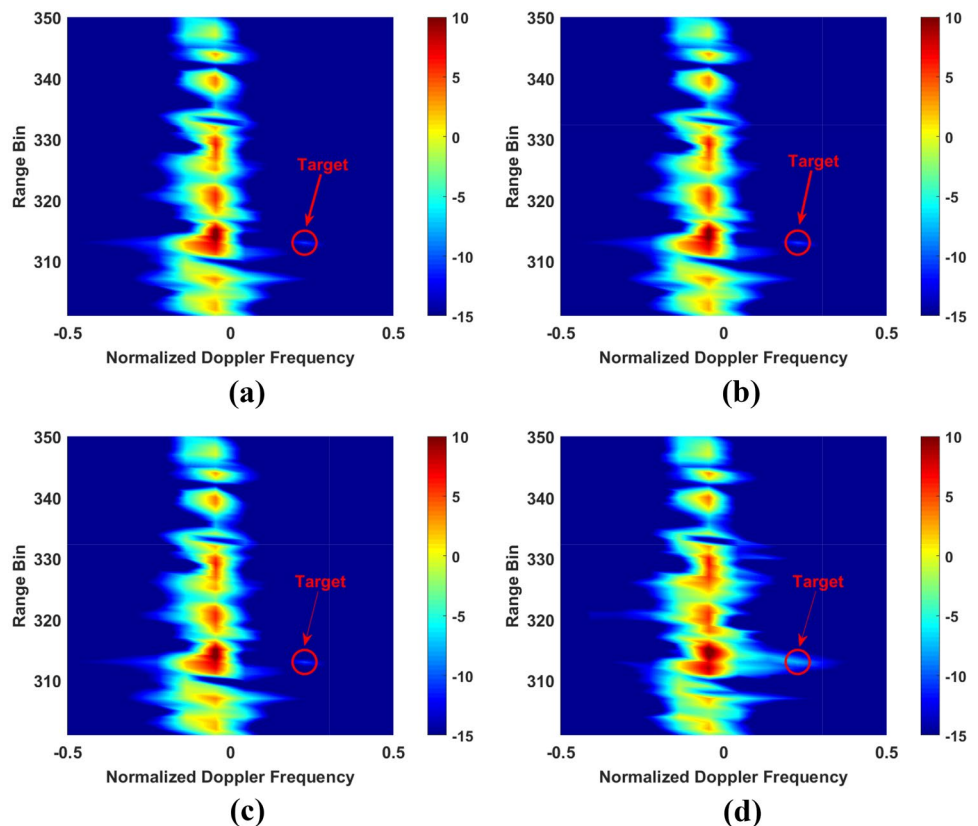


Fig. 8 The range-Doppler spectrum of raw data

Fig. 9 The RD spectrum of SR-STAP: **a** SR-STAP, **b** RD-FT-SR-STAP, **c** RD-LS-SR-STAP, **d** RD-C-SR-STAP



raw data as shown in Fig. 8, and the red circles represent the location of simulated moving targets.

The suppression of different SR-STAP methods in MCARM data is shown in Fig. 9. It should be pointed out that the spectrum-aid methods fail to recover the CCM in $\beta = 2$. The range-Doppler spectrum of SR-STAP is shown in Fig. 9a and the other three RD-SR-STAP methods are shown in Fig. 9b–d, respectively. It seems that there is no visible difference between the SR-STAP and the proposed methods, which illustrates that only adds a few noise grids to spectrum-aid methods can realize the performance of an overcomplete dictionary. Furthermore, we display the output power of the different range bin in normalized Doppler frequency 0.223 as shown in Fig. 10. It can be easy to find that the moving target emerges in the range cell 313 and the surrounded clutter is eliminated after STAP. For three methods, the RD-FT-SR-STAP and RD-LS-SR-STAP are better than RD-C-SR-STAP. This means that fewer grids cannot effectively recover the CCM even it is combined with noise grids. Therefore, the capon method is not suitable in application because of its computational load and poor suppression.

The computational complexity of the proposed methods is evaluated in the last experiment and the expansion factor β is used as the variable to change the size of input data. As shown in Fig. 11, the red bar areas represent the addition grid number of SR-STAP compared to the other methods. This means that the SR-STAP needs large grids to recover CCM perfectly. As for the proposed methods, the grid number only has a slight increase compared to spectrum-aid methods but far smaller than the SR-STAP. Next is the running time as shown in Fig. 12, and the symbol 'X' represents that a failed recovery in current β . The spectrum-aid methods often lost efficacy in the small input size, especially the C-SR-STAP cannot normally work in such conditions. Therefore, pure spectrum-aid methods lack practical significance. The proposed methods can still stay the effective recovery in all the input size, which means they are more robust than spectrum-aid methods in real data. The only price is a slight time extension, but the proposed methods can be easier to accept in using than SR-STAP and spectrum-aid methods.

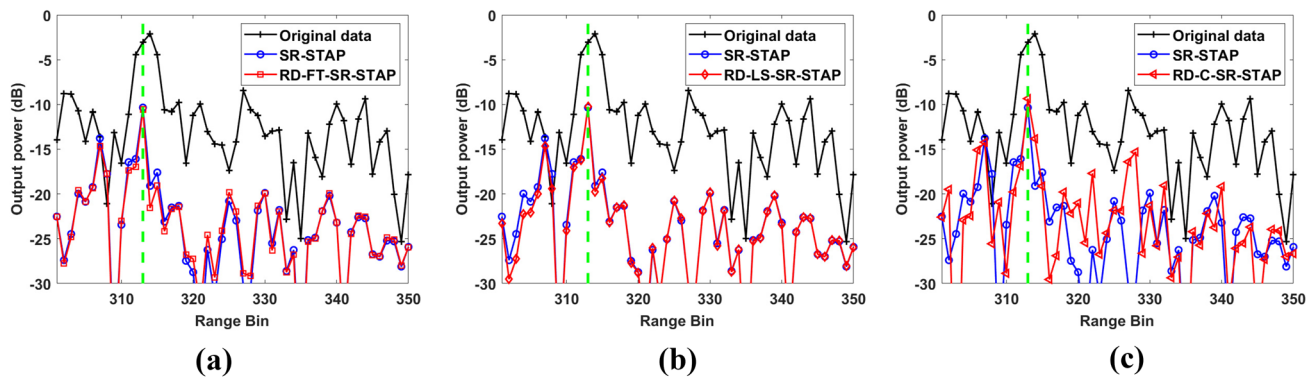


Fig. 10 The output power of defferent range bin: **a** SR-STAP, **b** RD-FT-SR-STAP, **c** RD-LS-SR-STAP, **d** RD-C-SR-STAP

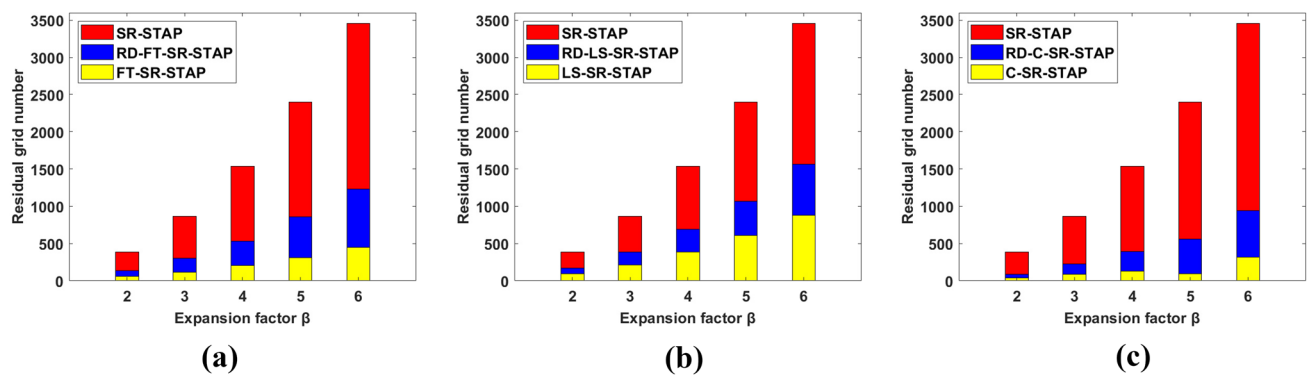


Fig. 11 The residual grid number for RD dictionary in different input size: **a** SR-STAP, **b** RD-FT-SR-STAP, **c** RD-LS-SR-STAP, **d** RD-C-SR-STAP

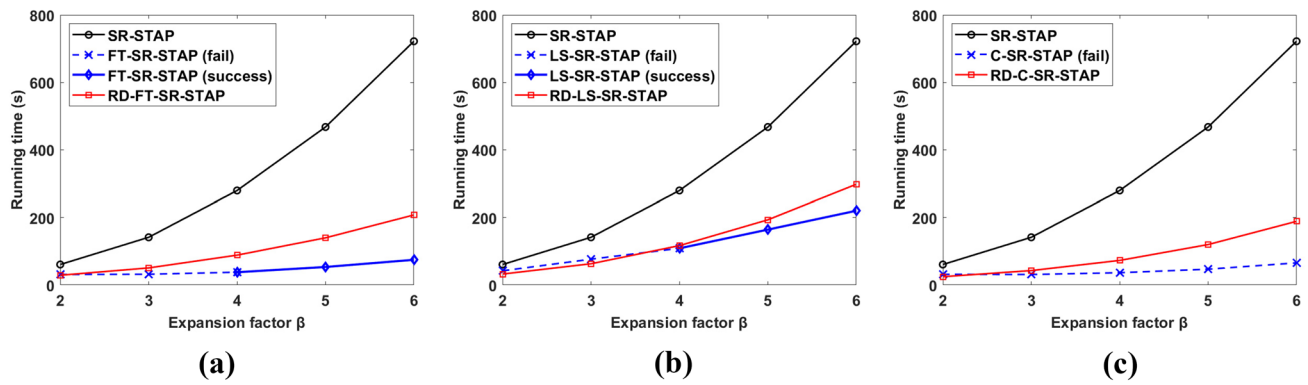


Fig. 12 The running time for RD dictionary in different input size: **a** SR-STAP, **b** RD-FT-SR-STAP, **c** RD-LS-SR-STAP, **d** RD-C-SR-STAP

7 Conclusion

For conventional SR-STAP methods, the complete dictionary leads to a huge computational load, and this limits the application of SR-STAP methods. The spectrum-aid methods have a great effect in reducing the computation of SR-STAP, but they are sensitive to the input data size. Except that, the spectrum-aid methods have poor performance because they discard lots of grids in recovery. In order to solve the above problems, the proposed methods first introduce the noise grids as the auxiliary information to enhance the sparsity of the recovery variable. Furthermore, a second grids screening method is developed to reject useless grids as much as possible. The exhaustive experiments demonstrate that it is feasible to get a similar recovery as the overcomplete dictionary by adding a few noise grids. Although we give three methods to extract spectrum information, the experiments show that the RD-FT-SR-STAP is more suitable than other two methods because of its balance between performance and computational load. In a word, the proposed method has a significant improvement and its computation is only a slight increase compared to the spectrum-aid methods, hence the proposed methods are more realistic significance. However, the proposed method only adopts the simple random sampling method to eliminate the useless noise grid points, hence a more appropriate screening scheme of noise grids is worth exploring. Furthermore, extending this method to other existing solving algorithms is also one of the main works in the future.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Data availability statement The MCARM data used to support the findings of this study are included within the supplementary information file(s).

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Brennan LE, Reed LS (1973) Theory of adaptive radar. *IEEE Aerosp Electron Syst Mag* 2:237–252
- Melvin WL (2004) A stap overview. *IEEE Aerosp Electron Syst Mag* 19(1):19–35
- Guerci JR (2014) Space-time adaptive processing for radar. Artech House, Norwood
- Reed IS, Mallett JD, Brennan LE (1974) Rapid convergence rate in adaptive arrays. *IEEE Trans Aerosp Electron Syst* 6:853–863
- Melvin WL (2000) Space-time adaptive radar performance in heterogeneous clutter. *IEEE Trans Aerosp Electron Syst* 36(2):621–633
- DiPietro RC (1992) Extended factored space-time processing for airborne radar systems. In: [1992] Conference record of the twenty-sixth asilomar conference on signals, systems & computers. IEEE, pp 425–430

7. Wang H, Cai L (1994) On adaptive spatial-temporal processing for airborne surveillance radar systems. *IEEE Trans Aerosp Electron Syst* 30(3):660–670
8. Brown RD, Schneible RA, Wicks MC, Wang H, Zhang Y (2000) STAP for clutter suppression with sum and difference beams. *IEEE Trans Aerosp Electron Syst* 36(2):634–646
9. Haimovich A, Berin M (1997) Eigenanalysis-based space-time adaptive radar: performance analysis. *IEEE Trans Aerosp Electron Syst* 33(4):1170–1179
10. Goldstein JS, Reed IS (1997) Reduced-rank adaptive filtering. *IEEE Trans Signal Process* 45(2):492–496
11. Goldstein JS, Reed IS, Scharf LL (1998) A multistage representation of the Wiener filter based on orthogonal projections. *IEEE Trans Inf Theory* 44(7):2943–2959
12. Zenghui Z, Wenchong X, Weidong H, Wenxian Y (2008) Local degrees of freedom of airborne array radar clutter for STAP. *IEEE Geosci Remote Sens Lett* 6(1):97–101
13. Wang Z, Xie W, Duan K, Wang Y (2017) Clutter suppression algorithm based on fast converging sparse Bayesian learning for airborne radar. *Signal Process* 130:159–168
14. Guo Y, Liao G, Feng W (2017) Sparse representation based algorithm for airborne radar in beam-space post-Doppler reduced-dimension space-time adaptive processing. *IEEE Access* 5:5896–5903
15. Han S, Fan C, Huang X (2016) A novel STAP based on spectrum-aided reduced-dimension clutter sparse recovery. *IEEE Geosci Remote Sens Lett* 14(2):213–217
16. Li S, Wei Y, Zhou X, Ding M (2019) A new sensing matrix construction strategy based on sparse recovery STAP. In: 2019 6th Asia-Pacific conference on synthetic aperture radar (APSAR), pp 1–6. IEEE
17. Mallat SG, Zhang Z (1993) Matching pursuits with time-frequency dictionaries. *IEEE Trans Signal Process* 41(12):3397–3415
18. Tsaig Y, Donoho DL (2006) Extensions of compressed sensing. *Signal Process* 86(3):549–571
19. Ward J (1998) Space-time adaptive processing for airborne radar. IEE colloquium on space-time adaptive processing (Ref. No. 1998/241), London, UK
20. Chen SS, Donoho DL, Saunders MA (2001) Atomic decomposition by basis pursuit. *SIAM Rev* 43(1):129–159
21. Sun K, Meng H, Wang Y, Wang X (2011) Direct data domain STAP using sparse representation of clutter spectrum. *Signal Process* 91(9):2222–2236
22. Stoica P, Moses R (2005) Spectral analysis of signals. Pearson Education, Upper Saddle River, NJ, USA
23. Boyd S, Boyd SP, Vandenberghe L (2004) Convex optimization. Cambridge University Press, Cambridge
24. Gonzalez RC, Woods RE, Eddins SL (2004) Digital image processing using MATLAB. Pearson Education India, Chennai
25. Grant M, Boyd S (2014) CVX: Matlab software for disciplined convex programming, version 2.1. In
26. Babu BS, Torres JA, Melvin WL (1996) Processing and evaluation of multichannel airborne radar measurements (MCARM) measured data. In: Proceedings of international symposium on phased array systems and technology. IEEE, pp 395–399

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.