A Robust Space-time Adaptive Processing Algorithm based on Particle Swarm Optimization for Non-stationary Clutter Suppression

Hanchao Wang, Lili Fang*, Chuanfang Zhang
School of Information and Electronics, Beijing Institute of Technology, Beijing, China

*Corresponding author: fanglili@bit.edu.cn

Abstract. A novel robust sparse recovery (SR) space-time adaptive processing (STAP) algorithm based on particle swarm optimization (PSO) for non-stationary clutter suppression is presented in this paper. A cost function for PSO in the presence of parameter errors is theoretically derived. An improved estimation process of clutter spectrum based on this cost function which is called PSO-SR is proposed and analyzed. A more accurate estimation result of clutter spectrum could be provided by this algorithm than the previous proposed algorithms in the presence of considerable parameter errors. Simulation results demonstrate the robust performance of this algorithm.

Keywords: Space-time adaptive processing (STAP), Conformal array, Sparse recovery (SR), Clutter compensation

1. Introduction
Space-time adaptive processing (STAP) technology is a successful method for stationary clutter suppression [1] [2]. However, the performance of conventional STAP will suffer a serious degradation when the characteristics of clutter of the conformal array are non-stationary, since the independent and identically distributed (IID) condition could not be met.

To mitigate the range dependence of the non-stationary clutter, a series of compensation methods have been proposed. Parametric methods, such as Doppler Warping [3], angle Doppler compensation [4] and registration-based compensation (RBC) [5], try to compensate the angle-Doppler response of clutter at adjacent range bins to the range bin under test. However, the estimation of non-stationary clutter spectra of the near range is not convenient, since the account of IID training samples cannot meet the Reed-Mallett-Brennan (RMB) rules [2] because of the severe range dependence in the near range. Non-parametric methods include: the derivative-based updating [6], prediction of the inverse clutter covariance matrix (CCM) [7] and Taylor series expansion-based eigen-canceller [8] etc., which do not require the knowledge of the system parameters. The performance of these methods depends on the accurate non-parametric models, which are not available in practice.

Sparse recovery (SR) theory could recover the signal with the minimum number of vectors from an over complete dictionary [9]. Since the non-stationary clutter has the characteristics of sparse distribution [10], SR theory could be used to estimate the non-stationary clutter spectra with the support of low IID samples. Recently several STAP algorithms based on sparsity have been proposed,
such as a fast converging sparse Bayesian learning (FCSBL) [11] and subspace-augmented multiple signal classification (SA-MUSIC) [12]. Since the over-complete dictionary is configured with ideal space-time steering vectors, the performance of the algorithms will degrade seriously in non-ideal condition where the unknown system errors are existed. In essential, the configuration of over-complete dictionary under blind parameters is a problem of multi-parameters optimization, which could be solved by the intelligent algorithms [13]. Recently, the intelligent algorithms such as Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) [14].

A novel sparse recovery STAP algorithm based on PSO for non-stationary clutter suppression is presented in this paper. A new cost function based on the \(l_2\)-norm of difference between the estimated clutter which is obtained by multiplying the significant vectors of spectra after sparsity recovery by a constantly revised over-complete dictionary and the echoed clutter, has been proposed for optimization. Simulation results show that the presented algorithm is robust when the system parameters are inaccurate.

2. Signal Model

2.1. General Clutter Model

Considering an array with \(N\) elements, \(M\) identical pulses are transmitted during each coherent processing interval (CPI) at a constant pulse repetition frequency (PRF) \(f_r\). The platform altitude is \(H\), the velocity is \(v\), and the radar wavelength is \(\lambda\). For each PRF, \(L\) fast time samples are collected to cover the detection region. The echoed data \(x\) could be expressed as a \(NM\)-dimensional vector:

\[
x = \alpha_0 v_0 + x_i
\] (1)

Where \(\alpha_0\) and \(v_0\) are the complex amplitude and space-time steering vector of the target, respectively. \(x_i = e + n\), where \(e\) is clutter matrix and \(n\) is noise matrix.

The space-time clutter snapshot could be expressed as the superposition of \(N_c\) independent clutter patches as follows:

\[
c = \sum_{i=1}^{N_c} \xi_i v_i = \sum_{i=1}^{N_c} \xi_i (v_{i,i} \otimes v_{s,i})
\] (2)

Where \(\xi_i\), \(v_i\), \(v_{i,i}\) and \(v_{s,i}\) denote the complex amplitude, the space-time steering vector, the doppler steering vector and the spatial steering vector of the \(i\)th patch, respectively.

\[
v_{i,i} = [1 \exp(j\pi f_{di,i}), \cdots, \exp(j\pi(K-1)f_{di,i})]^T
\] (3)

\[
v_{s,i} = g(\theta_i, \phi_i) \odot [e^{j2\pi f_{di,i}/\lambda}, \cdots, e^{j2\pi f_{di,i}/\lambda}]^T
\] (4)

Where \(f_{di,i}\) and \(f_{si,i}\) are the normalized doppler frequency and the spatial frequency of the \(i\)th clutter patch respectively. \(d_n\) is the position of the \(n\)th element, \(g\) is the gain of the array.

Taking the non-ideal factors such as amplitude and phase inaccuracy into account, the spatial steering vector \(v_{s,i}\) in (2) should be taken over by the real spatial steering vector \(\hat{v}_{s,i}\):

\[
\hat{v}_{s,i} = v_{s,i} \odot t
\] (5)

Where \(t = [t_i, t_i, \cdots, t_N]^T\) is the error vector,
\[ t_n = (1 + \gamma_n) e^{i \beta_n} \quad n = 1, \ldots, N \]  

Where \( \gamma_n \) is the normalized amplitude error for the \( n \)th element antenna, \( \beta_n \) is the phase error. \( \gamma_n \) follows the Gaussian model with mean zero and the standard deviation \( \gamma_{\text{max}} \), \( \beta_n \) also follows the Gaussian model with mean zero and the standard deviation \( \beta_{\text{max}} \). Therefore,

\[ \hat{v}_i = v_{i,i} \odot \hat{v}_{i,i} \]  

Then, the real clutter snapshot \( \hat{c} \) is:

\[ \hat{c} = \sum_{i=1}^{N_t} \xi_i \hat{v}_i = \sum_{i=1}^{N_t} \xi_i (v_{i,i} \odot v_{i,i}) \odot (I_M \odot t) \]

\[ = \sum_{i=1}^{N_t} \xi_i v_i \odot (I_M \odot t) \]  

Where \( 1_M = [1,1,\ldots,1]^T \in \mathbb{R}^{M \times 1} \).

2.2. Sparse Recovery

To construct over-complete dictionary for SR theory, the spatial-doppler plane should be meshed into \( N_s \times N_d \) grid nodes, where \( N_s = \rho_s N \) and \( N_d = \rho_d M \) (\( \rho_s > 1, \rho_d > 1 \)). \( \rho_s \) and \( \rho_d \) are the grid interval along the spatial and the doppler axis, respectively. The echoed signal can be expressed as:

\[ x = \sum_{m=1}^{N_s} \sum_{n=1}^{N_d} \alpha(m,n) \hat{w}_{m,n} + n \]  

Where \( \alpha(m,n) \) is the complex amplitude for each grid node. Equation (10) can be rewritten as:

\[ x = \hat{D} a + n \]  

Where \( \hat{D} = TD \), \( T = I_M \odot \text{diag}(t) \), \( I_M \) is an identity matrix of size \( M \), \( D \) is the constructed space-time dictionary, \( a \) is the reflection coefficient vector of the clutter. Traditionally, the SR technique could represent \( x \) as a linear combination of the atoms of \( D \) as few atoms as possible. The procedure could be written as:

\[ \min \| a \|_0 \quad \text{s.t.} \quad x - Da \leq \varepsilon \]  

Where \( \| \cdot \|_0 \) denotes \( l_0 \) norm and \( \| \cdot \|_1 \) denotes \( l_1 \) norm. \( \varepsilon \) is the allowance. The resulting \( a \) is used for the estimation of \( x \).

3. Principle of PSO-SR STAP Algorithm

In order to suppress the non-stationary clutter in the presence of inaccuracy of system parameters, a novel cost function of particle swarm optimization in the non-ideal case is theoretically derived, which is used to improve the position of significant component in the estimated clutter spectrum.

3.1. Construction of the position of PSO

The position of PSO is constructed as the combination of two parts: significant component matrix of clutter spectra and error matrix of the over-complete dictionary. The initial value of the swarm position \( S_0 \) could be written as:

\[ S_0 = [O_t, T_e] \in \mathbb{R}^{2(N_s + N_d) \times G} \]
Where $O_0 \in \mathbb{R}^{2N_x \times N_r}$ is the initial value of significant component matrix, $T_0 \in \mathbb{R}^{2N_x \times N_r}$ is the initial value of error matrix of dictionary. $O_0$ could be constructed as follows:

$$O_0 = \text{repmat}([x^T; y^T], N_r) + c_0 r_3$$

(13)

Where $[x^T; y^T] \in \mathbb{R}^{2N_x \times 1}$ is the coordinate of the significant component for the $l$th range cell. \text{repmat}$(\cdot)$ represents the operation of duplication along the column. Item $c_0 r_3$ is introduced for searching and updating. $T_0$ could be built up as follows:

$$T_0(1: N,:,:) = c_1 r_1$$
$$T_0(N+1: 2N,:,:) = c_2 r_2$$

(14)

Where $T_0(1:N,:,:)$ and $T_0(N+1:2N,:,:)$ represent the initial values of the amplitude and the phase errors, $r_1$ is a matrix with standard normal distribution along the column, $c_1$ and $c_2$ are random constants.

After the iteration, the swarm position could be obtained

$$S_t = [O_t T_t]$$

(15)

Where subscript $t$ represents the number of iteration and the element of $S_t$ correspond to the position in PSO.

### 3.2. Derivation of Cost Function

The range-dependent characteristics of the clutter data in adjacent range cells will decrease with the increasement of the slant range [15]. After SR, the resulting $a$ is used as the column to construct a joint coefficient matrix, whose row vectors are superimposed by $l_2$ norm to enhance the significant component of the estimated spectrum.

To make the constructed $D$ approximate to the real dictionary, an optimization problem which include error matrix and reflection coefficient vector is constructed as follows:

$$\min_{\hat{T}, \hat{a}} \|\hat{T} \hat{a} - x\|_2$$

(16)

Where $\hat{T}$ is the error matrix to be estimated, $\hat{a}$ is the reflection coefficient vector of clutter.

After the $t$th iteration, an error matrix $T_{t-1}$ and a significant component matrix $O_{t-1}$ for the $p$th particle could be acquired as follows:

$$T_{t-1} = I_{N_r} \otimes \text{diag}(1 + T_0(1:N,:)) \otimes \exp(j T_0(N+1:2N,:))$$

(17)

$$O_{t-1} = [O_t(1:N_r,:), O_t(N_r+1:2N_r,:)]$$

(18)

The constructed dictionary $D_{t-1}$ after the $t$th iteration is:

$$D_{t-1} = T_{t-1} D$$

(19)

A matrix $D_{t-1}$ is constructed by decimating $D_{t-1}$ along the columns using the positions of significant components in $O_{t-1}$. $a_{t-1}$ could be estimated by the least squares criterion:
To make the estimated clutter data approximate to the echoed data, the cost function is set to the $l_2$-norm as follows:

$$F(S_i(:, p)) = \left\| \tilde{c}_{p, t} - x \right\|_2 < \varepsilon$$

(21)

Where $\tilde{c}_{p, t} = T_{p, t}D\tilde{a}_{p, t}$, $\varepsilon$ is the iterative error threshold. The searching direction of PSO is to decrease $F(\bullet)$ to meet the threshold. After the iteration, $T_{p, t}$ and $\tilde{a}_{p, t}$ in the global optimum position of the particle swarm, are chose as the value of $T$ and $\tilde{a}$, and then the over-complete dictionary of SR could be expressed as:

$$D_r = TD$$

(22)

3.3. Estimation of Clutter Ridge

The results of PSO are used for clutter fitting and further clutter suppression. Define velocity vector of radar $v(\varphi, \theta_r)$. $\varphi$ and $\theta_r$ denote the azimuth and elevation angle of flight direction respectively, and generally assume $\theta_r = 0$. The normalized doppler frequency is given by

$$f_d = 4(\cos \theta \cos \varphi \cos \sin \varphi \sin \theta / (\lambda f_c))$$

$$= 4(\cos \varphi \cos \theta \sin \varphi \sin \theta / (\lambda f_c))$$

$$= (4v \cos \theta \cos \varphi / (\lambda f_c)) \cos \varphi + (4v \cos \theta \sin \varphi / (\lambda f_c)) \sin \varphi$$

(23)

Equation (31) could be written as:

$$f_d^2 + k_1 \cos^2 \varphi + k_2 f_d \cos \varphi = k_3$$

(24)

Where, $k_1 = (4v \cos \theta / (\lambda f_c))^2$, $k_2 = -8v \cos \varphi \cos \varphi / (\lambda f_c)$, $k_3 = (4v \cos \theta \sin \varphi / (\lambda f_c))^2$. Equation (32) could be rewritten in matrix form:

$$\eta^T p + u = 0$$

(25)

Where $\eta = [k_1 k_2 k_3]^T$, $p = [\cos^2 \theta \ f_d \cos \theta - 1]^T$, $u = f_d^2$. $p$ and $u$ could be determined when $\tilde{a}$ is given. Then $\eta$ could be estimated by the least squares criterion:

$$\eta = R_p^{-1}R_{ps}$$

(26)

Where $R_p = \sum_{n=1}^{N_p} p_n p_n^T$, $R_{ps} = -\sum_{n=1}^{N_p} p_n u_n$. A diagonal matrix $B$ used for CCM estimation could be constructed as follows:

$$B = \text{diag}(b)$$

(27)

$$b(o) = \begin{cases} \left\| (D'_o)^H D'_o \right\|^{-1} (D'_o)^H x, & o \in P \\ 0, & o \notin P \end{cases}$$

(28)

Where $P$ is the position set of clutter ridge, $D'$ is constructed by decimating $D$ along the columns using $P$. The estimation of CCM $\tilde{R}$ could be obtained as follows:
\[ \hat{R} = DBD^H \]  

Since the sparse structures of the echoed data in the several range cells where the slant ranges are 
far enough are consistent, a joint estimation of CCM \( \hat{R}_m \) could be achieved, 

\[ \hat{R}_m = DBD^H \]  

Where \( B_m = diag(b_m) \) is a diagonal matrix. 

\[ b_m(o) = \begin{cases} \frac{1}{L_m} \sum \left| (D')^H D' \right|^{-1} (D')^H X \right|_2^2, & o \in P \\ 0, & o \not\in P \end{cases} \]  

Where \( X = [x_1, x_2, \ldots, x_{L_m}] \) is the echoed data in the reference range cells. \( \sum(\bullet, 2) \) represents the 
sum of the matrix by column. 

3.4. Clutter Suppression Using RBC 
When the CCM is obtained, a transform matrix could be constructed to make the statistical character of 
clutter between the \( l \)th range cell and the cell under test (CUT) to be consist. \( \hat{R}_m \) is chosen as the 
reference cell to compensate the CCM estimation \( \hat{R} \) corresponding to the \( l \)th range cell. Using the 
eigenvalues decomposition, \( \hat{R}_m \) and \( \hat{R}_l \) can be expressed as: 

\[ \hat{R}_m = V_m \Lambda_m V_m^H = (V_m \Lambda_m^{1/2}) (V_m \Lambda_m^{1/2})^H \]  

\[ \hat{R}_l = (V_l \Lambda_l^{1/2}) (V_l \Lambda_l^{1/2})^H \]  

Where \( \Lambda_l \) and \( \Lambda_m \) are the diagonal matrixes containing the eigenvalues for the \( l \)th range cell and 
the reference cell,

![Figure 1. Lement arrangement of the Cylindrical array antennas.](image-url)
Respectively. $V_l$ and $V_m$ are the relative matrixes. The transform matrix $M_i$ for $\hat{R}_i$ in RBC could be constructed as:

$$M_i = V_m^{1/2} \Lambda M_i^{-1/2} V_l^{R}$$  \hspace{1cm} (34)

After the transformation, the compensated CCM of $\hat{R}_i$ is:

$$\hat{R}_{GBR.i} = M_i \hat{R}_i M_i^R$$  \hspace{1cm} (35)

The estimated CCM which is used for the calculation of weight vector in STAP is

$$\hat{R}_L = \frac{1}{L} \sum_{i=1}^{L} \hat{R}_{GBR.i} + \sigma_L^2 I$$  \hspace{1cm} (36)

Where $\sigma_L^2$ is a fixed diagonal loading level, $L$ is the number of training data. As is well known, the weight vector $w$ for STAP is given by [16]:

$$w = \mu \hat{R}_L^{-1} v_0$$  \hspace{1cm} (37)

Where $\mu$ is a normalization constant.

4. Simulation Results and Discussion

A conformal array radar with yaw angle $\varphi_p = 30^\circ$ is considered, and the parameters of the system is given in Table I. The range cell located at $R = 9.2$km is selected as the CUT, and the number of training samples is 10, whose range cells are around $R = 12$km. All the simulation results are the average results of 100 independent Monte Carlo runs.

| Symbol | Parameter | value |
|------|-----------|-------|
| $M$  | Number of pulses per CPI | 30 |
| $f_r$ | Pulse Repetition Frequency | 2000Hz |
| $\lambda$ | Radar operating wavelength | 0.23m |
| $v$  | Platform velocity | 150m/s |
| $H$  | Platform height | 9000m |
| $\varphi_p$ | Yaw angle | $30^\circ$ |
| CNR  | Clutter-to-noise ratio | 60dB |

**Figure 2.** The actual clutter spectrum corresponding to R=12km.
Figure 3. The estimated clutter spectrum corresponding to R=12km for when $\gamma_{\infty} = 0$ and $\beta_{\infty} = 0$.
(a) the traditional SR; (b) FCSBL; (c) PSO-SR.

Figure 4. The estimated clutter spectrum corresponding to R=12km for when $\gamma_{\infty} = 0.1$ and $\beta_{\infty} = 0.1$.
(a) the traditional SR; (b) FCSBL; (c) PSO-SR.

In Fig.1, a global coordinate system is built as follows: (1) the projection point of the geometric centre of the cylinder on the ground is set to be the origin; (2) the x-axis is aligned with the normal of the array; (3) the z-axis points vertically up. The radius of the cylinder is $R_y$, $N_x \times M_y$ elements with cosine-squared patterns are evenly arranged in $(Z_0, Z_0) \times (-\varphi_0, \varphi_0)$ area on the surface of the cylinder with the distance $d_0$ along the generatrix between adjacent elements. The parameters of the cylindric array are set as: $N_y=10$, $M_y=10$, $R_y=4\lambda$, $\varphi_0 = \pi / 3$, $d_0 = 0.5\lambda$.

4.1. Performance Assessment

In this section, the performance of the proposed PSO-SR-STAP algorithm is evaluated and compared with the existed SMV SR STAP algorithms, such as traditional SR STAP, FCSBL STAP algorithms. The improvement factor (IF) is defined as follows:

$$\text{IF} = \left| \frac{\text{tr}(R)}{\text{tr}(\tilde{R})} \right|$$

(38)
4.1.1. Clutter Spectrum Estimation. In order to compare the performance of different algorithms in the presence of parameter errors, clutter spectra estimation for the cylindric conformal array and the conical conformal array using traditional SR, FCSBL and PSO-SR are accomplished. The actual clutter spectra for the cylindric array corresponding to $R=12\text{km}$ are shown in Fig. 2. The estimations of CCM corresponding to $R=12\text{km}$ using different algorithms are shown in Fig. 3~Fig. 5.

In Fig. 3, there is no existence of parameter error, the estimated clutter ridges using three algorithms are all consistent with the actual one. In Fig. 4, the deviation between the estimated clutter ridges obtained by the traditional SR and FCSBL and the actual one gradually increases with the increasement of the errors, while the estimated results obtained by the PSO-SR are still relatively close to the real one.

4.1.2. IF curves. The IF curves in the ideal case for cylindric array is shown in Fig. 5(a). It is shown that all three algorithms could effectively form notches to suppress the clutter in the ideal condition. The IF curve of PSO-SR STAP is about $9\text{ dB}$ higher than that of traditional SR STAP and about $2\text{dB}$ higher than FCSBL STAP on average. Fig. 5(b) compares the IF curves of above mentioned algorithms in the case of non-ideal condition. It is shown that, the notches obtained by the traditional SR STAP and FCSBL STAP are widened in some degree, while that obtained by PSO-SR STAP is barely widened because of the negligible difference between the estimated clutter spectrum by PSO-SR and the actual one. The IF curve of PSO-SR STAP is about $20\text{ dB}$ higher than that of traditional SR STAP and about $7\text{dB}$ higher than FCSBL STAP on average.

5. Conclusion

A robust sparse recovery space-time adaptive algorithm based on particle swarm optimization for non-stationary clutter suppression is presented in this paper. The general clutter models for two conformal arrays with different configuration are modelled, and a cost function of PSO in the presence of parameter errors is theoretically derived. The clutter spectrum estimated by this algorithm is analyzed and simulated, and the IF curve compared with other algorithms is measured. Simulation results demonstrate that this novel PSO-SR STAP is more robust than other SMV STAP algorithms for non-
stationary clutter suppression in the presence of parameter errors, which has broad application prospects in scenarios such as conformal array radar, non-SLAR etc.

References
[1] Tao Fuyu, Wang Tong, Wu Jianxin, et al. A knowledge aided SPICE space time adaptive processing method for airborne radar with conformal array[J]. Signal Processing, 2018, 152: 54-62. doi: 10.1016/j.sigpro.2018.05.015.
[2] Brennan L E and Reed L S. Theory of adaptive radar [J]. IEEE Transactions on Aerospace and Electronic Systems, 1973, 9 (2): 237-252. doi: 10.1109/TAES.1973.309792.
[3] Kreyenkamp O, Klemm R. Doppler compensation in forward-looking STAP radar [J]. IEE Proceedings - Radar, Sonar and Navigation, 2001, 148 (5): 253-258.
[4] B. Himed and A. Hajjari, “STAP with angle-doppler compensation for bistatic airborne radars,” in Proceedings of the 2002 IEEE Radar Conference, pp. 22–25, Long Beach, Calif, USA, April 2002.
[5] Borsari G K. Mitigating effects on STAP processing caused by an inclined array [C]. IEEE National Radar Conference, Dallas, USA, 1998: 135-140. doi: 10.1109/NRC.1998.677990.
[6] Zatman, M. Circular array STAP [J]. IEEE Transactions on Aerospace & Electronic Systems, 2000, 36 (2): 510-517.
[7] C.-H. Lim and B. Mulgrew, “Linear prediction of range-dependent inverse covariance matrix (PICM) sequences,” Signal Process., vol. 87, no. 6, pp. 1412_1420, Dec. 2007.
[8] Beau S M S. Taylor series expansions for airborne radar space-time adaptive processing [J]. Radar, Sonar & Navigation, IET, 2011, 5(3):p.266-278.
[9] Donoho D L, Elad M, and Temlyakov V N. Stable recovery of sparse overcomplete representations in the presence of noise [J]. IEEE Transactions on Information Theory, 2006, 52: 6-18. doi: 10.1109/TIT.2005.860439.
[10] Yang Z, Li X, Wang H, et al. On clutter sparsity analysis in space-time adaptive processing airborne radar [J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10 (5): 1214-1218.
[11] Wang Z, Xie W, Duan K, et al. Clutter suppression algorithm based on fast converging sparse Bayesian learning for airborne radar [J]. Signal Processing, 2017, 130 (jan.): 159-168.
[12] Davies M E, Eldar Y C. Rank Awareness in Joint Sparse Recovery [J]. IEEE Transactions on Information Theory, 2012, 58 (2): 1135-1146.
[13] Chen K S, Zhu Y Y, Ni X L, et al. Low Sidelobe Sparse Concentric Ring Arrays Optimization Using Modified GA [J]. International Journal of Antennas & Propagation, 2016, 2015 (PT.2): 147247.1-147247.5.
[14] Shangce G, Mengchu Z, Yirui W, et al. Dendritic Neuron Model With Effective Learning Algorithms for Classification, Approximation, and Prediction [J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, PP: 1-14.
[15] J. Ward, "Space-time adaptive processing for airborne radar," IEE Colloquium on Space-Time Adaptive Processing (Ref. No. 1998/241), London, UK, 1998, pp. 2/1-2/6, doi: 10.1049/ic:19980240.