Reduced-Complexity Subarray-Level Sparse Recovery STAP for Multichannel Airborne Radar WGMTI Application

NING, CUI
KUN, XING
ZHONGJUN, YU, Member, IEEE
Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China

KEQING, DUAN, Member, IEEE
School of Electronics and Communication Engineering, Sun Yat-sen University (SYSU), Guangzhou, China

***Abstract*—For multichannel airborne radars, wide-area ground-moving target indication (WGMTI) processing can quickly obtain the dynamic distribution of moving targets in a wide area, which holds considerable significance in many fields. Nevertheless, the WGMTI mode suffers from the interference of powerful ground clutter, which frequently submerges slow-moving targets and causes many false alarms in subsequent moving target detection. Space–time adaptive processing (STAP) can successfully suppress clutter, but its performance depends critically on the available training samples. Consequently, an effective STAP method characterized by fast processing and a small sample size for WGMTI application in multichannel airborne radars must be developed. In this paper, a subarray-level sparse recovery STAP (SR-STAP) processing framework is proposed for multichannel airborne radars. First, the characteristics of the subarray-level received clutter are discussed in detail. Second, on the basis of this analysis, we further designed a joint space–time dictionary and developed a separable tensor-based sparse Bayesian learning (STSBL) method. In this method, two-stage decomposition is proposed to ensure that large-scale data can be degraded into small-scale data in processing, which significantly improves computation efficiency. Finally, the effectiveness of the proposed STSBL-STAP method in WGMTI processing was verified using real measurement data obtained from a developed dual-channel Ku-band airborne radar.**

***Index Terms*— Airborne radar, sparse recovery space-time adaptive processing (SR-STAP), wide-area ground moving target indication (WGMTI), subarray.**

This work was supported in part by the National Natural Science Foundation of China under Grant 61871397. (*Corresponding author: Zhongjun Yu and Keqing Duan.*)

Ning Cui is with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100864, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100094, China (e-mail: cuining181@mails.ucas.ac.cn); Kun Xing is with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100864, China (e-mail: xingkun@aircas.ac.cn); Keqing Duan is with the School of Electronics and Communication Engineering, Sun Yat-sen University (SYSU), Guangzhou 510275, China (e-mail: duankeqing@aliyun.com); Zhongjun Yu is with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100864, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100094, China (e-mail: yuzj@ucas.ac.cn).

I. INTRODUCTION

As a hot topic in engineering application, synthetic aperture radar ground-moving target indication (SAR-GMTI) for airborne radars has received extensive attention globally. SAR-GMTI can simultaneously offer relevant geomorphic information and moving target locations to clients with limited observation. Thus, it is important in many fields, such as commercial remote sensing and traffic monitoring [1]–[3]. To quickly obtain the results of large-area SAR-GMTI, combined with the advanced electronic beam scanning capability of an active electronically scanned array (AESA), a wide-area GMTI (WGMTI) working mode has been proposed by scholars [1]. This proposed method can achieve fast wide-area moving target detection at the cost of low-resolution SAR imaging. WGMTI processing, influenced by the motion of the platform, suffers from strong ground clutter. Additionally, slow-moving targets are easily submerged by the spread of Doppler clutter in this situation, which causes severe performance loss in the subsequent moving target detection stage. Existing airborne radar clutter suppression techniques mainly include displaced phase center antenna (DPCA) [4], along-track interferometer (ATI) [5], and space-time adaptive processing (STAP) [6]. Compared to the strict requirements of DPCA and ATI in engineering, STAP has relaxed implementation conditions. Additionally, with the wide use of multichannel radar systems in recent years, STAP has been successfully applied in many multichannel radar systems and can verifiably achieve better clutter suppression than the other two methods.

STAP is a space–time two-dimensional filter that mainly takes advantage of the distribution property of clutter. It adaptively estimates the clutter covariance matrix (CCM) to calculate the optimal filter weight using the neighbor range samples of the cell under test (CUT). Thus, it is suitable for a dynamically changing clutter environment. Nevertheless, STAP assumes that different training samples are independent and identically distributed (IID) with each other. The RMB rule [7], developed by Reed et al., proves that the number of IID training samples required to keep the average loss below 3 dB is twice the system degrees of freedom (DoFs). However, the real environment often appears as a heterogeneous distribution. In some extreme cases, every adjacent range cell has an entirely different clutter distribution [8], challenging the acquisition of adequate secondary data in STAP. Therefore, overcoming the prob-

lem of training sample shortage is an important research direction in the STAP field.

In recent decades, numerous STAP methods have been developed to address the above problem. Generally, most of these methods can be roughly classified into three types: reduced-dimensional (RD), reduced-rank (RR), and direct data domain (DDD). For RD-STAP [9], [10], a reduced-dimension transformation matrix was first applied to the original echo data to convert the full-dimension domain to the local data domain. Subsequently, the conventional STAP method for RD data was performed. Hence, the required number of training samples is reduced to twice the RD DoFs. Generally, RD-STAP often involves transformation among the element, pulse, beam, and Doppler domains. The advantages and disadvantages of these four domains are carefully analyzed in [9]. By contrast, the RR-STAP method uses the low-rank characteristic of the CCM and projects the signal onto its subspace [11], [12]. Subsequently, it reduces the training sample number to twice the clutter rank. However, owing to the huge computational burden of eigenvalue decomposition and uncertain clutter rank numbers in nonhomogeneous environments, such STAP methods are not used in practical engineering. The last type is DDD-STAP, which aims at extreme conditions [13]. It specifically constructs the available training samples using a space–time sliding window in CUT. The advantage of DDD-STAP is that it eliminates the requirement of neighbor training samples, but the performance degradation caused by aperture loss during its use is inevitable. Although the RD-STAP and RR-STAP methods can effectively reduce the training sample number, their requirements are still hard to satisfy in variable environments. DDD-STAP only needs a single sample, but its performance and calculations are unacceptable in real applications. Therefore, a small sample condition and a high-performance STAP method are still required in engineering.

With the rapid development of compressive sensing (CS), signal sparsity has gradually aroused the attention of many scholars [14]. The CS theory proves that a sparse signal can be accurately recovered with a sample number below the Nyquist sampling theorem. Inspired by the CS theory, sparse recovery techniques has been introduced in the field of radar signal processing [15]–[17], and correspondingly, sparse recovery STAP (SR-STAP) methods have been developed by researchers to estimate CCM with few training samples,even in a single sample [18]–[23]. Owing to the coupling characteristic of clutter space–time, it sparsely distributes in the angle-Doppler two-dimensional domain, which means that the clutter signal is satisfied with the sparsity assumption [24], [25]. According to the different solution algorithms, SR-STAP can be roughly separated into three types: greedy, convex relaxation, and Bayesian algorithms. The first type of algorithm only searches the best-matched single atom in each iteration until the number of atoms reaches the preset threshold, for example, orthogonal matching pursuit [26]. Greedy algorithms have low complexity, but their result is a local optimal solution, and recovery performance is severely affected by dictionary correlation. Thus, it is unsuitable for the SR-STAP application, where the space–time dictionary has high coherence. A more commonly used method is convex relaxation, which converts the nonconvex sparse problem formulation to the ℓ_1-norm problem or its equivalent form [27], [28], and solves it with the help of the convex optimization method. Typical of the prime-dual interior point method and the gradient descent method, the drawback of this method is that it relies heavily on selecting parameters [29], which cannot exactly be determined in advance. Compared to the methods mentioned above, the Bayesian type recasts the sparse problem as a maximum a posteriori estimation (MAP) and estimates the sparse coefficients with their mean value [30]. For instance, sparse Bayesian learning (SBL) obtains the sparsest solution by imposing different prior distributions in the MAP [31]. The study [32] demonstrates that SBL can achieve a sparser solution than the above two types of methods. Although SR-STAP has a significant advantage in reducing the sample requirement, it requires huge time resources in optimization, especially in large data sizes. This is the major constraint of SR-STAP for real applications.

To accelerate the calculation of SR-STAP algorithms, some strategies have been put forward in recent years. From our view, these methods include two classes: the RD dictionary and fast-optimal algorithms. The size of the discrete dictionary directly determines the scale of processing. Thus, replacing the original dictionary with its RD form can effectively reduce computation. A classic route involves extracting grids of clutter areas with prior knowledge [33]–[35]. The other route is similar to the idea of RD-STAP; a data-independent transformation is applied in the dictionary to realize RD operation [36]. When these two routes are compared, it can be seen that although the first route reduces the number of atoms, the second route reduces the size of the atoms. The shortcoming of such methods is that RD performance loss exists during recovery. For fast-optimal algorithms, it usually takes a better simple processing structure to replace the complex original methods and thus to achieve an improvement in speed (e.g., fast SBL method) [37]. However, these methods often need to weigh between performance and speed. To achieve a high-efficiency algorithm without a performance loss, our research group first introduced the tensor processing structure in SR-STAP [18], and the results demonstrated that the computation loading of SR-STAP solving can be accelerated by several orders of magnitude. It is important to note that tensor-based processing is independent of algorithm assumptions, which means that there is no degradation in clutter suppression of tensor-based SR-STAP. However, there is still a distance from the tensor-based SR-STAP algorithm to its real-time realization because of its relatively large optimization time.

Additionally, we found that existing SR-STAPs are all focused on element-level processing. Almost all real

airborne radar system processes are at the subarray level [38], [39]. The reason is summarized as follows: for achieving long-range target detection, a large array with many antenna elements is required by airborne radars. However, such huge data processing is unrealizable in real-time systems because of limited hardware resources. Therefore, subarray-level (also called subchannel) signal processing is necessary for the application. This means that subarray-level SR-STAP research has more practical significance.

Driven by the aforementioned challenges, we propose a basic framework for subarray-level SR-STAP. To our knowledge, there has never been a detailed discussion of this. Initially, a fast separable tensor-based SBL (STSBL) processing structure was developed to solve the subarray-level SR-STAP problem. Subsequently, combined with the real measured data from our developed dual-channel airborne radar system, we verified the effectiveness of the proposed method in WGMTI processing. The main contributions of this work are summarized as follows:

- On the basis of multi-channel airborne radar echo model, we expand the conventional element-level SR-STAP problem to the channel-level, and establish its corresponding optimal mathematic model with a tensor form. This provides a solution for the difficult that element-level cannot be effectively applied in the real multi-channel airborne radar application.
- For multi-channel SR-STAP application, the traditional uniform space-time dictionary will produce huge time consumption in optimization procedure due to its large space-time freedom. For this purpose, we proposed a joint dictionary design method, which can achieve the same recovery performance as dense uniform dictionary with fewer atoms by combing the multi-channel clutter distribution characteristics, and the same time, it keeps the space-time separable property of dictionary. This way can effectively relieve the subsequent processing pressure.
- To make full use of the joint dictionary, a separable tensor-based SBL method is developed for simultaneously solving multiple dictionary SR-STAP problem. The proposed method separates the original tensor SR-STAP problem into multi sub-problem, and approximately independently performs the sub-problems in each iterations, thus can further improve the computational efficiency than the conventional SBL methods, and its unique separable structure also makes it more suitable for the realization in the existing multicore processors.
- Existing SR-STAP studies less involve the SR-STAP performance in real measurement data. In this article, the echo data sampled by our developed dual-channel airborne radar is used to verify the performance of the proposed STSBL-STAP methods in the whole WGMTI processing. From the processing results, it fully demonstrates that the proposed channel-level SR-STAP can achieve an acceptable clutter suppression ability with very few samples compared with the conventional STAP method.

It is necessary to explain that the proposed separable tensor-based SR-STAP is distinguished from our previous study [18], [31], [40] in following two aspects:

1) Regarding the goal, in [31], our research group first introduced the SBL method in the SR-STAP application. The results showed that SBL methods accurately estimate clutter distribution with very few training samples. Subsequently, in [40], our goal was to further improve the performance of SBL with clutter block structure prior knowledge, which we called prior block SBL (PBSBL). Unlike the above two methods, in [18], our research direction turned to a fast algorithm for improving performance because of the real hardware resource constraint. We first proposed a tensor-based SBL-STAP (TSBL-STAP). However, its computation loading is still huge compared with the requirements of real applications. Note that the above methods are all focused on element-level application. In this paper, we further combine tensor-based processing with a parallel idea and give a more practical processing separable structure of TSBL-STAP for subarray-level application, which we call the STSBL-STAP method.
2) Regarding the application, limited by the conventional vector-based structure, SBL-STAP and PBSBL-STAP are only suitable for small-scale dictionaries. Compared to these two methods, the TSBL-STAP proposed in [18] is specially designed for large-scale dictionary applications, which can separate a large-scale matrix calculation into multiple small-size matrix operations with the property of the Kronecker product. In this article, the proposed STSBL-STAP is aimed at multiple independent dictionaries, which can be realized simply by dividing the space–time plane. In other words, STSBL-STAP has two decompositions. The first is to separate a large dictionary into multiple medium dictionaries, and the second is to decompose the medium dictionary matrix into multiple small matrices. Therefore, it is more suitable for real parallel processing core applications.

The remainder of this paper is organized as follows. Section II describes the signal model of a planar subarray-level multichannel airborne radar. In Section III, a detailed analysis of the subarray-level clutter spectrum and clutter rank is discussed, and in Section IV, the tensor-based subarray-level SR-STAP problem formulation is developed. Subsequently, a joint dictionary design method and STSBL realization are provided. In Section V, we briefly introduce our WGMTI processing framework. Sections VI and VII are the results of the simulation experiment and the real measurement experiment, respectively. Finally, the conclusions are deduced in Section VIII.

Notations: Throughout this paper, we use nonbold case letters, bold lowercase letters, and bold uppercase letters to denote scalars, column vectors, and matrices, e.g., x, $\boldsymbol{x}$, $\boldsymbol{X}$. Double line body $\mathbb{R}$, $\mathbb{C}$ mean the real number domain and complex number domain. The superscripts $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ represent the conjugate, transpose, and conjugate transpose operation, respectively. $|\cdot|$ denotes the absolute operator. $(\cdot)^{-1}$ is the inverse of matrix. $\mathrm{diag}(\cdot)$ is the diagonalization of the vector. $\boldsymbol{I}$ is a diagonal matrix with $\boldsymbol{I}_{i,i} = 1$.

II. The Mathematics Model of a Multichannel Airborne Radar

The illumination geometry of multichannel airborne radar is depicted in Fig. 1. Assume that the aircraft platform locates at $(0, 0, H)$ and moves along the x-axis with a constant velocity v_p. A multichannel airborne radar is mounted under the fuselage and keeps working in the side-looking mode (i.e., the angle between the antenna array and flight direction $\theta_p = 0°$). For the antenna array, it is assumed to be a uniform planar array (UPA) that contains M and N elements in rows in columns, and the distance of any two adjacent element is d. Commonly, such a large number of elements would receive huge echo data and cause a challenge in the subsequent processing. A common choice is to merge multiple elements into a subarray and to process the data at the channel level [38]. In numerous subarray construction methods, the uniform non-overlap subarray is often a common choice in practical airborne radar systems for its simple realization structure. Following this choice, the UPA is uniformly split into N_s channels along the azimuth direction.

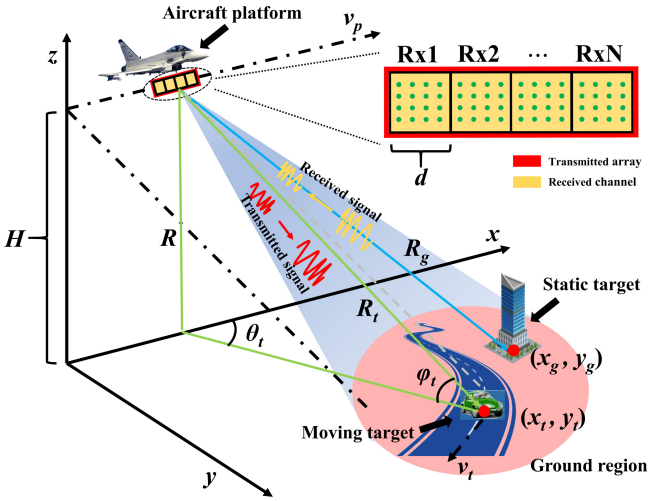


Fig. 1: Geometry model of a multichannel airborne radar.

Supposing that the carrier frequency is f_c, then the transmitted linear frequency modulation (LFM) signal is given as

$$s(t) = \mathrm{rect}\left(\frac{t}{T_p}\right)\exp\left[j2\pi\left(f_c t + \frac{\mu}{2}t^2\right)\right] \tag{1}$$

where

$$\mathrm{rect}(\frac{t}{T_p}) = \begin{cases} 1, if \ -\dfrac{T_p}{2} \le t \le \dfrac{T_p}{2} \\ 0, \mathrm{otherwise} \end{cases}$$

where T_p is the pulse width in each pulse period; $\mu = B/T_p$ represents the signal chirp rate; B is the working bandwidth. During each coherent processing interval, the multichannel airborne radar transmits K LFM waveform pulses to the target region with a fixed pulse repetition frequency (PRF) f_r.

Let the target azimuth angle and elevation angle be θ and φ, respectively. With the first-order Taylor approximation, the slant range of the constant-velocity ground target and the (n, m)th element can be approximately expressed with

$$\begin{aligned} R_{n,m}(t_k) \approx R_0^{n,m} - d(n\cos\phi + m\sin\varphi) \\ + \left(v_r - v_p\cos\phi + \frac{ndv_p}{R_0^{n,m}}\right)t_k \end{aligned} \tag{2}$$

where $R_0^{n,m}$ represents the initial slant range, v_r is the target radical velocity, and ϕ denotes the cone angle. t_k denotes the slow-time.

With a short delay time, the echo is collected by the received array. Passing a series of preprocessing steps, the baseband echo signal of the (n, m)th received element is as follows:

$$\begin{aligned} x_{n,m}(t, t_k) = \hat{\sigma}\omega(t_k)\,\mathrm{sinc}\left[B\left(t - \frac{2R_{n,m}(t_k)}{c}\right)\right] \\ \times \exp\left[-j\frac{4\pi}{\lambda}R_{n,m}(t_k)\right] \end{aligned} \tag{3}$$

where $\hat{\sigma}$ represents the complex amplitude in the range-compression target and $\lambda = c/f_c$ is radar wavelength. $\omega(\cdot)$ is the slow-time window. $\mathrm{sinc}(\cdot)$ is sinc function. By substituting Eq. (2) into Eq. (3), and organize the relevant term with its discrete form (i.e., $t_k = k/f_r$). The Eq. (3) can be simplified as

$$\begin{aligned} x_{n,m}(l, k) = \widetilde{\sigma}(l, k)\exp\left(j\frac{4\pi v}{\lambda}k\right) \\ \times \exp\left[j\frac{4\pi d}{\lambda}(n\cos\phi + m\sin\varphi)\right] \end{aligned} \tag{4}$$

where $\widetilde{\sigma}(l, k) = \hat{\sigma}_0\omega(k)\,\mathrm{sinc}\left[B\left(t - 2R_0^{n,m}/c\right)\right]$ is the complex reflection of the target; $l \in 1, \cdots, L$ is the l-th range sample point, and L is the total number of range gates in every pulse; $v = (v_r - v_p\cos\phi + ndv_p/R_0^{n,m})$ is the radical velocity of the relative target. To further synthesize the channel-level signal, it needs to incorporate a multi-element echo. As the partition rule mentioned before, each channel contains MN_{sn} elements, where $N_{sn} = N/N_s$, so the channel echo can be described as

follows:

$$x_{n_s}(l,k)=\sum_{n=1+(n_s-1)N_{sn}}^{n_sN_{sn}}\sum_{m=1}^{M}w_{n,m}\widetilde{\sigma}(l,k)\exp\left(j\frac{4\pi v}{\lambda f_r}k\right)\times\exp\left[j\frac{4\pi d}{\lambda}(n\cos\phi+m\sin\varphi)\right] \quad (5)$$

where $w_{n,m}$ is the element weight coefficient. In our radar system, direct synthesis is used here.

The above description is the received point target signal of multichannel airborne radar. In a real environment, a physical echo mainly involves three parts: clutter, moving target, and noise [9].

For the l-th range cell, the clutter from different directions (azimuth and elevation) can be modeled as follows:

$$x_{n_s}^{c}(l,k)=\sum_{n=1+(n_s-1)N_{sn}}^{n_sN_{sn}}\sum_{m=1}^{M}\int_{\theta}\int_{\varphi}\widetilde{\sigma}(l,k,\theta,\varphi)\exp\left(\frac{j4\pi v_k}{\lambda f_r}k\right)\times\exp\left[j\frac{4\pi d}{\lambda}(n\cos\phi+m\sin\varphi)\right]d\theta d\varphi \quad (6)$$

Restricted by the array structure design and sweep mode, $\theta\in(-\pi/2,\pi/2]$ and $\varphi\in(0,\pi/2]$. Additionally, the clutter is mainly composed of static ground targets, hence its radical velocity only depends on the aircraft platform, i.e., $v\approx v_p\cos\varphi\cos(\theta+\theta_p)$. $\widetilde{\sigma}(l,k,\theta,\varphi)$ is related to the spatial state of moving targets, it is often simulated with independent random distribution [41].

The targets can emerge at any location with an unknown initial velocity, so the multi-targets model can be expressed as

$$x_{n_s}^{t}(l,k)=\sum_{n=1+(n_s-1)N_{sn}}^{n_sN_{sn}}\sum_{m=1}^{M}\sum_{n_t=1}^{N_t}\widetilde{\sigma}_{n_t}(l,k)\exp\left(j\frac{4\pi v_{n_t}}{\lambda f_r}k\right)\times\exp\left[j\frac{4\pi d}{\lambda}(n\cos\phi+m\sin\varphi)\right] \quad (7)$$

where v_{n_t} is the relative n_t-th target radical velocity.

Noise is mainly caused by the radar hardware system, which leads to random fluctuation interference for real target signals. As a common assumption, a zero mean Gaussian white noise with a variance of σ^2, i.e.,$x_{n_s}^{n}(l,k)\sim\mathcal{N}(0,\sigma^2)$ is considered throughout this paper.

Combining the clutter, target, and noise models before, for a single-range piece, the discrete vectorization echo data can be rewritten as follows:

$$\begin{aligned}\boldsymbol{x}&=\boldsymbol{x}_t+\boldsymbol{x}_c+\boldsymbol{x}_n\\&=\boldsymbol{T}^H\sum_{n_t=1}^{N_t}\gamma_{n_t}\left[\boldsymbol{s}_a(\theta_{n_t},\varphi_r)\otimes\boldsymbol{s}_e(\varphi_{n_t})\otimes\boldsymbol{s}_d\left(f_d^{n_t}\right)\right]\\&+\boldsymbol{T}^H\sum_{n_r=0}^{N_r}\sum_{n_c=1}^{N_c}\gamma_{n_r,n_c}\boldsymbol{s}(\theta_{n_c},\varphi_{n_r})+\boldsymbol{T}^H\boldsymbol{x}_n\end{aligned} \quad (8)$$

where γ_{n_t} and γ_{n_r,n_c} denote the complex amplitude of targets and clutter patches. $\boldsymbol{s}(\theta_{n_c},\varphi_{n_r})\in\mathbb{C}^{N_sK\times1}$ is two-dimensional space-time steering vector and it can be separated as azimuth steering vector $\boldsymbol{s}_a(\theta_{n_c},\varphi_{n_r})$, elevation steering vector $\boldsymbol{s}_e(\varphi_{n_r})$ and temporal steering vector $\boldsymbol{s}_d(\theta_{n_c},\varphi_{n_r})$, that is,

$$\boldsymbol{s}(\theta_{n_c},\varphi_{n_r})=\boldsymbol{s}_a\left(f_a^{n_c,n_r}\right)\otimes\boldsymbol{s}_e\left(f_e^{n_r}\right)\otimes\boldsymbol{s}_d\left(f_d^{n_c,n_r}\right) \quad (9)$$

where

$$\begin{aligned}\boldsymbol{s}_e\left(f_e^{n_r}\right)&=\left[1,\cdots,\exp\left[j\pi(M-1)f_e^{n_r}\right]\right]^T\\\boldsymbol{s}_a\left(f_a^{n_c,n_r}\right)&=\left[1,\cdots,\exp\left[j\pi(N-1)f_a^{n_c,n_r}\right]\right]^T\\\boldsymbol{s}_d\left(f_d^{n_c,n_r}\right)&=\left[1,\cdots,\exp\left[j\pi(K-1)f_d^{n_c,n_r}\right]\right]^T\end{aligned}$$

where $f_a^{n_c,n_r}=2d\cos\varphi_{n_r}\cos\theta_{n_c}/\lambda$, $f_e^{n_r}=2d\sin\varphi_{n_r}/\lambda$ and $f_d^{n_c,n_r}=2v\cos\varphi_{n_r}\cos\theta_{n_c}/(\lambda f_r)$ are the normalized azimuth frequency, elevation frequency, and Doppler frequency respectively. N_r represents the range ambiguity number. $f_d^{n_t}=2v_{n_t}/(\lambda f_r)$ denotes the Doppler frequency of moving targets. $\boldsymbol{T}\in\mathbb{R}^{NMK\times N_sK}$ is a subarray transformation matrix, which can be decomposed as the Kronecker product of $N\times N_s$ azimuth transformation matrix $\boldsymbol{T}_a$, $M\times1$ unit elevation transformation matrix $\boldsymbol{T}_e$, and Doppler transformation $\boldsymbol{T}_d$, that is,

$$\boldsymbol{T}=\boldsymbol{T}_a\otimes\boldsymbol{T}_e\otimes\boldsymbol{T}_d \quad (10)$$

where

$$\boldsymbol{T}_a=\begin{bmatrix}1_1&0&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\1_1&0&\cdots&0\\0&1_2&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&1_2&\cdots&0\\\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&1_{N_s}\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1_{N_s}\end{bmatrix},\boldsymbol{T}_d=\begin{bmatrix}1&\cdots&0\\\vdots&\ddots&\vdots\\0&\cdots&1\end{bmatrix}$$

where the number of "1" in each column of $\boldsymbol{T}_a$ is N_{sn}. Note that when the beam points non-normal direction, the element of the $\boldsymbol{T}_a$ should be updated with its corresponding steering vector value. The covariance matrix is

$$\boldsymbol{R}=\mathbb{E}\left\{\boldsymbol{x}\boldsymbol{x}^H\right\}=\boldsymbol{R}_t+\boldsymbol{R}_c+\boldsymbol{R}_n \quad (11)$$

where $\boldsymbol{R}_n=\sigma^2\boldsymbol{I}\in\mathbb{C}^{N_sK\times N_sK}$ denotes the noise covariance matrix. Regarding the target covariance matrix $\boldsymbol{R}_t$ and CCM $\boldsymbol{R}_c$, they can be further expressed as follows:

$$\begin{aligned}\boldsymbol{R}_t&=\sum_{r=1}^{N_t}|\gamma_{n_t}|^2\left(\boldsymbol{T}^H\boldsymbol{s}_{n_t}\boldsymbol{s}_{n_t}^H\boldsymbol{T}\right)\\\boldsymbol{R}_c&=\sum_{n_r=0}^{N_r}\sum_{n_c=1}^{N_c}|\gamma_{n_r,n_c}|^2\left(\boldsymbol{T}^H\boldsymbol{s}_{n_r,n_c}\boldsymbol{s}_{n_r,n_c}^H\boldsymbol{T}\right)\end{aligned} \quad (12)$$

For the target-free case, the covariance matrix should be updated as $\boldsymbol{R}=\boldsymbol{R}_c+\boldsymbol{R}_n$, also called the clutter plus noise covariance matrix. The optimal STAP filter is derived from the linear constraint minimal variance rule, which can be described as follows:

$$\boldsymbol{w}=\mu\boldsymbol{R}^{-1}\boldsymbol{s}_o \quad (13)$$

where $\boldsymbol{s}_o$ is the target space-time steering vector and $\mu = 1/\boldsymbol{s}_o^H \boldsymbol{R}^{-1} \boldsymbol{s}_o$ is the normalized factor.

III. Analysis of Subarray-Level Clutter

A. Subarray Clutter Space–Time Distribution Property

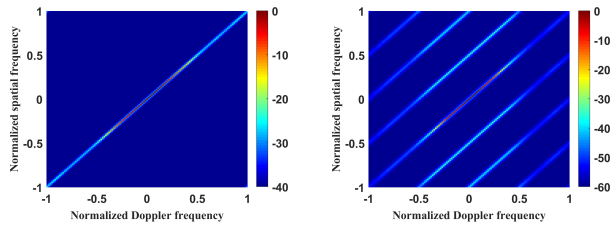
A multichannel radar usually adopts a wide-transmitting and narrow-receiving mode, thus a grating lobe inevitably emerges in channel-level application because the distance d_s of the adjacent synthesis channel is bigger than $\lambda/2$. The deviation between the main lobe and the h-th grating lobe [39] is

$$f_a - f_{ao} = \pm h, h = 1, 2, \cdots \tag{14}$$

where $f_a = N_{sn} d \cos\theta \cos\varphi / \lambda$. The azimuth angle of the h-th grating lobe can be inverted by

$$\hat{\theta} = \arccos\left(\cos\theta_o \pm \frac{h\lambda}{N_{sn} d \cos\varphi_o} \right) \tag{15}$$

Fig. 2 gives a clutter spectrum comparison results, where the parameter setting is as follows: $N = 8$, $M = 6$, and $K = 8$. In subarray processing, 2×2 channels are synthesized in rows and columns. Compared to these two spectrums, it can be found that the subarray-level clutter spectrum in Fig. 2(b) includes many grating-lobe clutter components, which holds the same distribution characteristics as the main lobe clutter.



(a) (b)

Fig. 2: Comparison of the clutter spectrum: (a) element level, (b) subarray level.

This phenomenon can be explained as shown in Fig. 3. When the space–time beam scans point "A" its synthesis clutter is mainly composed of the original clutter in the cross-curve area. For the Doppler beams, the main clutter is located in the sidelobe and has less contribution to the final result. However, for space beams, the main clutter may be exposed in the grating-lobe area, and then may lead to an increase in clutter power. Note that the same analysis holds for other cases, such as clutter periodic extension and non-side-looking mode.

More fundamentally, the above angle-Doppler spectrum is a projection from the 3D space to the 2D plane [42]. Consequently, to better understand the impact of channel synthesis, Fig. 4 displays the three-dimensional clutter spectrum of the full element and subarray. From these results, the effect of channel synthesis can be summarized as follows: a series of clutter (similar to the main clutter distribution) appears regularly in the 3D space.

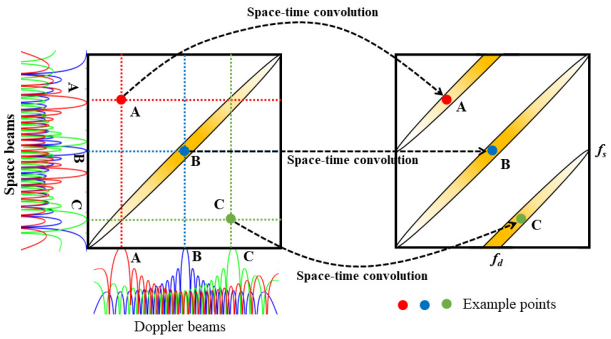
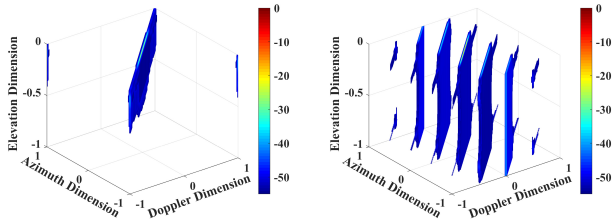


Fig. 3: The mechanism of the grating-lobe clutter.



(a) (b)

Fig. 4: Comparison of the clutter spectrum: (a) element level, (b) subarray level.

B. Subarray Clutter Rank Property

The clutter rank plays an important role in STAP because it determines the required DoFs of effective clutter suppression. On the basis of [43] and [38], we further give an estimation formula of planar subarray clutter rank as follows:

$$r_c = \min\left(M, N_r\right)\left(N_s + \frac{\beta}{N_{sn}}\left(M - N_w\right)\right) N_w \tag{16}$$

where

$$N_w = \min\left(N_{sn}, \frac{N_{sn}}{\gcd\left(\mathrm{mod}\left(\beta, N_{sn}\right), N_{sn}\right)}\right)$$

where gcd() represents the greatest common divisor and mod() is the modulus after division. Fig. 5 exhibits a clutter rank estimation result of different subarray levels ($N_s = 2$ and 4) and full elements in the same condition as in the last section, where $N_r = 2$. It is proven from the knee of the eigenvalue curve that the theoretical estimation is consistent with the simulation. This also illustrates that the available DoFs of the channel level are significantly smaller than the element level, which means that some degree of performance loss is inevitable in subarray-level STAP clutter suppression.

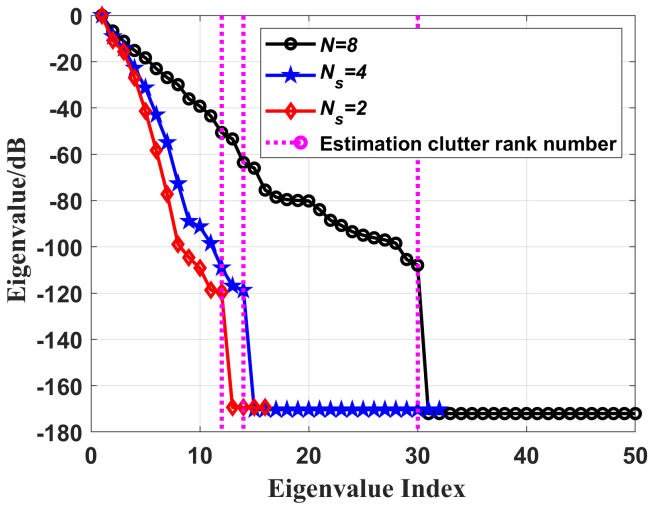


Fig. 5: The clutter rank results of the different subarrays.

IV. Fast Subarray Sparse Recovery STAP Method

A. Tensor Mathematics Foundation

A brief review of tensor algebra is provided in this subsection. $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{i_1 \times i_2 \times \cdots i_N}$ denotes N-order (also called N-mode) tensor, and i_n represents the element in nth mode. $\mathcal{V}_{\mathrm{I}_1}\{\cdot\}$ denotes the tensor vectorization operation, and its subscript means the rearrangement size. Analogously, $\mathcal{M}_{\mathrm{I}_1,\mathrm{I}_2}\{\cdot\}$ and $\mathcal{T}_{\mathrm{I}_1,\mathrm{I}_2,\cdots,\mathrm{I}_N}\{\cdot\}$ are matricization and tensorization operation, respectively. $\mathcal{G}_{\mathrm{I}_1,\mathrm{I}_2,\cdots,\mathrm{I}_N}\{\cdot\}$ represents the dimension exchange operation. Then, the Hadamard product and division are denoted as $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{X}} \odot \boldsymbol{\mathcal{Y}}$ and $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{X}} \oslash \boldsymbol{\mathcal{Y}}$. The inner product and outer product are $\langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}} \rangle$ and $\boldsymbol{\mathcal{X}} = \boldsymbol{s}_1 \circ \boldsymbol{s}_2 \circ \cdots \circ \boldsymbol{s}_{I_N}$. $(\boldsymbol{\mathcal{X}} \times_n \boldsymbol{S})_{i_1,i_2,\cdots,p,\cdots i_N}$ represents the n-mode product and $\|\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{Y}}\|_F$ is the tensor Fibonacci norm.

B. Subarray Tensor Sparse Recovery STAP Problem Formulation

Recalling Section III, the original clutter was continuously distributed in the whole angle-Doppler plane. Accordingly, to construct the appropriate overcomplete space–time dictionary, the entire 2D plane is first discretized into some grids. Suppose that the grid resolution is narrow enough to approximately represent the actual clutter locations; then, combined with Eq. (8), the clutter signal can be further expressed as follows:

$$\boldsymbol{x} = \boldsymbol{T}^H \boldsymbol{S}\boldsymbol{\gamma} + \hat{\boldsymbol{n}} \tag{17}$$

where $\hat{\boldsymbol{n}} = \boldsymbol{T}^H \boldsymbol{n}$ and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \cdots, \gamma_{N_s K_d}]^T$ denotes the sparse coefficients.. $\boldsymbol{S} = \boldsymbol{S}_a \otimes \boldsymbol{S}_e \otimes \boldsymbol{S}_d$ is the overcomplete space–time dictionary, and $\boldsymbol{S}_a$, $\boldsymbol{S}_e$, and $\boldsymbol{S}_d$, respectively, represent the azimuth space dictionary, elevation space dictionary, and Doppler dictionary. For a full-element SR-STAP application, $\boldsymbol{S}$ is composed of its corresponding space–time steering vector in the divided grids. However, owing to the subarray synthesis operation, the above dictionary construction is unsuitable for subarray applications. Therefore, it is necessary to find a more appropriate dictionary basis. In this article, the RD space–time steering vector is adopted to establish the dictionary, and the realization detail is discussed in Section IV.C. Subsequently, with tensor transformation operations, Eq. (17) can be rewritten as follows:

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{H}} \times_1 \widehat{\boldsymbol{S}}_a \times_2 \boldsymbol{S}_d + \boldsymbol{\mathcal{N}} \tag{18}$$

where $\boldsymbol{\mathcal{X}} = \mathcal{T}_{N_s,K,L}\{\boldsymbol{x}\}$, $\boldsymbol{\mathcal{H}} = \mathcal{T}_{N_a,K_d,L}\{\boldsymbol{\gamma}\}$ and $\boldsymbol{\mathcal{N}} = \mathcal{T}_{N_s,K,L}\{\hat{\boldsymbol{n}}\}$ are the tensor forms corresponding to the original variables, and N_a and K_d are the atom numbers of the azimuth space dictionary and Doppler dictionary, respectively. Note that $\boldsymbol{S}_e$ is ignored here because of the azimuth dual-channel synthesis and $\widehat{\boldsymbol{S}}_a$ is the subarray azimuth dictionary. The goal of the tensor-based SR-STAP is to explore the sparest solution of $\boldsymbol{\mathcal{H}}$ as much as possible, and its mathematical problem model is

$$\begin{aligned} &\min_{\boldsymbol{\mathcal{H}}} \left\|\mathcal{M}_{N_a K_d, L}\{\boldsymbol{\mathcal{H}}\}\right\|_{2,1} \\ &\text{s.t.} \left\|\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{H}} \times_1 \widehat{\boldsymbol{S}}_a \times_2 \boldsymbol{S}_d\right\|_F^2 \le \varepsilon \end{aligned} \tag{19}$$

where $\|\cdot\|_{2,1}$ is the combination of ℓ_2-norm and ℓ_1-norm; specifically, execute ℓ_1-norm in a column and ℓ_2-norm in a row. ε can be regarded as a threshold to adjust the acceptable noise level. Another commonly used expression of Eq. (19) is the unconstrained form:

$$\min_{\boldsymbol{\mathcal{H}}} \left\|\mathcal{M}_{N_a K_d, L}\{\boldsymbol{\mathcal{H}}\}\right\|_{2,1} + \rho \left\|\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{H}} \times_1 \widehat{\boldsymbol{S}}_a \times_2 \boldsymbol{S}_d\right\|_F^2 \tag{20}$$

where ρ is the penalty coefficient. Finally, the CCM estimation was calculated as follows:

$$\hat{\boldsymbol{R}} = \boldsymbol{S}\mathrm{diag}\left\{\frac{1}{L}\sum_{l=1}^{L} \mathcal{V}_{N_a K_d}\left\{\left|\boldsymbol{\mathcal{H}}_{:,:,l}\right|^2\right\}\right\} \boldsymbol{S}^H + \delta \boldsymbol{I} \tag{21}$$

where δ is the diagonal loading term [44]. For simplifying the analysis, let $\boldsymbol{S} = \widehat{\boldsymbol{S}}_a \otimes \boldsymbol{S}_d$ in the subsequent section.

C. Construction of a Prior-Aided Subarray Dictionary

The traditional uniform dictionary has a simple implementation. However, as the grid resolution increases, the number of atoms increases exponentially, which leads to a huge computation load in a practical SR-STAP application. In this section, a joint dictionary is developed to achieve a better compromise between grid number and performance.

1. *Uniform Subarray Space–Time Dictionary*

Generally, a grid resolution factor (GRF) β is adopted to adjust the grid resolution, that is, $N_a = \beta_a N$ in the spatial dimension and $K_d = \beta_d K$ in the temporal dimension. Hence, the space–time dictionary of the uniform subarray $\boldsymbol{S}_u \in \mathbb{C}^{N_s K \times N_a K_d}$ is given as follows:

$$\boldsymbol{S}_u = \left[\boldsymbol{s}\left(f_a^1, f_d^1\right), \boldsymbol{s}\left(f_a^2, f_d^2\right), \cdots, \boldsymbol{s}\left(f_a^{N_a}, f_d^{K_d}\right)\right] \tag{22}$$

where $\boldsymbol{s}\left(f_a^{n_a}, f_d^{k_d}\right) \in \mathbb{C}^{N_s K}$ is the subarray-level space–time steering vector corresponding to the n_a-th and k_d-th in the azimuth–Doppler plane, and it can be calculated with

$$\boldsymbol{s}\left(f_a^{n_a}, f_d^{k_d}\right) = \left[\hat{\boldsymbol{T}}_a^H \boldsymbol{s}_a\left(f_a^{n_a}\right)\right] \otimes \boldsymbol{s}_d\left(f_d^{k_d}\right) \tag{23}$$

It should be noted that each atom corresponds to a unique space transformation matrix $\hat{\boldsymbol{T}}_a$. Moreover, Eq. (23) only considers the azimuth multichannel airborne radar model mentioned in Section II. As for other subarray configurations (e.g., overlapping subarray and nonuniform subarray), the proposed dictionary construction is still working. In addition, if the subarray contains a pitch scanning function (planar subarray), the elevation transformation matrix $\hat{\boldsymbol{T}}_e$ can be constructed in the same way.

$$\hat{\boldsymbol{T}}_a = \begin{bmatrix} 1 & e^{j2\pi f_a^{n_a}} & \cdots & e^{j2\pi(N_a-1)f_a^{n_a}} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & e^{j2\pi f_a^{n_a}} & \cdots & e^{j2\pi(N_a-1)f_a^{n_a}} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & e^{j2\pi f_a^{n_a}} & \cdots & e^{j2\pi(N_a-1)f_a^{n_a}} \end{bmatrix}^T$$

The drawback of $\boldsymbol{S}_u$ can be summarized as follows:

1) Compared with clutter region grids, noise area grids contribute a little to recovery. Unfortunately, owing to the sparsity of clutter distribution, most existing grids in the dictionary are noise area grids, and this will be more noticeable in large-scale input data [34];
2) The actual clutter region is continuous, and their locations often fall outside the divided grids. Thus, the grid mismatch effect is commonly unavoidable when applying a discrete dictionary [45]. Increasing the GRF can effectively mitigate the performance degradation caused by mismatches. However, because of the uniform segmentation in the 2D space, many useless grids will occupy huge hardware resources in the estimation.

Recently, some nonuniform dictionary construction methods have been proposed to resolve the dilemma of a uniform dictionary. Nevertheless, such nonuniform dictionaries again introduce other questions while improving performance. One important problem is the coupling of space and time dimensions (i.e., the two dimensions are no longer independent because the divided grids are directly selected in the angle-Doppler domain). Consequently, the 2D space–time dictionary cannot be decomposed in its independent dimension, which leads to many fast realization structures losing effectiveness. Additionally, a complicated two-dimensional grid segmentation method brings additional time costs.

2. *Clutter Dictionary*

To retain the separable ability of the space–time dictionary, the rule that must be satisfied is that the segmentation should be executed in each dimension, which means that the design freedom is reduced from $N_a K_d$ to $N_a + K_d$. Consequently, the rational selection of the division density of different areas with finite DoFs becomes a crucial problem in space–time dictionary design. In [46], a simple but potent idea is proposed by selecting the appropriate f_d based on the prior radar parameter. Drawing on this concept, we propose a new design method for the grating-lobe clutter ridge dictionary.

a. Azimuth Dimension

The first step is to lock the distribution region of the grating-lobe clutter. According to the analysis in Section III.A, the h-th grating-lobe space frequency relative to the main lobe clutter is as follows:

$$f_a = f_{ao} + \frac{h\lambda}{N_{sn} d \cos\varphi_o} \tag{24}$$

To further reduce complexity, the sparsely sampling of sidelobe clutter area is achieved by the subsequent uniform dictionary, in the clutter dictionary construction, only consider the powerful main lobe clutter area and its corresponding grating-lobe clutter area, which is a reasonable assumption in application because the strong clutter is primarily caused by these areas. Simultaneously, to reasonably allocate dense sampling points in the focus area, we introduce a nonlinear transformation function and adopt it to select the appropriate sampling points. The sampling transformation function can be expressed as follows:

$$\begin{cases} y = \exp\left(a\left|x - x_0\right|\right) \\ x = x_0 \pm \dfrac{\log\left(y\right)}{a} \end{cases} \tag{25}$$

where a and x_0 are the adjustable parameters used to adjust the function shape and location of the center sampling point. Fig. 6 compares linear and nonlinear sampling; the sampling points in y and x are the linear and nonlinear sampling results, respectively. Subsequently, with the help of a nonlinear sampling function, dense sampling near the main clutter and grating-lobe clutter area can be achieved in construction.

b. Doppler Dimension

For the Doppler dimension, a reasonable sampling strategy takes dense sampling in the main clutter area and sparse sampling in other areas concerning the main beam region. The method that we implemented was the inverse of nonlinear space domain sampling. A detailed realiza-

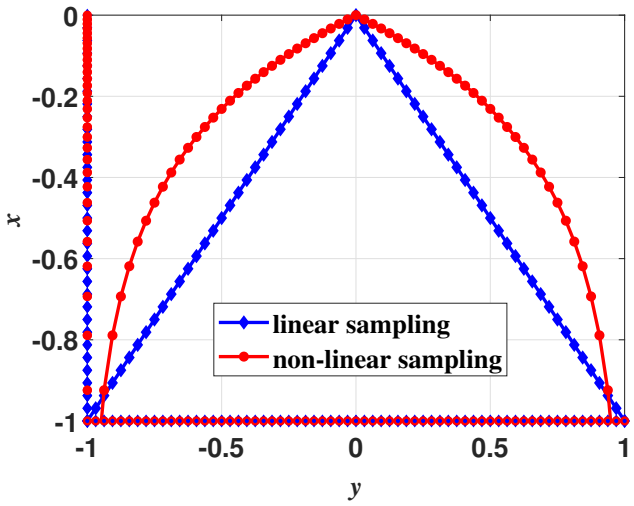


Fig. 6: Comparison of linear and nonlinear sampling.

tion list is as follows: 1) uniformly dividing the azimuth angle in the range from -1 to 1 with K_d points; 2) with the help of nonlinear transformation function, converting the uniform space sampling points to nonuniform sampling; and 3) getting the mapping results from space domain to Doppler domain by the relation between f_d and f_a, that is,

$$f_d = \beta \left[f_a \cos\theta_p \frac{2d}{\lambda} - \sin\theta_p \sqrt{\left(\frac{2d}{\lambda}\cos\varphi\right)^2 - f_a^2} \right] \quad (26)$$

Fig. 7 shows the result of the dictionary construction of the grating-lobe clutter ridge for the dual channel, where the black dot is the sampling point in the 2D plane.

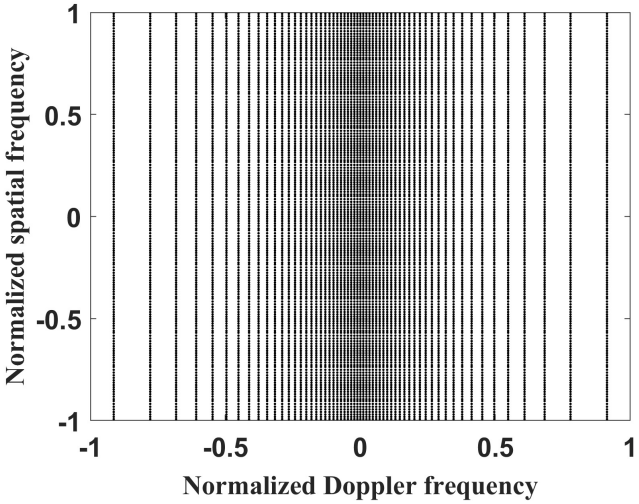


Fig. 7: An example of a grating-lobe clutter ridge dictionary.

3. *Joint Dictionary*

The traditional dictionary uniformly samples the space–time plane, but its useless grids waste many calculation units. By contrast, the ridge dictionary of the grating-lobe clutter decreases the grid number at the cost of reducing the sampling in the nonclutter region, which may cause potential instability in recovery. To combine the advantages of both dictionaries, a joint dictionary is provided as follows:

$$\boldsymbol{S} = \boldsymbol{S}_u \cup \boldsymbol{S}_c = \left[\begin{array}{cc} \boldsymbol{S}_u & \boldsymbol{S}_c \end{array} \right] \quad (27)$$

where $\boldsymbol{S}_c$ represents the clutter dictionary. As expected, in this way, a rough $\boldsymbol{S}_u$ with $\boldsymbol{S}_c$ can achieve the same recovery performance as a dense $\boldsymbol{S}_u$. The hardware that is occupied by useless noise grids is reduced and partially converted to effective clutter grids. The whole flow of establishing a joint dictionary is summarized in Algorithm 1.

Algorithm 1: Joint Dictionary Construction

Clutter Dictionary:
Step 1: Calculate the f_a by Eq.(24) ;
Step 2: Calculate the nonlinear sampling locations in the grating-lobe by Eq.(25) ;
Step 3: Calculate the space and Doppler sampling points ;
Step 4: Generate the $\boldsymbol{S}_c$ by Eq.(23) ;
Uniform Dictionary:
Step 1: Divide space and Doppler grids uniformly ;
Step 2: Generate the $\boldsymbol{S}_u$ by Eq.(23) ;
Final step: $\boldsymbol{S}$;

D. Separable Tensor-Based Sparse Bayesian Learning STAP

Let $\hat{\boldsymbol{S}}_a^u \in \mathbb{C}^{N_s \times N_a^u}$, $\hat{\boldsymbol{S}}_a^c \in \mathbb{C}^{N_s \times N_a^c}$, $\boldsymbol{S}_d^u \in \mathbb{C}^{K \times K_d^u}$,$\boldsymbol{S}_d^u \in \mathbb{C}^{K \times K_d^c}$, $\boldsymbol{S}_u = \hat{\boldsymbol{S}}_a^u \otimes \boldsymbol{S}_d^u$ and $\boldsymbol{S}_c = \hat{\boldsymbol{S}}_a^c \otimes \boldsymbol{S}_d^c$. $N_a = N_a^u + N_a^c$ and $K_d = K_d^u + K_d^c$. Using Eq. (33), the joint dictionary tensor subarray SR-STAP problem can be modeled as follows:

$$\begin{aligned} &\min_{\boldsymbol{\mathcal{H}}_u, \boldsymbol{\mathcal{H}}_c} \left\| \left[\mathcal{M}_{N_a K_d, L} \left\{ \boldsymbol{\mathcal{H}}_u \right\}; \mathcal{M}_{N_a K_d, L} \left\{ \boldsymbol{\mathcal{H}}_c \right\} \right] \right\|_{2,1} \\ &\text{s.t.} \left\| \boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{H}}_u \times_1 \hat{\boldsymbol{S}}_a^u \times_2 \boldsymbol{S}_d^u - \boldsymbol{\mathcal{H}}_c \times_1 \hat{\boldsymbol{S}}_a^c \times_2 \boldsymbol{S}_d^c \right\|_{\mathrm{F}}^2 \leq \varepsilon \end{aligned} \quad (28)$$

where $\boldsymbol{\mathcal{H}}_u$ and $\boldsymbol{\mathcal{H}}_c$ are the sparse coefficient of the uniform dictionary and clutter ridge dictionary, respectively. Traditional SBL methods spend a huge amount of time and resources on large-scale matrix multiplication and inversion operations. In [18], tensor processing was applied to decompose the large-size multidimensional dictionary matrix into a series of small-size single-dimension dictionaries. It can achieve a considerable improvement in solving by three or four orders of magnitude, but for a real application, it has the following problems:

- ***Computation loading***: Although the dictionary dimension of the tensor-based SBL reduces from $N_sK \times N_aK_d$ to $N_s \times N_a$ for $\boldsymbol{S}_a$ and $K \times K_d$ for $\boldsymbol{S}_d$, it is still a large burden for the hardware; for example, assuming that $N_s = 8$, $K = 64$ and $\beta = 6$, then the whole data to be processed for a single-range cell is $\beta\left(N_s^2 + K^2\right) \approx 2.5 \times 10^4$, which is impossible to accept in real-time processing systems.
- ***Data Transfer***: The problem of transmission delay is not evident in small-scale data, but by increasing the size of the input data, the time cost of this operation will quickly rise even more than the calculation.

This motivated us to introduce the separable structure and propose the STSBL, which is aimed at utilizing the advantages of parallel processing to further speed up the time of calculation and data transfer.

Suppose that $\boldsymbol{\mathcal{H}}_u$ and $\boldsymbol{\mathcal{H}}_c$ are satisfied with the complex Gaussian distribution. Hence, the joint probability distribution $p\left(\boldsymbol{\mathcal{H}}_u|\boldsymbol{\Gamma}_u\right)$ and $p\left(\boldsymbol{\mathcal{H}}_c|\boldsymbol{\Gamma}_c\right)$ can be expressed as follows:

$$\begin{aligned} p\left(\boldsymbol{\mathcal{H}}_u|\boldsymbol{\Gamma}_u\right) &= \frac{1}{\pi^{N_a^u K_d^u L}\left(\prod_{i=1}^{N_a^u}\prod_{j=1}^{K_d^u}\boldsymbol{\Gamma}_u^{i,j}\right)^L} \\ &\quad \cdot \exp\left\{-\sum_{l=1}^{L}\left[\sum_{i=1}^{N_a^u}\sum_{j=1}^{K_d^u}\left(\left|\boldsymbol{\mathcal{H}}_u^{:,:,l}\right|^2 \oslash \boldsymbol{\Gamma}_u\right)_{i,j}\right]\right\} \\ p\left(\boldsymbol{\mathcal{H}}_c|\boldsymbol{\Gamma}_c\right) &= \frac{1}{\pi^{N_a^c K_d^c L}\left(\prod_{i=1}^{N_a^c}\prod_{j=1}^{K_d^c}\boldsymbol{\Gamma}_c^{i,j}\right)^L} \\ &\quad \cdot \exp\left\{-\sum_{l=1}^{L}\left[\sum_{i=1}^{N_a^c}\sum_{j=1}^{K_d^c}\left(\left|\boldsymbol{\mathcal{H}}_c^{:,:,l}\right|^2 \oslash \boldsymbol{\Gamma}_c\right)_{i,j}\right]\right\} \end{aligned} \tag{29}$$

where $\boldsymbol{\Gamma}_u^{i,j}$ and $\boldsymbol{\Gamma}_c^{i,j}$ correspond to the variance of (i,j) element in $\boldsymbol{\mathcal{H}}_u^{:,:,l}$ and $\boldsymbol{\mathcal{H}}_c^{:,:,l}$. As described in Section II, the noise obeys the complex Gaussian distribution with zero mean and variance σ^2, that is,

$$p\left(\boldsymbol{\mathcal{N}}|0,\sigma^2\right) = \frac{1}{\left(\pi\sigma^2\right)^{N_sKL}}\exp\left\{-\frac{1}{\sigma^2}\sum_{l=1}^{L}\left\|\boldsymbol{\mathcal{N}}_{:,:,l}\right\|_{\mathrm{F}}^2\right\} \tag{30}$$

When combined with Eq. (30), the likelihood function of $\mathcal{X}$ can be written as follows:

$$\begin{aligned} p\left(\boldsymbol{\mathcal{X}}|\boldsymbol{\mathcal{H}}_u,\boldsymbol{\mathcal{H}}_c,\sigma^2\right) &= \frac{1}{\left(\pi\sigma^2\right)^{N_sKL}} \\ &\quad \cdot \exp\left\{\frac{-1}{\sigma^2}\sum_{l=1}^{L}\left\|\begin{array}{c}\boldsymbol{\mathcal{X}}_{:,:,l} \\ -\boldsymbol{\mathcal{H}}_u^{:,:,l}\times_1\widehat{\boldsymbol{S}}_a^u\times_2\boldsymbol{S}_d^u \\ -\boldsymbol{\mathcal{H}}_c^{:,:,l}\times_1\widehat{\boldsymbol{S}}_a^c\times_2\boldsymbol{S}_d^c\end{array}\right\|_{\mathrm{F}}^2\right\} \end{aligned} \tag{31}$$

Following the Bayesian rule, the posterior probability $p\left(\boldsymbol{\mathcal{H}}_u,\boldsymbol{\mathcal{H}}_c|\boldsymbol{\mathcal{X}},\boldsymbol{\Gamma}_u,\boldsymbol{\Gamma}_c,\sigma^2\right)$ can be approximately decomposed as follows:

$$\begin{aligned} p\left(\boldsymbol{\mathcal{H}}_u,\boldsymbol{\mathcal{H}}_c|\boldsymbol{\mathcal{X}},\boldsymbol{\Gamma}_u,\boldsymbol{\Gamma}_c,\sigma^2\right) &\propto p\left(\boldsymbol{\mathcal{X}}|\boldsymbol{\mathcal{H}}_u,\boldsymbol{\mathcal{H}}_c,\sigma^2\right) \\ &\quad \times p\left(\boldsymbol{\mathcal{H}}_u|\boldsymbol{\Gamma}_u\right)p\left(\boldsymbol{\mathcal{H}}_c|\boldsymbol{\Gamma}_c\right) \end{aligned} \tag{32}$$

As expected, $p\left(\boldsymbol{\mathcal{H}}_u,\boldsymbol{\mathcal{H}}_c|\boldsymbol{\mathcal{X}},\boldsymbol{\Gamma}_u,\boldsymbol{\Gamma}_c,\sigma^2\right)$ also complies with complex Gaussian distribution. Nevertheless, it is difficult to process the problem with the tensor form. Here, we adopt a similar route as our previous study [18] (i.e., searching a tensor solution from its vector-based expression). Let $\boldsymbol{\Gamma} = \mathrm{diag}\left(\left[\mathcal{V}_{N_a^uK_d^u}\left(\boldsymbol{\Gamma}_u\right);\mathcal{V}_{N_a^cK_d^c}\left(\boldsymbol{\Gamma}_c\right)\right]\right)$ and $\boldsymbol{C} = \sigma^2\boldsymbol{I} + \boldsymbol{S}\boldsymbol{\Gamma}\boldsymbol{S}^H$. Following separable dictionaries, $\boldsymbol{C}$ can be further decomposed as follows:

$$\boldsymbol{C} = \boldsymbol{C}_c + \boldsymbol{C}_u \tag{33}$$

where

$$\begin{cases} \boldsymbol{C}_u = \dfrac{\sigma^2\boldsymbol{I}}{2} + \mathcal{M}_{N_sK,N_sK}\{\mathcal{G}_{N_sK,N_a^u,K_d^u}\{\boldsymbol{\mathcal{Z}}_u\}\times_2\hat{\boldsymbol{S}}_a^u\times_3\boldsymbol{S}_d^u\} \\ \boldsymbol{C}_c = \dfrac{\sigma^2\boldsymbol{I}}{2} + \mathcal{M}_{N_sK,N_sK}\{\mathcal{G}_{N_sK,N_a^c,K_d^c}\{\boldsymbol{\mathcal{Z}}_c\}\times_2\hat{\boldsymbol{S}}_a^c\times_3\boldsymbol{S}_d^c\} \end{cases} \tag{34}$$

where $\boldsymbol{\mathcal{Z}}_u = \mathcal{T}_{N_s,K,N_a^uK_d^u}\{\boldsymbol{S}_u\boldsymbol{\Gamma}_u\}$ and $\boldsymbol{\mathcal{Z}}_c = \mathcal{T}_{N_s,K,N_a^cK_d^c}\{\boldsymbol{S}_c\boldsymbol{\Gamma}_c\}$. For the vector-based SBL, the solutions of mean $\boldsymbol{\Xi}$ and variance $\boldsymbol{\Sigma}$ are

$$\begin{aligned} \boldsymbol{\Xi} &= \boldsymbol{\Gamma}\boldsymbol{S}^H\boldsymbol{C}^{-1}\mathcal{M}_{N_sK,L}\{\boldsymbol{\mathcal{X}}\} \\ \boldsymbol{\Sigma} &= \left(\sigma^2\boldsymbol{S}^H\boldsymbol{S} + \boldsymbol{\Gamma}\right)^{-1} \end{aligned} \tag{35}$$

Accordingly, $\boldsymbol{\Xi}$ can be expressed with its separable form as follows:

$$\boldsymbol{\Xi} = [\boldsymbol{\Xi}_c;\boldsymbol{\Xi}_u] \tag{36}$$

where

$$\begin{cases} \boldsymbol{\Xi}_c = \boldsymbol{\Gamma}_c\boldsymbol{S}_c^H\boldsymbol{C}^{-1}\mathcal{M}_{NK,L}\{\boldsymbol{\mathcal{X}}\} \\ \boldsymbol{\Xi}_u = \boldsymbol{\Gamma}_u\boldsymbol{S}_u^H\boldsymbol{C}^{-1}\mathcal{M}_{NK,L}\{\boldsymbol{\mathcal{X}}\} \end{cases} \tag{37}$$

As for $\boldsymbol{\Sigma}$, only its diagonal matrix is used in subsequent steps, and it can be expressed as follows:

$$\begin{cases} \boldsymbol{\Sigma}_c = \boldsymbol{\Gamma}_c - \boldsymbol{\Gamma}_c\boldsymbol{S}_c^H\boldsymbol{C}^{-1}\boldsymbol{S}_c\boldsymbol{\Gamma}_c \\ \boldsymbol{\Sigma}_u = \boldsymbol{\Gamma}_u - \boldsymbol{\Gamma}_u\boldsymbol{S}_u^H\boldsymbol{C}^{-1}\boldsymbol{S}_u\boldsymbol{\Gamma}_u \end{cases} \tag{38}$$

Substitute its corresponding tensor processing with Eq. (37). As a result, $\boldsymbol{\Xi}$ can be updated as Eq. (39). $\boldsymbol{\mathcal{W}}_u = \mathcal{T}_{N_a^u,K_d^u,N_sK}\left\{\left[\mathcal{V}_{N_a^uK_d^u}\left(\boldsymbol{\Gamma}_u\right),\cdots,\mathcal{V}_{N_a^uK_d^u}\left(\boldsymbol{\Gamma}_u\right)\right]\right\}$ and $\boldsymbol{\mathcal{W}}_c = \mathcal{T}_{N_a^c,K_d^c,N_sK}\left\{\left[\mathcal{V}_{N_a^cK_d^c}\left(\boldsymbol{\Gamma}_c\right),\cdots,\mathcal{V}_{N_a^cK_d^c}\left(\boldsymbol{\Gamma}_c\right)\right]\right\}$, and the repeatedly copying times is N_sK. A complete derivation of Eq. (39) is given in [18].

$$\begin{aligned} \boldsymbol{\Xi}_u &= \mathcal{T}_{N_a^u,K_d^u,L}\left\{\mathcal{M}_{N_a^uK_d^u,N_sK}\left\{\boldsymbol{\mathcal{W}}_u\odot\left[\mathcal{T}_{N_s,K,N_sK}\left\{\boldsymbol{C}^{-1}\right\}\times_1\left(\boldsymbol{S}_a^u\right)^H\times_2\left(\boldsymbol{S}_d^u\right)^H\right]\right\}\mathcal{M}_{N_sK,L}\left\{\boldsymbol{\mathcal{X}}\right\}\right\} \\ \boldsymbol{\Xi}_c &= \mathcal{T}_{N_a^c,K_d^c,L}\left\{\mathcal{M}_{N_a^cK_d^c,N_sK}\left\{\boldsymbol{\mathcal{W}}_c\odot\left[\mathcal{T}_{N_s,K,N_sK}\left\{\boldsymbol{C}^{-1}\right\}\times_1\left(\boldsymbol{S}_a^c\right)^H\times_2\left(\boldsymbol{S}_d^c\right)^H\right]\right\}\mathcal{M}_{N_sK,L}\left\{\boldsymbol{\mathcal{X}}\right\}\right\} \end{aligned} \tag{39}$$

Subsequently, the sparse coefficient $\boldsymbol{\mathcal{H}}$ can be estimated with $\boldsymbol{\Xi}$. The unknown parameters can be calculated via the expectation maximization (EM) method. The EM method can be described as follows:

$$\boldsymbol{\Theta}^{t+1} = \arg\max_{\boldsymbol{\Theta}} \int \ln p\left(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{\Theta}\right) p\left(\boldsymbol{Y} | \boldsymbol{X}, \boldsymbol{\Theta}^{t}\right) d\boldsymbol{X} \tag{40}$$

where $\boldsymbol{\Theta}$ represents the parameters to be estimated and $\boldsymbol{X}$ is the hidden variable. With the help of Eq. (40), the logarithm object function can be written as

$$\mathcal{Q}\left(\boldsymbol{\Gamma}, \sigma^2\right) \triangleq \mathbb{E}\left\{\ln\left[\begin{array}{l} p\left(\boldsymbol{\mathcal{X}} | \boldsymbol{\mathcal{H}}_u, \boldsymbol{\mathcal{H}}_c, \sigma^2\right) \\ \times p\left(\boldsymbol{\mathcal{H}}_u | \boldsymbol{\Gamma}_u\right) p\left(\boldsymbol{\mathcal{H}}_c | \boldsymbol{\Gamma}_c\right) \end{array}\right]\right\} \tag{41}$$

Substitute Eqs. (31) and (29) into Eq. (41), and the concrete form of Eq.(41) in the proposed method can be further written as follows:

$$\begin{aligned} \mathcal{Q}\left(\boldsymbol{\Gamma}, \sigma^2\right) \propto & -N_s K L \ln \sigma^2 - \frac{1}{\sigma^2} \sum_{l=1}^{L} \mathbb{E}\left\{\left\|\boldsymbol{\mathcal{X}}_{:,:,l} - \boldsymbol{\mathcal{H}}_u^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^u \times_2 \boldsymbol{S}_d^u - \boldsymbol{\mathcal{H}}_c^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^c \times_2 \boldsymbol{S}_d^c\right\|_{\mathrm{F}}^2\right\} - L \ln\left(\prod_{i=1}^{N_a^u} \prod_{j=1}^{K_d^u} \boldsymbol{\Gamma}_u^{i,j}\right) \\ & - \sum_{l=1}^{L}\left[\sum_{i=1}^{N_a^u} \sum_{j=1}^{K_d^u}\left(\mathbb{E}\left\{\left|\boldsymbol{\mathcal{H}}_u^{:,:,l}\right|^2\right\} \oslash \boldsymbol{\Gamma}_u\right)_{i,j}\right] - L \ln\left(\prod_{i=1}^{N_a^c} \prod_{j=1}^{K_d^c} \boldsymbol{\Gamma}_c^{i,j}\right) - \sum_{l=1}^{L}\left[\sum_{i=1}^{N_a^c} \sum_{j=1}^{K_d^c}\left(\mathbb{E}\left\{\left|\boldsymbol{\mathcal{H}}_c^{:,:,l}\right|^2\right\} \oslash \boldsymbol{\Gamma}_c\right)_{i,j}\right] \end{aligned} \tag{42}$$

where

$$\begin{aligned} & \mathbb{E}\left\{\left|\boldsymbol{\mathcal{H}}_u^{:,:,l}\right|^2\right\} = \left|\boldsymbol{\mathcal{H}}_u^{:,:,l}\right|^2 + \boldsymbol{D}_u \\ & \mathbb{E}\left\{\left|\boldsymbol{\mathcal{H}}_c^{:,:,l}\right|^2\right\} = \left|\boldsymbol{\mathcal{H}}_c^{:,:,l}\right|^2 + \boldsymbol{D}_c \\ & \mathbb{E}\left\{\left\|\boldsymbol{\mathcal{X}}_{:,:,l} - \boldsymbol{\mathcal{H}}_u^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^u \times_2 \boldsymbol{S}_d^u - \boldsymbol{\mathcal{H}}_c^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^c \times_2 \boldsymbol{S}_d^c\right\|_{\mathrm{F}}^2\right\} \\ & = \left\|\boldsymbol{\mathcal{X}}_{:,:,l} - \boldsymbol{\mathcal{H}}_u^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^u \times_2 \boldsymbol{S}_d^u - \boldsymbol{\mathcal{H}}_c^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^c \times_2 \boldsymbol{S}_d^c\right\|_{\mathrm{F}}^2 \\ & + \sigma^2\left(\sum_{i=1}^{N_a^u} \sum_{j=1}^{K_d^u}\left(1 - \boldsymbol{D}_u \oslash \boldsymbol{\Gamma}_u\right)_{i,j} + \sum_{i=1}^{N_a^c} \sum_{j=1}^{K_d^c}\left(1 - \boldsymbol{D}_c \oslash \boldsymbol{\Gamma}_c\right)_{i,j}\right) \end{aligned} \tag{43}$$

where $\boldsymbol{D}_u = \mathcal{M}_{N_a^u, K_d^u}\left\{diag\left\{\boldsymbol{\Sigma}_u\right\}\right\}$ and $\boldsymbol{D}_c = \mathcal{M}_{N_a^c, K_d^c}\left\{diag\left\{\boldsymbol{\Sigma}_c\right\}\right\}$. Our goal is to acquire the maximum value of $\mathcal{Q}\left(\boldsymbol{\Gamma}, \boldsymbol{\sigma}^2\right)$, which can be accomplished by taking the first-order derivative of $\boldsymbol{\Gamma}_u$, $\boldsymbol{\Gamma}_c$ and $\boldsymbol{\sigma}^2$. For $\boldsymbol{\Gamma}$, its first-order derivation is as follows:

$$\begin{aligned} \frac{\partial \mathcal{Q}\left(\boldsymbol{\Gamma}, \sigma^2\right)}{\partial \boldsymbol{\Gamma}_u^{i,j}} &= -\frac{1}{\boldsymbol{\Gamma}_u^{i,j}} + \frac{\sum_{l=1}^{L} \mathbb{E}\left\{\left|\boldsymbol{\mathcal{H}}_u^{:,:,l}\right|^2\right\}}{\left(\boldsymbol{\Gamma}_u^{i,j}\right)^2} \\ \frac{\partial \mathcal{Q}\left(\boldsymbol{\Gamma}, \sigma^2\right)}{\partial \boldsymbol{\Gamma}_c^{i,j}} &= -\frac{1}{\boldsymbol{\Gamma}_c^{i,j}} + \frac{\sum_{l=1}^{L} \mathbb{E}\left\{\left|\boldsymbol{\mathcal{H}}_c^{:,:,l}\right|^2\right\}}{\left(\boldsymbol{\Gamma}_c^{i,j}\right)^2} \end{aligned} \tag{44}$$

Setting Eq. (44) equal to zero, $\Gamma_{i,j}$ can be solved as follows:

$$\begin{aligned} \hat{\boldsymbol{\Gamma}}_u^{i,j} &= \frac{1}{L} \sum\nolimits_{l=1}^{L}\left|\boldsymbol{\mathcal{H}}_u^{:,:,l}\right|^2 + \boldsymbol{D}_u^{i,j} \\ \hat{\boldsymbol{\Gamma}}_c^{i,j} &= \frac{1}{L} \sum\nolimits_{l=1}^{L}\left|\boldsymbol{\mathcal{H}}_c^{:,:,l}\right|^2 + \boldsymbol{D}_c^{i,j} \end{aligned} \tag{45}$$

Similar to the processing procedure of $\boldsymbol{\Gamma}$, the derivation of $\mathcal{Q}\left(\boldsymbol{\Gamma}, \sigma^2\right)$ regarding σ^2 is

$$\frac{\partial \mathcal{Q}\left(\boldsymbol{\Gamma}, \sigma^2\right)}{\partial \sigma^2} = \frac{\sum_{l=1}^{L} \mathbb{E}\left\{\left\|\begin{array}{l} \boldsymbol{\mathcal{X}}_{:,:,l} - \boldsymbol{\mathcal{H}}_u^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^u \times_2 \boldsymbol{S}_d^u \\ -\boldsymbol{\mathcal{H}}_c^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^c \times_2 \boldsymbol{S}_d^c \end{array}\right\|_{\mathrm{F}}^2\right\}}{\left(\sigma^2\right)^2} - \frac{N_s K L}{\sigma^2} \tag{46}$$

And updated solution of σ^2 is

$$\begin{aligned} \widehat{\sigma^2} = & \frac{1}{N_s K L} \sum_{l=1}^{L}\left\|\begin{array}{l} \boldsymbol{\mathcal{X}}_{:,:,l} - \boldsymbol{\mathcal{H}}_u^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^u \times_2 \boldsymbol{S}_d^u \\ -\boldsymbol{\mathcal{H}}_c^{:,:,l} \times_1 \widehat{\boldsymbol{S}}_a^c \times_2 \boldsymbol{S}_d^c \end{array}\right\|_{\mathrm{F}}^2 \\ & + \frac{\sigma^2}{N_s K L}\left(\begin{array}{l} \sum_{i=1}^{N_a^u} \sum_{j=1}^{K_d^u}\left(1 - \boldsymbol{D}_u \oslash \boldsymbol{\Gamma}_u\right)_{i,j} \\ + \sum_{i=1}^{N_a^c} \sum_{j=1}^{K_d^c}\left(1 - \boldsymbol{D}_c \oslash \boldsymbol{\Gamma}_c\right)_{i,j} \end{array}\right) \end{aligned} \tag{47}$$

The proposed STSBL algorithm is summarized in Algorithm 2.

E. Preliminary Discussion

1. *Complexity Analysis*

As in many earlier SR-STAP studies, only the multiplication time is considered in the following analysis due to its higher complexity in the application. For the whole processing of the STSBL-STAP method, it includes five different variables estimation i.e., $\boldsymbol{C}$, $\boldsymbol{\Xi}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\Gamma}$ and σ^2. By analyzing the operations of these estimations, the time-consuming steps emerge in the $\boldsymbol{C}$, $\boldsymbol{\Xi}$ and $\boldsymbol{\Sigma}$. Firstly, the calculation of $\boldsymbol{C}_c$ mainly involves $\boldsymbol{\mathcal{Z}}_c$ and two times mode product operations (i.e., $\mathcal{G}_{N_s K, N_a^c, K_d^c}\left\{\boldsymbol{\mathcal{Z}}_c\right\} \times_2 \hat{\boldsymbol{S}}_a^c \times_3 \boldsymbol{S}_d^c$), the required multiplication times of this step is $\left(N_s K\right)^2 N_a^c K_d^c + N_s K N_a^c K_d^c + N_s^2 K K_d^c N_a^c + N_s^2 K^2 K_d^c$, thus its complexity can be approximately evaluated with $\mathcal{O}\left\{\left(N_s K\right)^2 N_a^c K_d^c\right\}$.

Algorithm 2: STSBL-STAP

input : $\widehat{\boldsymbol{S}}_a^u$, $\boldsymbol{S}_d^u$, $\widehat{\boldsymbol{S}}_a^c$, $\boldsymbol{S}_d^c$ and $\boldsymbol{\mathcal{X}}$

Initialize: $\boldsymbol{\Gamma}_c = 1$, $\boldsymbol{\Gamma}_u = 1$ and $\sigma^2 = 10$;
for $p = 1$ **to** $P_{\max}$ **do**
  Calculate $\boldsymbol{C}_u$ and $\boldsymbol{C}_c$ by eqn. (34);
  Calculate $\boldsymbol{C}$ by eqn. (33);
  Calculate $\boldsymbol{\Xi}_c$ and $\boldsymbol{\Xi}_u$ by eqn. (39);
  Calculate $\boldsymbol{\Sigma}_c$ and $\boldsymbol{\Sigma}_u$ by eqn. (38);
  Calculate $\boldsymbol{\Gamma}_c$ and $\boldsymbol{\Gamma}_u$ by eqn. (45);
  Calculate σ^2 by eqn. (47);
  if $\left\|[\boldsymbol{\Xi}_c^p; \boldsymbol{\Xi}_u^p]\right\|_{\mathrm{F}}^2 / \left\|\left[\boldsymbol{\Xi}_c^{p-1}; \boldsymbol{\Xi}_u^{p-1}\right]\right\|_{\mathrm{F}}^2 \leq \mu$ **then**
    break;
Calculate $\hat{\boldsymbol{R}}$ based on eqn. (21) ;

output: $\hat{\boldsymbol{R}}$

TABLE I: The complexity of different methods

Methods	Complexity
SBL	$\mathcal{O}\left\{\left(N_a^u K_d^u + N_a^c K_d^c\right)^3\right\}$
STSBL	$\mathcal{O}\left\{\max\left(N_a^u K_d^u, N_a^c K_d^c\right)(N_s K)^2\right\}$

As for $\boldsymbol{C}_u$ except for the dictionary, which has the same processing flow as $\boldsymbol{C}_c$, hence the complexity of it is $\mathcal{O}\left\{(N_s K)^2 N_a^u K_d^u\right\}$. Now, different from the previous research work, benefit from the advantages of separable property, the final complexity of $\boldsymbol{C}$ is $\mathcal{O}\left\{\max\left(N_a^u K_d^u, N_a^c K_d^c\right)(N_s K)^2\right\}$. Then for $\boldsymbol{\Xi}_c$, two times mode product (i.e.,$\mathcal{T}_{N_s,K,N_s K}\left\{\boldsymbol{C}^{-1}\right\} \times_1 \hat{\boldsymbol{S}}_a^c \times_2 \boldsymbol{S}_d^c$), one time matrix multiplication and Hadmard product are executed in this step, the total multiplication time is $N_s^2 K^2 N_a^c + N_s K^2 N_a^c K_d^c + N_s K N_a^c K_d^c + N_s K L N_a^c K_d^c$, because $K \gg L$ in SR-STAP, hence its complexity is $\mathcal{O}\left\{N_s K^2 N_a^c K_d^c\right\}$. Similar to the operation of $\boldsymbol{\Xi}_c$, the complexity of $\boldsymbol{\Xi}_u$ can be assessed with $\mathcal{O}\left\{N_s K^2 N_a^u K_d^u\right\}$, and thus the complexity of the $\boldsymbol{\Xi}$ is $\mathcal{O}\left\{\max\left(N_a^u K_d^u, N_a^c K_d^c\right) N_s K^2\right\}$. In the estimation of $\boldsymbol{\Sigma}_c$ and $\boldsymbol{\Sigma}_u$, it should be noted that only the diagonal elements are used in a subsequent operation, hence the complexity of this step is $\mathcal{O}\left\{\max\left(N_a^u K_d^u, N_a^c K_d^c\right)(N_s K)^2\right\}$. From the above analysis, it can be obtained that the complexity of the proposed STSBL method is $\mathcal{O}\left\{\max\left(N_a^u K_d^u, N_a^c K_d^c\right)(N_s K)^2\right\}$, which is far smaller than the conventional SBL method with large-scale dictionary $S \in {}^{N_s K \times (N_a^u K_d^u + N_a^c K_d^c)}$ (i.e., $\mathcal{O}\left\{\left(N_a^u K_d^u + N_a^c K_d^c\right)^3\right\}$).The complexity of different methods is summarized in Table I.

2. General Application

The STSBL could use a novel separable processing structure, and thus, it can effectively overcome the shortage of the conventional SBL methods mentioned above. Note that the core idea of the separable structure is dictionary segmentation. Concretely, in SR-STAP applications, this means the regional division of the angle-Doppler plane.

- ***Remark 1***: The selection of dividing methods can be regular or irregular, which has no impact on the final estimation accuracy but will increase the pressure on multicore processing. This property gives great flexibility to the application.
- ***Remark 2***: In segmentation, there is no definitive number constraint of a separable dictionary grid. In other words, when the atom number of separable dictionaries is one, the above structure can still work effectively.

Fig. 8 presents an example of segmentation in 2D space, where a 6×6 grid set is divided into nine areas. The proposed division idea is not limited to the space–time two-dimensional plane, and it is suitable for applying multidimensional data, such as 3D echo data (azimuth–elevation–Doppler).

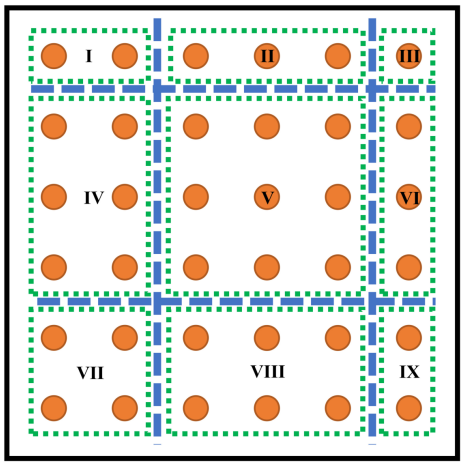


Fig. 8: Region segmentation of the angle-Doppler grids.

The STSBL processing frame can guarantee that each subdictionary can perform individually. An alternative processing structure is shown in Fig. 9, and the entire processing structure contains two stages. In the first stage, Q block dictionaries are input to the multicore system and $\boldsymbol{C}$ can be solved by the sum from $\boldsymbol{C}_1$ to $\boldsymbol{C}_Q$. In the second stage, the original data $\boldsymbol{S}$ is separated into Q block and input into every core to calculate $\boldsymbol{\Xi}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\Gamma}$ and σ^2. It is shown that the most time-consuming step is the calculation of $\boldsymbol{\Sigma}$, which is $\mathcal{O}\left(N_a K_d (N_s K)^2\right)$. Compared with the TSBL, the advantages of the STSBL in application is the separable processing structure; hence, its speed-up ratio can be approximately evaluated with its processor number, that is, $\mathcal{O}\left(N_a K_d (N_s K)^2 / N_{\text{core}}\right)$, where N_{core} is the number of cores. However, it is important to point out that this ratio is ideal because it ignores the effects of real hardware. In other words, the real design should also consider an appropriate core

number according to actual demand. Introducing this processing technology (e.g., multicore DSP and FPGA), the STSBL parallel processing structure can achieve a further significant improvement in computation loading compared with the TSBL.

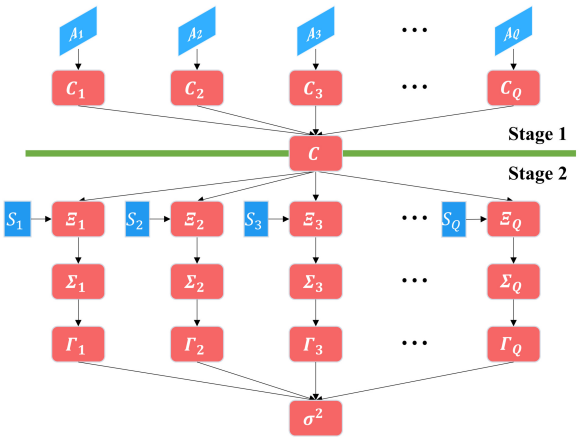


Fig. 9: Generalized parallel processing structure of STSBL.

V. Wide-Area GMTI with Fast Subarray SR-STAP

The WGMTI working mode aims to detect potential moving targets and simultaneously mark them in the SAR image; thus, it puts a high-level real-time processing requirement on the hardware system and processing algorithms. Under these constraints, conventional SAR imaging algorithms, such as the range–Doppler algorithm (RDA), chirp-scaling algorithm, and ωk algorithm, are no longer suitable for the WGMTI application. A trade-off selection is Doppler beam sharpening (DBS), which has a simple processing structure with the help of FFT algorithms. The whole WGMTI processing framework of our developed radar system is depicted in Fig. 10, which mainly contains four parts: data preprocessing, DBS SAR imaging, clutter suppression and moving target detection, and moving target relocation. Subsequently, the realization of each stage will be briefly discussed. It is necessary to clarify that in this paper, we focused on the clutter suppression stage. For the detailed realization of the other stage, the reader can refer to the relative literature in each subsection.

A. Data Preprocessing

Influenced by the real environment and nonideal radar system configuration, the actual Doppler center frequency deviates from the ideal setting. Thus, the actual Doppler center frequency must be estimated first [47]. Correlation function estimation and energy balance are used as the estimation methods for the Doppler center frequency. Compared with the second type, the correlation function is often applied in real radar systems because of its simple implementation. Therefore, we adopt the correlation function method in this step. The next step is pulse compression. The purpose of this step is to improve the range resolution by compressing the received LFM signal.

After pulse compression, a channel calibration technique should be performed to compensate for the unknown gain/phase error between different received channels. Note that the performance of this step has a vital impact on subsequent STAP clutter suppression. The classic calibration is the adaptive channel balancing method [48], which can be executed alternately. In most cases, three iterations could meet the application requirement.

B. DBS Imaging and Stitching

Before the azimuth FFT operation, a range migration correction (RMC) step should be performed to compensate for the range cell change in moving targets caused by the long-time observation [49]. Nevertheless, when the target observation time is too short to cross a single-range cell, the RMC step can be ignored to reduce complexity. Subsequently, the azimuth FFT operation is applied in the points pulse signal to generate the Doppler frequency spectrum. Note that the ground clutter in the DBS imaging stage represents the real geomorphology environment; thus, it should be retained in the DBS imaging stage. The next step is extracting the main Doppler region. The reason for this operation is that only the main beam illumination area has sufficient CNR to reserve a sharp outline of the terrain. In an ideal case, a 3 dB width of the main Doppler clutter area is extracted in imaging. However, in the application, it usually adjusts a little bigger than 3 dB to ensure that all powerful areas are covered in processing. The final step is image stitching. This operation mainly involves geometric mapping and insertion. This operation maps the ground point from the range–Doppler space to the geometric coordinate system. It should be noted that the above is the result of the original SAR image. For subsequent SAR image analysis, some image preprocessing is required in the application [50].

C. Clutter Suppression and Moving Target Detection

For an airborne radar, strong ground clutter would significantly affect the performance of the detection of moving targets. Therefore, clutter must be eliminated before the detection of the target. The SR-STAP method was used to execute this operation. The processing flow can be roughly divided into two steps: sample selection and filtering. For the filter step, the proposed STSBL-STAP was used to calculate the optimal filter coefficient, which has been given an exhaustive description in Section III. The sample selection strategy takes the sliding window. For other window strategies, refer to [51].

The next step is the detection of moving targets. In this step, it is crucial to choose an appropriate constant false

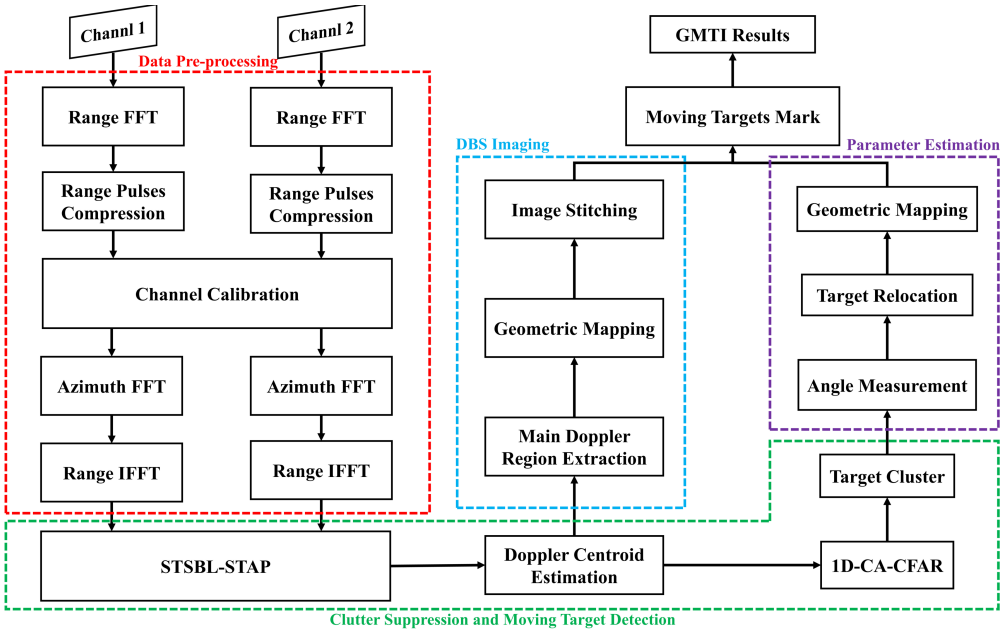


Fig. 10: The WGMTI processing framework.

alarm rate (CFAR) algorithm following different application environments. Although numerous CFAR algorithms have been developed to process different application scenarios [52], in real applications, especially in large input data size (e.g., hundreds of pulses and thousands of range gates) conditions, too many complicated CFAR methods are impossible to realize in real-time processing. Thus, the cell averaging CFAR (CA-CFAR) is still an optimal choice in the WGMTI mode. The common CA-CFAR has two main types: one-dimensional and two-dimensional CA-CFAR. 1D-CA-CFAR was performed in the range dimension, and its computational loading was the lowest of these types. Compared to 1D-CA-CFAR, 2D-CA-CFAR works in the 2D range–Doppler plane and requires a huge computational load. A compromise of 2D-CA-CFAR is to use the "+" window, which solves the threshold with the elements of the cross part; thus, computation can be further reduced. Additionally, a separable 1D-CA-CFAR idea can be useful in engineering [53], which cleverly decomposes 2D-CA-CFAR as two pieces of 1D-CA-CFAR and individually performs in range dimension and Doppler dimension. Considering computational effort and time, 1D-CA-CFAR is chosen in WGMTI.

D. Target Parameter Estimation and Relocation

The target range can be directly inversed through the location detected in the range–Doppler spectrum, but the azimuth angle of moving targets cannot be calculated similarly because of the unknown target initial velocity. Thus, it needs to measure the actual azimuth angle to relocate the moving targets in the true ground location. The commonly used methods in engineering are monopulse angle measurement [54] and the maximum likelihood method [3]. In these methods, the maximum likelihood method has the highest measurement accuracy, but its large DoFs and huge calculation requirements are unsuitable for the application of the dual-channel radar system WGMTI mode. Fortunately, the accuracy of the measurement of the conventional monopulse angle is sufficient for application. It is easy to implement in engineering. Once the actual azimuth angle of moving targets is obtained, relocation can be executed in the next step of measurement, marking the moving targets in the DBS image. Here, single-cycle WGMTI processing has been completed.

VI. Simulation Experiments

In this section, we mainly verify the performance of STSBL-STAP in the subarray-level application, and some performance evaluation experiments were carried out in the form of simulations. The raw echo data were generated based on the multichannel airborne radar model that was described in Section II, and some relative radar system parameters are summarized in the following: the center frequency f_c was Ku-band; the platform altitude and velocity were 2000 m and 60 m/s, respectively; the element spacing was $\lambda/2$; the PRF was set equal to $4v/\lambda$; and the pulse number K, azimuth element number N and elevation element number were 32, 24, and 6, respectively. During the simulation, the radar main beam was adjusted to point $\theta = 90^\circ$ and $\varphi = 5^\circ$. Each range cell was uniformly divided into 180 patches, and each was satisfied with the Weibull distribution. The CNR here was set at 60 dB, which corresponds to the real environment. Additionally, the SBL results were used for the comparison. All algorithms were executed with a similar processing flow in MATLAB R2020b. The

hardware configuration was a CPU eight core i9-9900KF 3.6GHz 460GFLOPS with 64 GB RAM.

A. Space-time Spectrum Estimation

In this experiment, the CCM estimation accuracy was verified by the joint dictionary under the subarray synthetic condition. Three common looking modes are chose in here, which are side-looking mode ($\theta_p = 0°$), squint-looking mode ($\theta_p = 30°$) and forward-looking mode ($\theta_p = 90°$). For a joint dictionary, let the total grid number be close to the β was 4 (i.e., the number of grids $N_a K_d$ was 12513). A minimum variance distortionless response spectrum was applied to evaluate the performance of recovery, which can be achieved as follows:

$$\boldsymbol{P}\left(f_s, f_d\right) = \frac{1}{\left|\boldsymbol{s}^H\left(f_s, f_d\right) \tilde{\boldsymbol{R}}^{-1} \boldsymbol{s}\left(f_s, f_d\right)\right|} \tag{48}$$

Unless otherwise stated, the training sample L was 6, and the initialization parameter values of the SBL-type algorithm throughout the experiments were the following: $\sigma^2 = 10$, $\mu = 10^{-3}$ and $P_{max} = 50$.

Fig. 11 shows the spectrum estimation result of two synthesis channel numbers (i.e., $N_s = 24$ and 2). Compared to the joint dictionary spectrum estimation result with the original clutter space-time spectrum, it can be known that the proposed method can achieve a good performance at the element-level. Then, with gradually decreasing synthesis channel numbers, as in the analysis before, the number of grating-lobe clutter ridges is rapidly increasing. In this scene, the joint dictionary can still hold an acceptable estimation accuracy. This is benefited from the property that the joint dictionary can adjust denser sampling in the main clutter area. The above result implies that the joint dictionary is effective in channel-level SR-STAP applications.

B. The Signal-to-Clutter-Plus-Noise Ratio (SCNR) Loss

In this section, the clutter suppression ability of the different synthesis numbers is further tested using the STSBL-STAP method. Additionally, a commonly used post-Doppler STAP method [51] in engineering is added for better comparison, in which the three Doppler channels are jointly utilized in processing. In addition, according to the RMB rule, setting $L = 30$ and 160 respectively represent the sufficient sample condition of the Post-Doppler method at the channel-level and element-level. And the SCNR loss is used as a performance evaluation criteria [9], that is,

$$\text{SCNR Loss} = \frac{\sigma_n^2}{NK} \frac{\left|\boldsymbol{w}^H s_o\right|^2}{\left(\boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w}\right)} \tag{49}$$

The parameter setting is the same as in the last experiment, and to reduce the effect of random fluctuation, all curves are the average result of multiple Monte Carlo trials.

TABLE II: Grid size

	SBL	STSBL
$\beta = 2$	$N_{\text{sbl}} = 3185$	$N_c = 1562, N_u = 1530$
$\beta = 3$	$N_{\text{sbl}} = 7081$	$N_c = 4046, N_u = 3185$
$\beta = 4$	$N_{\text{sbl}} = 12513$	$N_c = 6864, N_u = 6120$
$\beta = 5$	$N_{\text{sbl}} = 19481$	$N_c = 9686, N_u = 9605$
$\beta = 6$	$N_{\text{sbl}} = 27985$	$N_c = 13370, N_u = 13635$

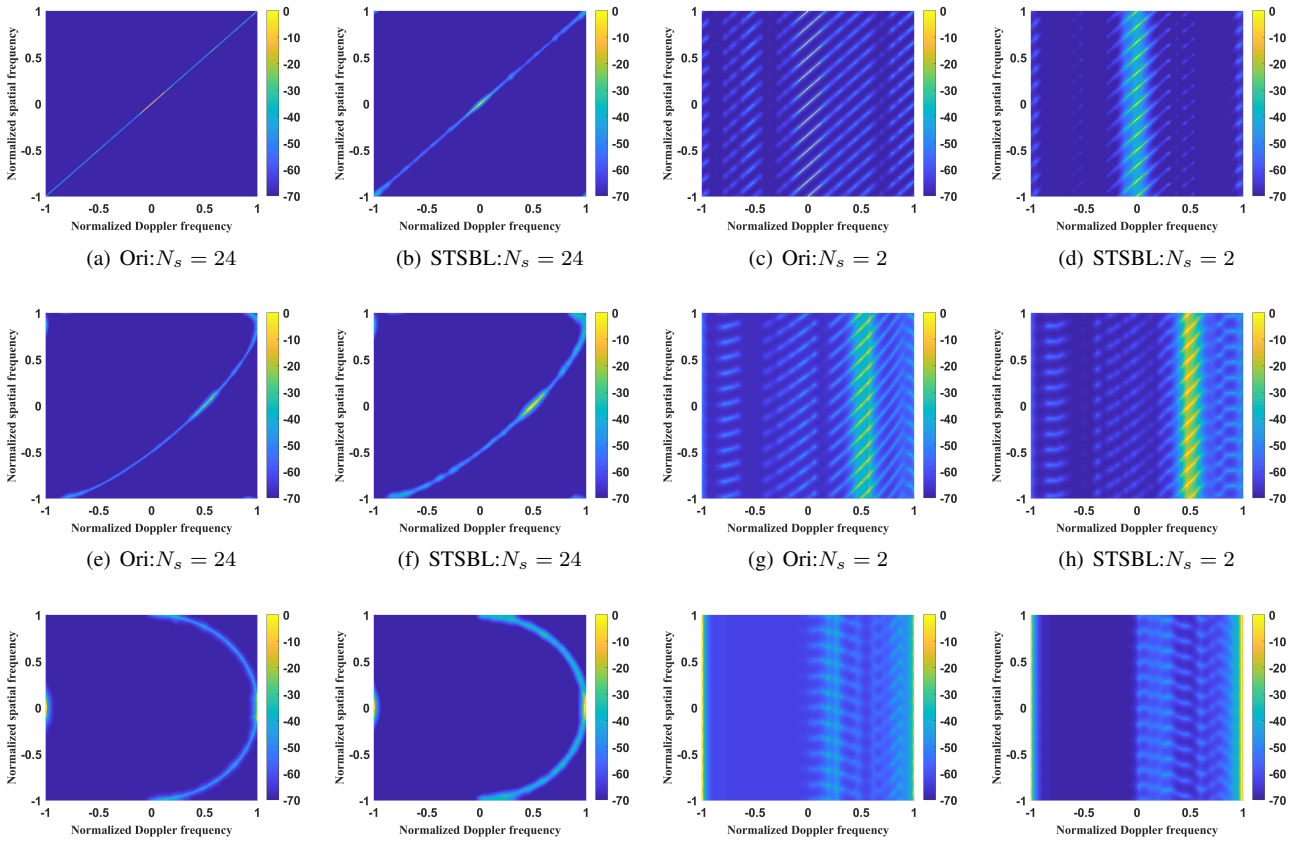
Fig. 12shows the SCNR loss results of different methods in three looking modes. The post-Doppler STAP had the worst clutter performance in the severe shortage of training samples. With increasing the sample number, its performance has significant improvement, however, even in this condition, the performance of the Post-Doppler method is still lower than SR-STAP method. Conversely, the SR-STAP methods have excellent clutter suppression ability even in such sample conditions. STSBL has a narrower notch than SBL due to an appropriate sampling strategy. Subsequently, with the subarray synthesis ($N_s = 2$), Fig. 12(d)-(f) gives the SCNR loss result of different methods in $N_s = 2$. As analyzed earlier, grating-lobe clutter causes performance degradation of STAP methods. Nevertheless, the SBL and STSBL methods are still effective at the subarray level. Furthermore, it can be found that STSBL-STAP is several dB higher than SBL whether in the main or non-main clutter region, which illustrates the validity of the STSBL method.

C. Running Time

In the final simulation experiment, the computation loading of the proposed STSBL method was assessed with its running time. As listed in Table II, under the condition of $N_s = 2$, $N = 24$, $K = 32$ and $L = 6$, the size of a conventional space-time dictionary is set from $\beta = 2$ to 6 (i.e. $N_{sbl} = 3185$ to 27985). Subsequently, to make a fair comparison, the joint dictionary is adjusted such that the number of atoms is close to its corresponding conventional dictionary. As shown in Fig. 13, the speed of STSBL was significantly reduced compared with the SBL method. For example, SBL requires 9540.632 s in $\beta = 6$, but STSBL only needs 2.655 s, which is approximately 3600 times faster than SBL. Our test here is only executed in MATLAB software. The powerful advantages of the proposed separable structure can be further exploited in processing tools, such as DSP and FPGA.

VII. Real Measurement Experiments

In this section, we used real measurement data to verify the clutter suppression performance of the proposed STSBL in WGMTI processing. This flight experiment was performed in October 2021, and the observation area was in Xiangyang, a city in the province of Hubei in China. Fig. 14 shows the flight route, which begins at $\left(112°54'\text{E}, 31°975'\text{N}\right)$ and ends at



(a) Ori:$N_s = 24$ (b) STSBL:$N_s = 24$ (c) Ori:$N_s = 2$ (d) STSBL:$N_s = 2$

(e) Ori:$N_s = 24$ (f) STSBL:$N_s = 24$ (g) Ori:$N_s = 2$ (h) STSBL:$N_s = 2$

(i) Ori:$N_s = 24$ (j) STSBL:$N_s = 24$ (k) Ori:$N_s = 2$ (l) STSBL:$N_s = 2$

Fig. 11: Spectrum estimation results of different methods. (a)-(d) side-looking mode, (e)-(h) squint-looking mode, (i)-(l) forward-looking mode.

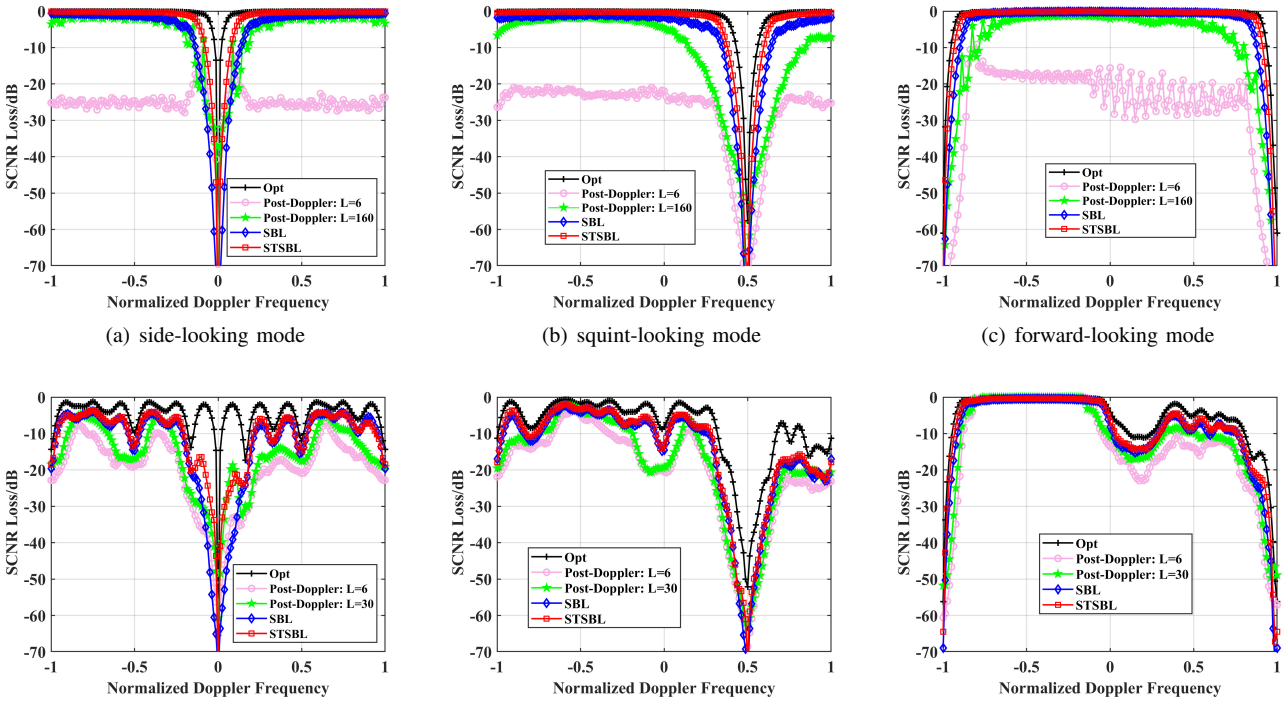
TABLE III: Radar system parameters

Parameters	Symbols	Values
Carrier frequency	f_c	Ku-band
Bandwidth	B	20MHz
PRF	f_r	3000Hz
Pulse number	K	256
Channel number	N_s	2
Platform altitude	H	2500m
Platform velocity	v_p	60m/s
Azimuth angle	θ_o	90°
Elevation angle	φ_o	10°

$\left(112°63'\mathrm{E}, 31°975'\mathrm{N}\right)$. The Ku-band echo data were recorded with our developed dual-channel radar system, and the main radar system parameters are summarized in Table III. As shown in Fig. 15, the radar was installed under the airborne platform, and a special design radome was applied to avoid the impact of the outside environment. The antenna array was a UPA with 16 elements in a row and 32 elements in a column. The received original echo data of these elements were directly synthesized by double channel output. In the flight experiment, the radar worked in the side-looking mode and illuminated the left area of the airplane platform. The real measurement data from the 77th cycle were selected for the experiment, composed of 15 beam positions. Every beam position was extracted from 1000 range gates for testing. The following is a formal analysis.

Fig. 16(a) shows the range–Doppler spectrum of echo data. The strong clutter occupies the most Doppler bins in the entire two-dimensional space, and only the powerful targets in the pure area can be effectively identified in the original data. Subsequently, Fig. 16(b) gives a phase image of this beam position. After the operation of channel balance, there is good channel consistency between Channels 1 and 2, which provides a good basis for subsequent clutter suppression. In STAP, a successive sliding window (from range bin 1 to 1000) is taken to select the training samples. If there were no special instructions, four samples were used for STSBL-STAP and post-Doppler STAP. Other parameters' initialization is the same as in the simulation experiments.

In the first experiment, the estimation ability of the STSBL in real measurement data was evaluated following the space–time spectrum results of the different samples. Fig. 17(a)-(c) are the 566th range gate estimation results of the traditional maximum likelihood (ML) methods ($\hat{R} = XX^H/L$) with $L = 4$, $L = 60$ and $L = 1000$.



(a) side-looking mode (b) squint-looking mode (c) forward-looking mode

(d) side-looking mode (e) squint-looking mode (f) forward-looking mode

Fig. 12: The SCNR loss results from different methods: (a)-(c) full element and (d)-(f) $N_s = 2$.

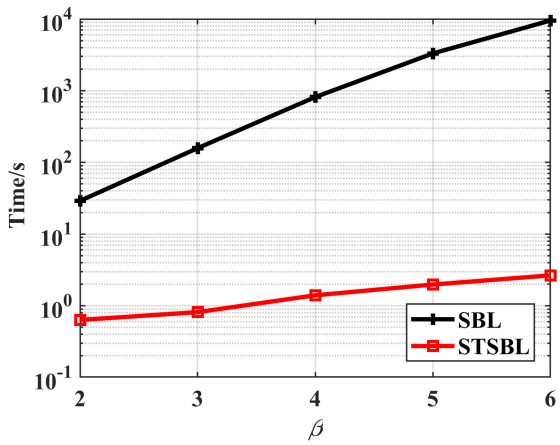


Fig. 13: The running time of different methods.

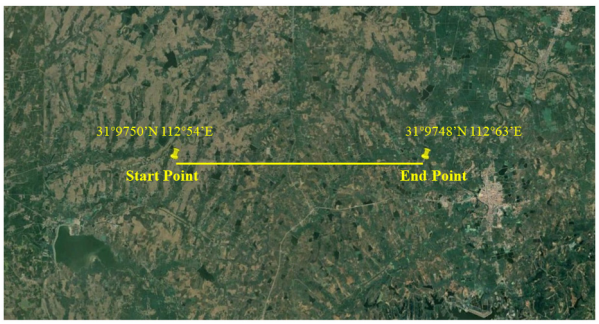


Fig. 14: The diagram of flight route.

The ML method has a slow convergence speed; thus, it cannot achieve a good estimation with a small sample number. Once the sample number is adequate (see Fig. 17(c)), the actual space–time clutter distribution is similar to the simulation result in Section V. Compared to ML (Fig. 17(d), the STSBL can effectively estimate the clutter even in $L = 4$, and the accuracy continuously rises with the addition of the samples (see Figs. 17(e) and (f)). This means that the sparse recovery is suitable for the dramatic change in environment application.

In the second experiment, we checked the clutter suppression performance of STSBL-STAP with the average Doppler residual power:

$$\boldsymbol{P}_r(k) = \frac{1}{L}\sum_{l=1}^{L}\boldsymbol{P}(l,k) \tag{50}$$

As shown in Fig. 18, the original main clutter area is approximately 60 Doppler bin width. Additionally, it should be noted that sidelobe clutter severely affects target detection. After clutter suppression, post-Doppler STAP and STSBL-STAP can effectively reduce the main clutter area to approximately $\Delta K \approx 40$ Doppler bin, which is less than its original result. The nonmain clutter area, such as the sidelobe clutter, is also well suppressed by STAP. Comparing these two methods, they can have a similar performance in the main clutter area. However, in other areas, STSBL-STAP seems to have a better result than post-Doppler STAP (performance improvement of

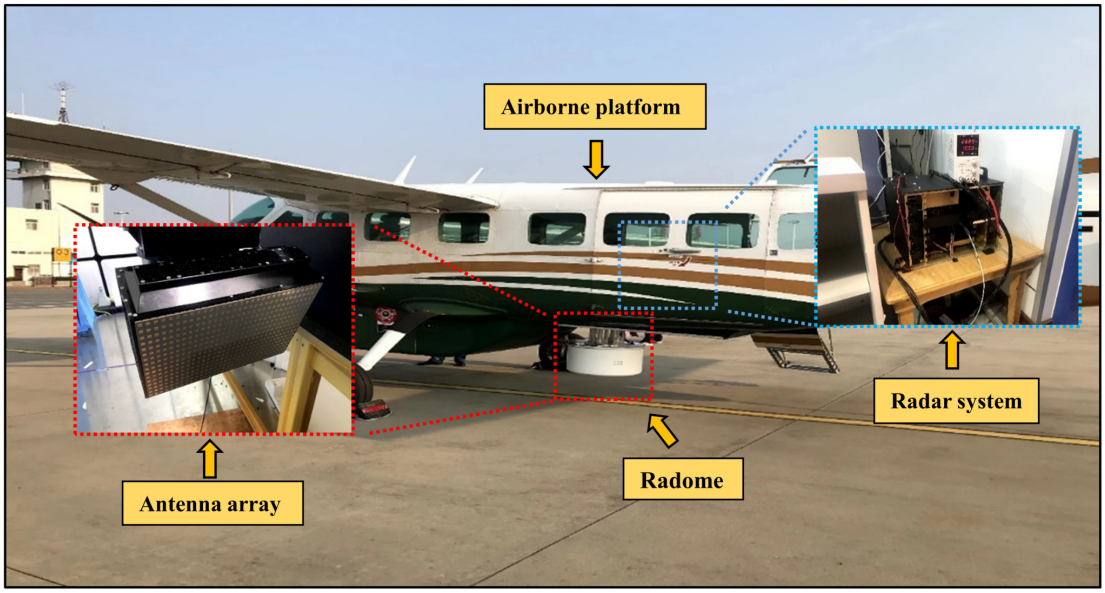
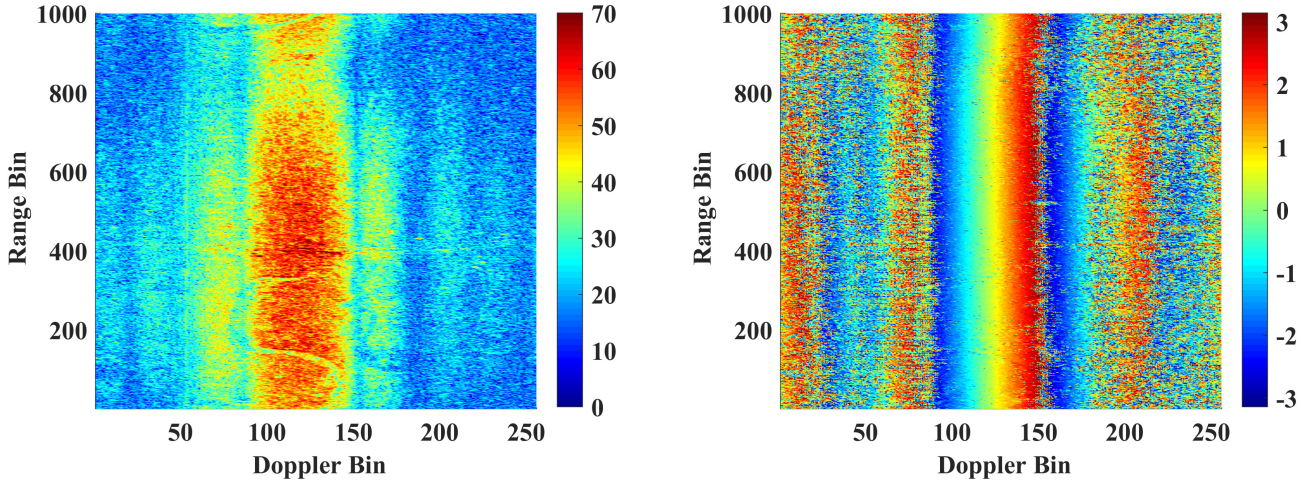


Fig. 15: The flight experiment platform and radar system.

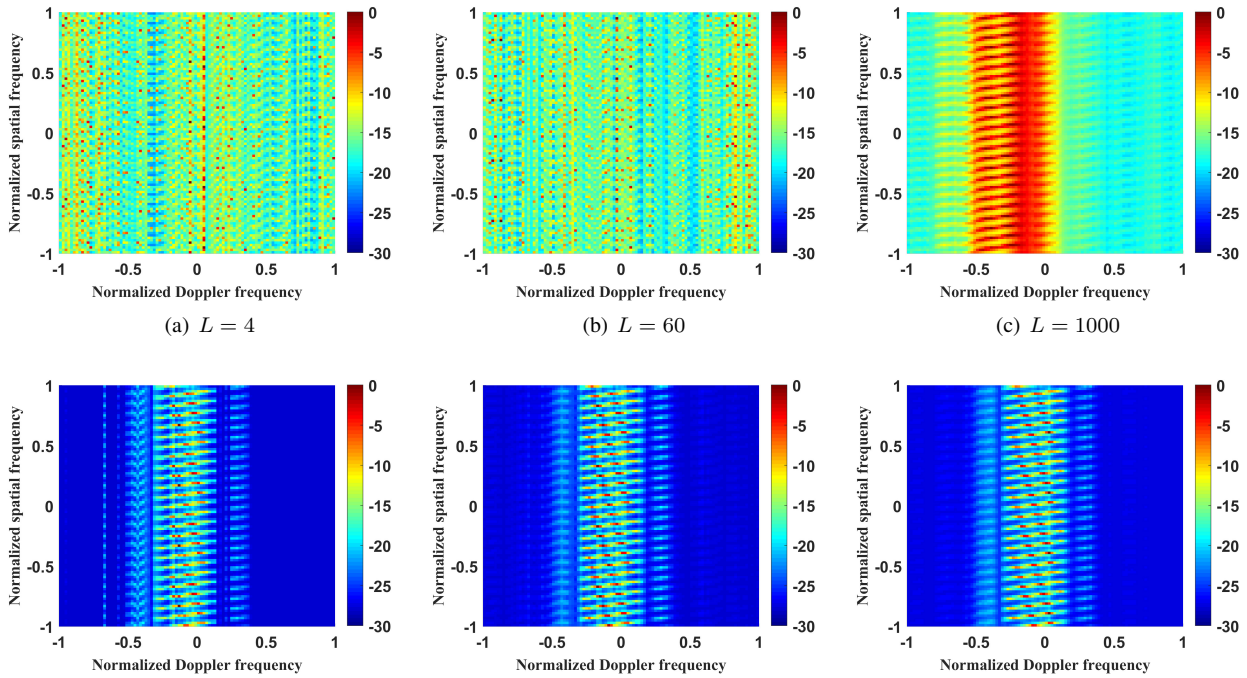


(a) (b)

Fig. 16: The results of echo data. (a) range-Doppler spectrum. (b) the phase difference between two channels.

several dB). This is because four training samples for post-Doppler STAP are inadequate in the application.

To clarify the clutter suppression performance of STSBL, the whole range-Doppler processing results of different methods are displayed in Fig. 19. Through the clutter suppression operation, it can be seen that the main clutter has a significant reduction. Compared the STSBL-STAP methods with the conventional Post-Doppler methods, it has a better suppression in sidelobe clutter area than other methods. Then for better comparison, the local area results are given in Fig. 20. Due to the heterogeneity of the range training sample, the Post-Doppler directly lost the effective target components. As for the STSBL-STAP method, it can still hold an effective target while suppressing the clutter. The above results demonstrate that STSBL-STAP is highly suitable for the heterogeneous environment. In addition, it should be pointed out that it seems that the Post-Doppler method has a similar clutter suppression ability in the $L = 4$ and $L = 30$. However, as shown in Fig. 21(a), the result of post-Doppler STAP processing in sample shortage would generate many isolate peaks after clutter suppression, which may cause lots of false target points. This is because the estimation of conventional STAP methods has latent instability given a shortage of training samples. When more samples are supplemented, as shown in Fig. 21(b), the processing result becomes smooth compared to Fig. 21(a).



(a) $L = 4$ (b) $L = 60$ (c) $L = 1000$

(d) $L = 4$ (e) $L = 60$ (f) $L = 1000$

Fig. 17: The spectrum estimation of different training sample numbers: (a)-(c) are the results of the original echo, and (d)-(f) are the results of STSBL.

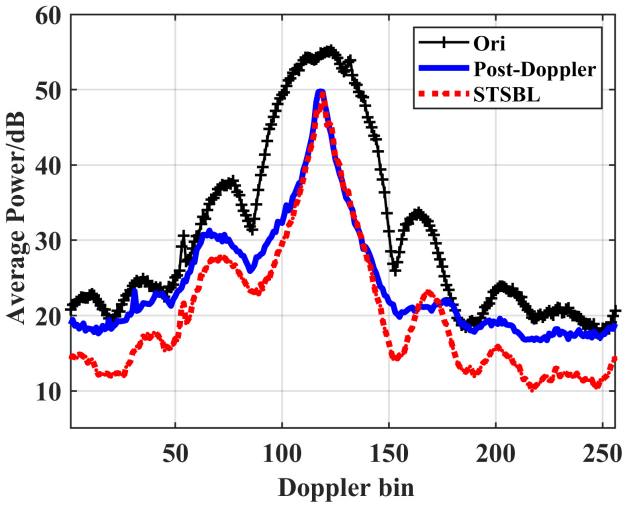
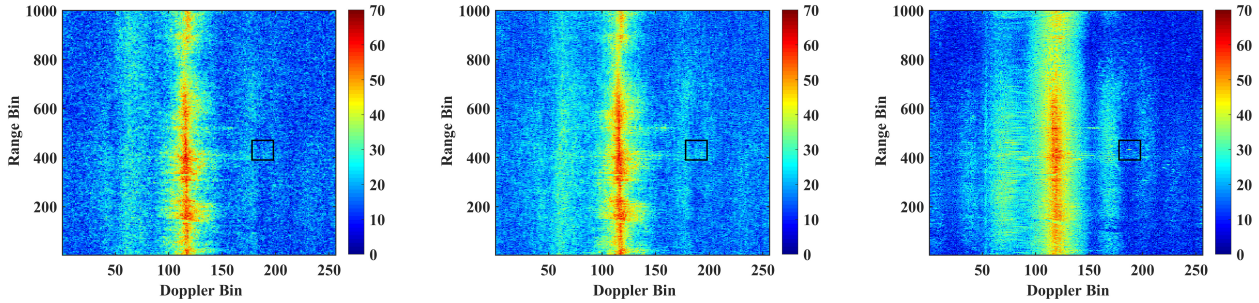


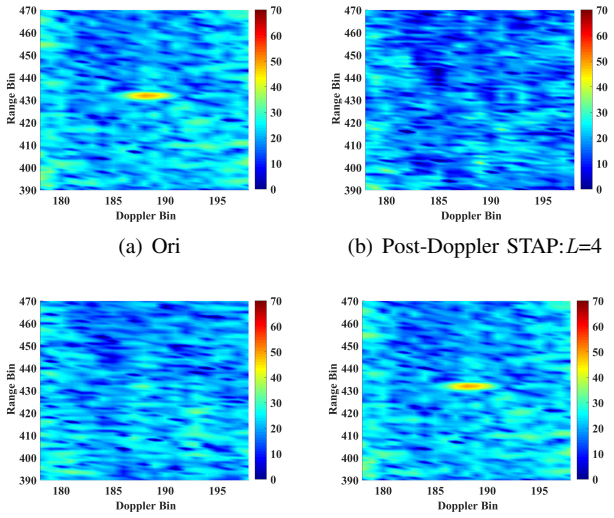
Fig. 18: Average Doppler residual power.

Finally, to evaluate the effectiveness of STSBL-STAP in GMTI processing, the whole beam positions echo data of the 77th cycle were used for the experiment. According to the radar system parameters, the range resolution was 6 m and the whole imaging range was approximately 6 km. Fig. 22(b) shows the wide-area DBS image stitching result, and Fig. 22(a) is a screenshot of the real environment from Google Earth. Compared to the DBS image result with its corresponding real environment, the basic image texture is clear and satisfies the requirement of the WGMTI mode. It should be pointed out that contrary to conventional SAR imaging, WGMTI requires fast imaging and moving target detection in a wide area; thus, the image is often rough. Subsequently, Fig. 22(c) shows the WGMTI processing result without clutter suppression. Many isolated moving targets are shown to distribute throughout the whole image area; this is a normal phenomenon in WGMTI processing. The reasons are as follows: (1) to obtain accurate ground information, the CFAR threshold is set a little lower than that of the other modes; and (2) for a large imaging area, lots of moving targets may exist. In WGMTI, we are concerned about the moving targets that appear on the road. Fig. 22(d) presents the WGMTI results after clutter suppression. Compared with Fig. 22(c), it can be seen that many false alarm targets caused by strong clutter are rejected in Fig. 22(d). A local area detection result of the road area is extracted in Fig. 23. Fig. 23(a) is a local screenshot of the real environment, and Fig. 23(b) is the moving target detection result without clutter suppression processing. Fig. 23(c) and (d) are the results of post-Doppler STAP and STSBL-STAP, respectively. As previously analyzed, many false alarm targets are detected in Fig. 23(c) because of the unstable estimation. From a comprehensive perspective, the STSBL-STAP method performs better than the traditional post-Doppler STAP method in small sample conditions



(a) Post-Doppler STAP:L=4 (b) Post-Doppler STAP:L=30 (c) STSBL-STAP

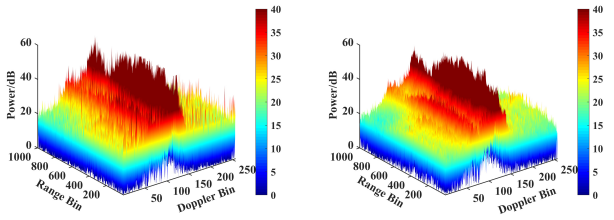
Fig. 19: The clutter suppression results of different methods.



(a) Ori (b) Post-Doppler STAP:L=4

(c) Post-Doppler STAP:L=30 (d) STSBL

Fig. 20: The local clutter suppression results of different methods.



(a) $L = 4$ (b) $L = 30$

Fig. 21: The range-Doppler results of post-Doppler STAP in different sample numbers.

VIII. Conclusions

In this paper, a general subarray-level SR-STAP framework was proposed for multichannel airborne radar applications. Specifically, a joint space–time dictionary was designed, combining the subarray-level clutter property to achieve improved recovery. A novel separable SR-STAP structure was presented on this basis. This structure aims to decompose the large-size optimal problem into multiple suboptimal problems and reduce the data dimension for fast processing. Therefore, the proposed method can be applied in practical engineering through multi-core programming technology. The processing results of the real measurement data obtained from our developed airborne radar illustrate that in WGMTI application, the proposed STSBL-STAP method has significant advantages compared to the conventional post-Doppler STAP in small sample conditions. In the future, the engineering implementation of the proposed STSBL-STAP method in WGMTI will be our main focus in follow-up studies.

REFERENCES

[1] D. Cerutti-Maori, J. Klare, A. Brenner, and J. Ender, "Wide-area traffic monitoring with the SAR/GMTI system PAMIR," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 3019–3030, Oct. 2008.

[2] E. Makhoul, S. V. Baumgartner, M. Jäger, and A. Broquetas, "Multichannel SAR-GMTI in maritime scenarios with F-SAR and TerraSAR-X sensors," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 11, pp. 5052–5067, Nov. 2015.

[3] C. Song, B. Wang, M. Xiang, et al., "A General Framework for Slow and Weak Range-Spread Ground Moving Target Indication Using Airborne Multichannel High-Resolution Radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022, doi: 10.1109/TGRS.2022.3174858.

[4] D. Cerutti-Maori and I. Sikaneta, "A generalization of DPCA processing for multichannel SAR/GMTI radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 560–572, Jan. 2013.

[5] A. Budillon, C. H. Gierull, V. Pascazio, and G. Schirinzi, "Alongtrack interferometric SAR systems for ground-moving target indication: Achievements, potentials, and outlook," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 2, pp. 46–63, Jun. 2020.

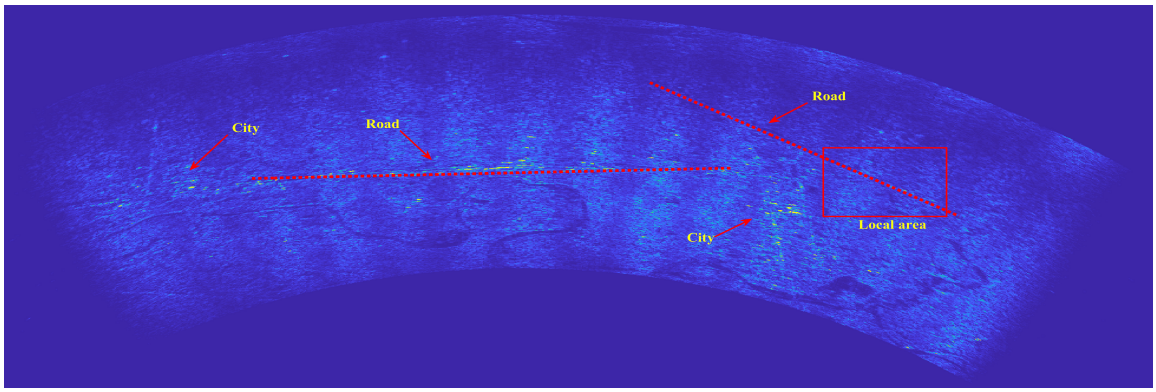
[6] Z. Li, H. Ye, Z. Liu, et al., "Bistatic SAR clutter-ridge matched STAP method for nonstationary clutter suppression," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022, doi: 10.1109/TGRS.2021.3125043.

[7] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, no. 6, pp. 853–863, Nov. 1974.

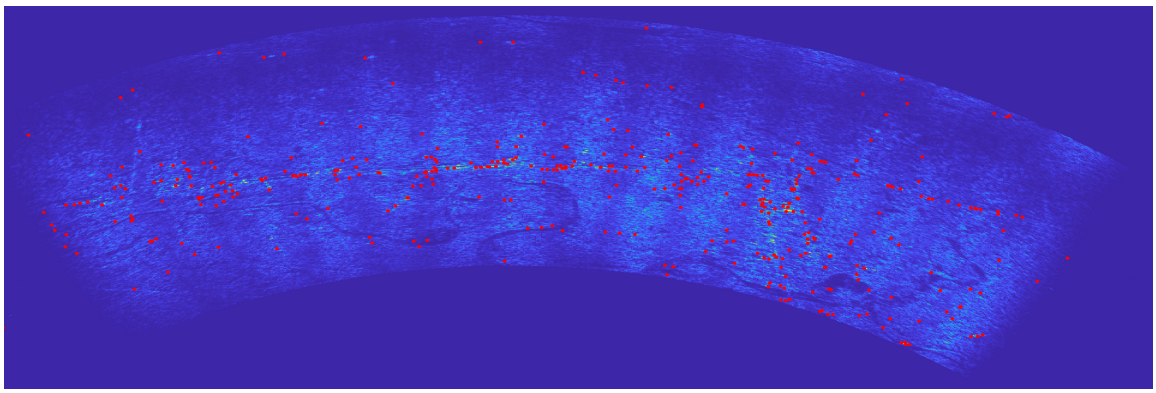
[8] W. L. Melvin, "Space-time adaptive radar performance in heterogeneous clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 621–633, Apr. 2000.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2023.3274104

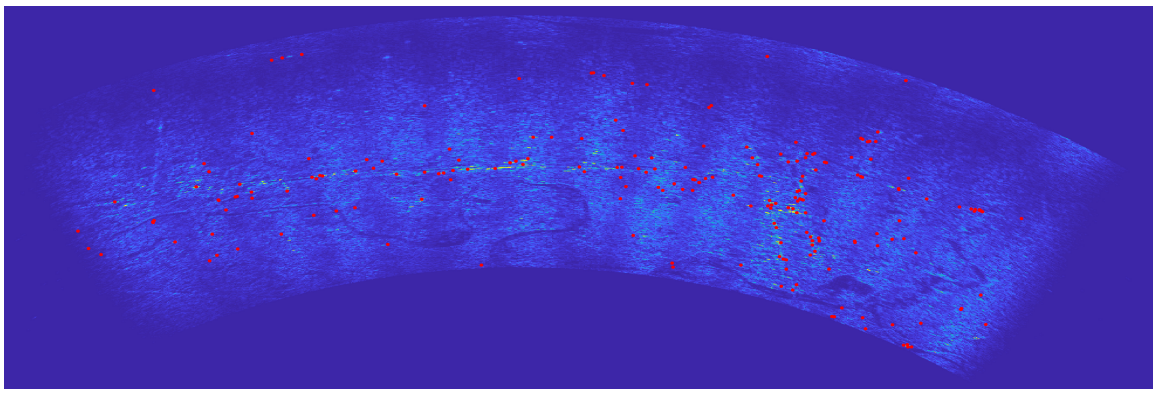
(a) Optical image



(b) DBS image

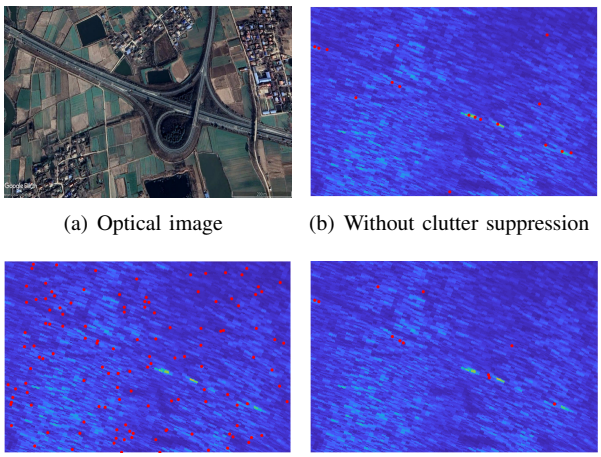


(c) WGMTI without clutter suppression



(d) WGMTI

Fig. 22: The moving target detection result of WGMTI in the 77th cycle.



(a) Optical image (b) Without clutter suppression

(c) Post-Doppler STAP (d) STSBL-STAP

Fig. 23: Local area detection results of different methods.

[9] J. Ward, Space-time adaptive processing for airborne radar. Lincoln Lab. Massachusetts Inst. Technol, Lexington, KY, USA, Tech. Rep. 1015, 1994.

[10] H. Wang and L. Cai, "On adaptive spatial-temporal processing for airborne surveillance radar systems," *IEEE Trans. Aerosp. Electron. Syst*., vol. 30, no. 3, p. 660-670, 1994.

[11] A. M. Haimovich and M. Berin, "Eigenanalysis-based space-time adaptive radar: Performance analysis," *IEEE Trans. Aerosp. Electron. Syst*., vol. 33, no. 4, pp. 1170–1179, Oct. 1997.

[12] J. S. Goldstein and I. S. Reed, "Subspace selection for partially adaptive sensor array processing," *IEEE Trans. Aerosp. Electron. Syst*., vol. 33, no. 2, pp. 539–544, Apr. 1997.

[13] T. K. Sarkar, H. Wang, S. Park, et al., "A deterministic least-squares approach to space-time adaptive processing (STAP)," *IEEE Trans. on Antennas Propag*., vol. 49, no. 1, pp. 91-103, Jan 2001.

[14] M. Rani, S. Dhok, and R. Deshmukh, "A systematic review of compressive sensing: Concepts, implementations and applications," *IEEE Access*., vol. 6, pp. 4875–4894, 2018.

[15] J. Li, X. Zhu, P. Stoica, and M. Rangaswamy, "High resolution angledoppler imaging for MTI radar," *IEEE Trans. Aerosp. Electron. Syst*., vol. 46, no. 3, pp. 1544–1556, Jul. 2010.

[16] S. Maria and J. J. Fuchs, "Application of the global matched filter to STAP data an efficient algorithmic approach," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process*., 2006, p. IV.

[17] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed,"Space-time adaptive processing and motion parameter estimation in multistatic passive radar using sparse Bayesian learning," *IEEE Trans. Geosci. Remote Sens*., vol. 54, no. 2, pp. 944–957, Feb. 2016.

[18] N. Cui, K. Xing, Z. Yu, and K. Duan, "Tensor-Based Sparse Recovery Space-Time Adaptive Processing for Large Size Data Clutter Suppression in Airborne Radar," *IEEE Trans. Aerosp. Electron. Syst*., 2022, doi:10.1109/TAES.2022.3192223.

[19] Z. Li, H. Ye, Z. Liu, et al., "Optimally Matched Space-Time Filtering Technique for BFSAR Nonstationary Clutter Suppression," *IEEE Trans. Geosci. Remote Sens*., vol. 60, pp. 1–17, 2022, doi: 10.1109/TGRS.2021.3090462.

[20] H. Li, G. Liao, J. Xu, et al., "Robust STAP for non-homogeneous clutter suppression with a single snapshot," *Dig. Signal Process*., vol. 126, 103477, Jun 2022, doi: 10.1016/j.dsp.2022.103477.

[21] K. Sun, H. Zhang, G. Li, H. Meng, and X. Wang, "A novel STAP algorithm using sparse recovery technique," in *Proc. IGARSS*, 2009, pp. 336–339.

[22] S. Sen, "OFDM radar space-time adaptive processing by exploiting spatio-temporal sparsity," *IEEE Trans. Signal Process*., vol. 61, no. 1, pp. 118–130, Jan. 2012.

[23] S. Sen, "Low-rank matrix decomposition and spatio-temporal sparse recovery for STAP radar," *IEEE J. Sel. Topics Signal Process*., vol. 9, no. 8, pp. 1510–1523, Dec. 2015.

[24] Z. Yang, X. Li, H. Wang, and W. Jiang, "On clutter sparsity analysis in space-time adaptive processing airborne radar," *IEEE Geosci. Remote Sens. Lett*., vol. 10, no. 5, pp. 1214–1218, Sep. 2013.

[25] K. Sun, H. Meng, Y. Wang, et al., "Direct data domain STAP using sparse representation of clutter spectrum," *Signal Process*., vol. 91, no. 9, pp. 2222–2236, Apr. 2011.

[26] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[27] Z. Yang, X. Li, H. Wang, and W. Jiang, "Adaptive clutter suppression based on iterative adaptive approach for airborne radar," *Signal Process*., vol. 93, no. 12, pp. 3567–3577, Dec. 2013.

[28] Z. Wang, Y. Wang, K. Duan, and W. Xie, "Subspace-augmented clutter suppression technique for STAP radar," *IEEE Geosci. Remote Sens. Lett*., vol. 13, no. 3, pp. 462–466, Mar. 2016.

[29] Z. Yang, R. C. de Lamare, and X. Li, "L_1-regularized STAP algorithms with a generalized sidelobe canceler architecture for airborne radar," *IEEE Trans. Signal Process*., vol. 60, no. 2, pp. 674–686, Feb. 2012.

[30] J. T. Parker and L. C. Potter, "A Bayesian perspective on sparse regularization for STAP post-processing," in *Proc. IEEE Radar Conf*., May 2010, pp. 1471–1475.

[31] K. Duan, Z. Wang, W. Xie, et al., "Sparsity-based STAP algorithm with multiple measurement vectors via sparse Bayesian learning strategy for airborne radar," *IET Signal Process*., vol. 11, no. 5, pp. 544–553, May 2017.

[32] D. P. Wipf, B. D. Rao, and S. Nagarajan, "Latent variable Bayesian models for promoting sparsity," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, pp. 6236–6255, Sep. 2011.

[33] S. Han, C. Fan, and X. Huang, "A novel STAP based on spectrum-aided reduced-dimension clutter sparse recovery," *IEEE Geosci. Remote Sens. Lett*., vol. 14, no. 2, pp. 213–217, Feb. 2017.

[34] N. Cui, Z. Yu and K. Xing, "A novel reduced-dimensional dictionary grid screening strategy for SR-STAP," *Social Netw. Appl. Sci*., vol. 3, no. 2, pp. 1-12, 2021.

[35] Z. Yang et al., "Knowledge-aided STAP with sparse-recovery by exploiting spatio-temporal sparsity," *IET Signal Process*., vol. 10, pp. 150–161, Mar. 2016.

[36] Y. Guo, G. Liao, and W. Feng, "Sparse representation based algorithm for airborne radar in beam-space post-Doppler reduced-dimension space-time adaptive processing," *IEEE Access*, vol. 5, pp. 5896–5903, 2017.

[37] Z. Wang, W. Xie, K. Duan, et al., "Clutter suppression algorithm based on fast converging sparse Bayesian learning for airborne radar," *Signal Process*., vol. 130, pp. 159–168, Jan. 2017.

[38] K. Duan, H. Xu, H. Yuan, H. Xie, and Y. Wang, "Reduced-DOF threedimensional STAP via subarray synthesis for non-sidelooking planar array airborne radar," *IEEE Trans. Aerosp. Elect. Syst*., vol. 56, no. 4, pp. 3311–3325, Aug. 2020.

[39] F. Jia, G. Sun, Z. He, and J. Li, "Grating-lobe clutter suppression in uniform subarray for airborne radar STAP," *IEEE Sensors J*., vol. 19, no. 16, pp. 6956–6965, Aug. 2019.

[40] N. Cui, K. Xing, K. Duan and Z. Yu, "Knowledge-aided block sparse Bayesian learning STAP for phased-array MIMO airborne radar," *IET Radar, Son. Navig*., vol. 15, no. 12, pp. 1628–1642, Dec. 2021.

[41] E. Conte, A. De Maio, and G. Galdi, "Statistical analysis of real clutter at different range resolutions," *IEEE Trans. Aerosp. Electron. Syst*., vol. 40, no. 3, pp. 903–918, 2004.

[42] N. Cui, K. Duan, K. Xing and Z. Yu, "Beam-space reduced-dimension 3D-STAP for nonside-looking airborne radar," *IEEE Geosci. Remote Sens. Lett*., vol. 19, pp. 1–5, May. 2021.

[43] L. B. "Fertig Estimation of space-time clutter rank for subarrayed data," In Proc. Asilomar Conf. Signal, Syst. Comput., Pacific Grove, CA, USA, Nov. 2004, pp. 289–292.

[44] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst*., vol. AES-24, no. 4, pp. 397–401, Jul. 1988.

[45] S. Bernhardt, R. Boyer, S. Marcos, and P. Larzabal, "Compressed sensing with basis mismatch: Performance bounds and sparse-based estimator," *IEEE Trans. Signal Process*., vol. 64, no. 13, pp. 3483–3494, Jul. 2016.

[46] K. Duan, W. Liu, G. Duan, et al., "Off-grid effects mitigation exploiting knowledge of the clutter ridge for sparse recovery STAP," *IET Radar, Son. Navig*., vol. 15, no. 5, pp. 557–564, Apr. 2018.

[47] S. N. Madsen, "Estimating the Doppler centroid of SAR data," *IEEE Trans. Aerosp. Electron. Syst*., vol. 25, no. 2, pp. 134–140, Mar. 1989.

[48] J. H. G. Ender, "Space-time processing for multichannel synthetic aperture radar," *Electron. Commun. Eng. J*., vol. 11, no. 1, pp. 29–38, Feb. 1999.

[49] X. Fang, G. Xiao, Z. Cao, R. Min, and Y. Pi, "Migration correction algorithm for coherent integration of low-observable target with uniform radial acceleration," *IEEE Trans. Instrum. Meas*., vol. 70, pp. 1–13, 2021, doi: 10.1109/TIM.2020.3023512.

[50] J. Li, Z. Wang, W. Yu, et al., "A Novel Speckle Suppression Method with Quantitative Combination of Total Variation and Anisotropic Diffusion PDE Model," *Remote Sens*., vol. 14, no. 3, 796, 2022, doi: 10.3390/rs14030796.

[51] A. B. C. da Silva, S. V. Baumgartner, and G. Krieger, "Training data selection and update strategies for airborne post-Doppler STAP," *IEEE Trans. Geosci. Remote Sens*., vol. 57, no. 8, pp. 5626–5641, Aug. 2019.

[52] P. Gandhi and S. Kassam, "Analysis of CFAR processors in nonhomogeneous background," *IEEE Trans. Aerosp. Electron. Syst*., vol. AES-24, no. 4, pp. 427–445, Jul. 1988.

[53] Q. Zheng et al., "A target detection scheme with decreased complexity and enhanced performance for range-Doppler FMCW radar," *IEEE Trans. Instrum. Meas*., vol. 70, 2021, doi: 10.1109/TIM.2020.3027407.

[54] U. Nickel, E. Chaumette, and P. Larzabal, "Estimation of extended targets using the generalized monopulse estimator: Extension to a mixed target model," *IEEE Trans. Aerosp. Electron. Syst*., vol. 49, no. 3, pp. 2084–2096, Jul. 2013.

Ning Cui (S'19) received the B.S. degree in electronic information science and technology from the Wuhan University of Technology, Wuhan, China. He is currently pursuing the Ph.D. degree in the school of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences at Beijing, China.

His research interests include array signal processing, space-time adaptive processing, compressive sensing, GMTI/MMTI and their application in real radar systems.

Kun Xing was born in Xi'an, China, in 1980. He received the B.S. degrees in computer science and technology from Xi'an Jiaotong University, Xi'an, China, in 2003, and M.S. degrees in electronic engineering from Xidian University, Xi'an, China, in 2006, respectively. In 2006, he joined the Institute of Electrics, Chinese Academy of Sciences (IECAS).

Now he is the Associate researcher and Master's tutor in the Aerospace Information Research Institute, Chinese Academy of Sciences. His current research interests include design and performance simulation of airborne multifunction radar, ground and marine moving target detection, tracking, classification and recognition.

Zhongjun Yu was born in Chengdu, China, in 1980. He received the B.S. and M.S. degrees in electronic engineering from the University Of Electronic Science And Technology Of China, Chengdu, China, in 2003, 2005, and Ph.D. degree in School of Electronic, Electrical and Communication Engineering from the University of Chinese Academy of Sciences, Beijing, China, in 2011, respectively. From 2012 to 2013, he was a Visiting Scholar with City University of Hong Kong, Hong Kong, China. In 2005, he joined the Institute of Electrics, Chinese Academy of Sciences (IECAS).

Now he is the Researcher and doctorial tutor in the Aerospace Information Research Institute, Chinese Academy of Science. His current research interests include design and evaluation of microwave integrated systems and component devices.

Keqing Duan (M'15) received the B.S. and M.S. degrees from Wuhan Early Warning Academy, Wuhan, China, and the Ph.D. degree from National University of Defense Technology (NUDT), Changsha, China, in 2003, 2006, and 2010, respectively, all in electrical engineering.

Now he is the Associate professor and Master's tutor in the Sun Yat-sen University (SYSU), Guangzhou, China, and the judge expert of the National natural science foundation of China. His research interests include the area of signal processing, including array signal processing, space-time adaptive processing (STAP), compressive sensing, and its application to radar systems.