

A fast inverse synthetic aperture radar translational motion compensation method for low SNR scenes

Ping Cheng^{1*}, Xinkai Yang¹, Weiyang Chen¹, Jiaqun Zhao^{2*}, Zhiyu Qu¹

Accurate translational motion compensation is crucial for inverse synthetic aperture radar (ISAR) imaging. At low signal-to-noise ratio (SNR), the performance of nonparametric methods is declining because the correlation of adjacent range profiles is severely destroyed by the noise. Fortunately, parametric methods can still work well at low SNR. However, existing parametric methods are time-consuming, which decreases ISAR imaging efficiency significantly. So, a fast parametric translational motion compensation method is studied for low SNR scenes in the paper, which is based on modified Adam (Madam) algorithm. By changing the updating strategy for weights and learning rate in Adam, Madam algorithm can significantly accelerate the convergence when estimating the motion parameters. To further improve the efficiency, a new objective function by combining image contrast and image entropy is proposed. When applied into measured Yak-42 data at low SNR, the proposed method can image about 10 times faster than particle swarm optimization (PSO) method and 2 times faster than Adam method. Therefore, the proposed method is an efficient translational motion compensation method for low SNR ISAR imaging.

Keywords: inverse synthetic aperture radar (ISAR), translational motion compensation, Adam, image contrast, low SNR

1 Introduction

Inverse synthetic aperture radar (ISAR) can obtain high-resolution images of targets all day and all weather by using the relative motion of targets to radar, so it has found wide application in civil and military fields [1, 2]. The high range resolution is obtained by transmitting wide band signal such as linear frequency modulated (LFM) signal, stepped frequency signal, etc. And the high cross-range resolution is achieved by the relative rotation between target and radar. But the target motion consists of translational motion and rotational motion. And the translational motion leads to range misalignment and phase errors. Hence, a crucial step in ISAR imaging is to compensate translational motion well.

By now, many translational motion compensation methods have been proposed, which include nonparametric methods and parametric methods. Nonparametric methods usually consist of range alignment and phase autofocus. Conventional range alignment methods include maximum correlation-based range alignment (MCRA) method and minimum average range profile entropy method [3], etc. And representative phase autofocus methods include Doppler centroid tracking method, phase gradient autofocus (PGA) method [4], etc. These nonparametric methods work well under high signal-to-noise ratio (SNR) conditions. Nevertheless, they cannot get satisfied results at low SNR as they are based on the correlation of adjacent range profiles which is severely destroyed by the strong noise.

Parametric methods model the translational motion as a polynomial with unknown coefficients and estimate the coefficients by using optimization methods. Based on the estimated coefficients, compensation function is constructed to compensate for the translational motion. This kind of methods can achieve good results under low SNR conditions as they have low requirement for the correlation of profiles. Intelligent optimization algorithms such as particle swarm optimization (PSO) and genetic algorithm (GA) are often utilized to estimate translational motion parameters [5,6]. Although they achieve good ISAR images, the computational complexity is very high.

Compared with intelligent optimization algorithms such as PSO and GA, gradient-based optimization algorithms can get good results with much less time. Adam algorithm is an excellent gradient-based optimization algorithm which is widely used in deep learning now [7]. Although Adam has made many improvements based on gradient descent (GD) algorithm, it still has some disadvantages. First, due to the use of exponential moving average (EMA), its update is affected by past gradients. Hence, when approaching global optimal point, the convergence speed of Adam decreases. Second, Adam's learning rate increases when approaching global optimal point, which causes the decrease in convergence speed or even not converge.

Hence, a novel modified Adam (Madam) method is proposed for translational motion parameters estimation

¹ College of Information Science and Engineering, Hohai University, Nanjing 211100, China

² College of Mechanics and Engineering Science, Hohai University, Nanjing 211100, China

*Corresponding authors: chengping1219@163.com, zhaojq@hhu.edu.cn

in the paper. The contributions of this paper are as follows.

- (1) Adam algorithm is innovatively used to estimate translational motion parameters. Compared with intelligent optimization methods, Adam is much more efficient.
- (2) Madam method is proposed to significantly accelerate the convergence of Adam, which modifies the updating strategy of weights and learning rate reasonably.
- (3) A new objective function by combining image contrast and image entropy is proposed to further improve efficiency and performance.

2 Signal model

Suppose linear frequency modulated signal is used as radar transmitted signal, which is

$$S_T(\hat{t}, t_m) = \text{rect}\left(\frac{\hat{t}}{T_p}\right) \exp\left\{j2\pi\left(f_c \hat{t} + \frac{1}{2}\gamma \hat{t}^2\right)\right\} \quad (1)$$

where $\hat{t}$, t_m , T_p , f_c and γ are fast time, slow time, pulse width of transmitted signal, carrier frequency and chirp rate respectively, and

$$\text{rect}\left(\frac{\hat{t}}{T_p}\right) = \begin{cases} 1, & |\hat{t}| \leq T_p/2 \\ 0, & |\hat{t}| > T_p/2 \end{cases} \quad (2)$$

The echo reflected by a strong scattering point of target can be written as

$$S_R(\hat{t}, t_m) = \sigma \cdot \text{rect}\left(\frac{\hat{t} - \frac{2R_p(x, y, t_m)}{c}}{T_p}\right) \cdot \exp\left\{j2\pi f_c \left[\hat{t} - \frac{2R_p(x, y, t_m)}{c}\right]\right\} \cdot \exp\left\{j\pi\gamma \left[\hat{t} - \frac{2R_p(x, y, t_m)}{c}\right]^2\right\} \quad (3)$$

where σ is scattering coefficient, x and y are cross-range coordinate and range coordinate respectively, $R_p(x, y, t_m)$ is distance between the strong scattering point and radar, and c is speed of light.

After stretch processing and remaining video phase compensation, the received signal can be written as [6]

$$S_R(f_r, t_m) \approx \sum_{k=1}^K \sigma_k \exp\left\{-j \frac{4\pi(f_c + f_r) \cdot \Delta R_k(x_k, y_k, t_m)}{c}\right\} \quad (4)$$

where K is the number of strong scattering points, σ_k , x_k and y_k are scattering coefficient, cross-range

coordinate and range coordinate of the k -th strong scattering point respectively, and f_r is range frequency which can be expressed as [6]

$$f_r = \gamma\left(\hat{t} - \frac{2R_{ref}}{c}\right) \quad (5)$$

where R_{ref} is reference distance of stretch processing. As R_{ref} may change with t_m and bring additional phase error, its influence needs to be compensated. $\Delta R_k(x_k, y_k, t_m)$ is the difference between $R_p(x_k, y_k, t_m)$ and R_{ref} , which can be expressed as

$$\Delta R_k(x_k, y_k, t_m) = R_t(t_m) + y_k \cos \theta(t_m) + x_k \sin \theta(t_m) \quad (6)$$

where $R_t(t_m)$ is translational distance and $\theta(t_m)$ is rotational angle. Substituting (6) into (4), signal (4) can be rewritten as

$$S_R(f_r, t_m) \approx \exp\left\{-j \frac{4\pi(f_c + f_r) \cdot R_t(t_m)}{c}\right\} \times \sum_{k=1}^K \sigma_k \exp\left\{-j \frac{4\pi(f_c + f_r)}{c} \cdot \begin{bmatrix} y_k \cos \theta(t_m) \\ + x_k \sin \theta(t_m) \end{bmatrix}\right\} \quad (7)$$

Usually, $R_t(t_m)$ can be approximated as [8]

$$R_t(t_m) \approx R_0 + \sum_{l=1}^L a_l t_m^l \quad (8)$$

where R_0 is the initial range which can be omitted as it has no impact on imaging [8].

In radar imaging, parameters are discretized by sampling. Replacing f_r and t_m in (7) with their discrete form $m \cdot \Delta f_r$ and $n \cdot \Delta t_a$ respectively, the discrete form of (7) is obtained

$$S_R(m, n) \approx \exp\left\{-j \frac{4\pi(f_c + m\Delta f_r) \cdot R_t(n\Delta t_a)}{c}\right\} \times \sum_{k=1}^K \sigma_k \exp\left\{-j \frac{4\pi(f_c + m\Delta f_r)}{c} \cdot \begin{bmatrix} y_k \cos \theta(n\Delta t_a) \\ + x_k \sin \theta(n\Delta t_a) \end{bmatrix}\right\} \quad (9)$$

where m and n are range frequency index and azimuth time index respectively, Δf_r and Δt_a are sampling interval in range frequency and azimuth time respectively.

The summation part in (9) supplies high resolution information for ISAR imaging. However, the first part in (9) caused by the translational motion is interference for imaging, which should be removed because it destroys ISAR imaging.

According to (8), the translational motion is determined by the polynomial coefficients. So, if the coefficients are obtained, the first part in (9), i.e., the impact of translational motion can be compensated.

The signal to compensate the translational motion can be constructed as

$$C(m, n) = \exp\left(j \frac{4\pi(f_c + m\Delta f_r)}{c} \times \sum_{l=1}^L a_l (n\Delta t_a)^l\right) \quad (10)$$

And the signal (9) is compensated as

$$S_C(m, n) = S_R(m, n) \cdot C(m, n) \quad (11)$$

Then by using inverse Fourier transform along the range direction and Fourier transform along azimuth direction, a focus ISAR image is obtained.

3 Fast translational motion parameters estimation method

3.1 Objective function

To get high-quality ISAR image, translational motion should be compensated. In parametric methods, translational motion is modelled as a polynomial and the polynomial coefficients is estimated by optimizing ISAR image quality. Usually, two image quality evaluation indicators image entropy (IE) and image contrast (IC) are used as objective function [2,9].

Supposing ISAR image has P rows and Q columns, and $I(p, q, a)$ is the complex value of the pixel in the p -th row and the q -th column of ISAR image after translational motion compensation with polynomial coefficient $a = [a_1, a_2, \dots, a_L]^T$, image entropy is defined as

$$EN = - \sum_{p=1}^P \sum_{q=1}^Q \left(|I(p, q, a)|^2 / S \right) \ln \left(|I(p, q, a)|^2 / S \right) \quad (12)$$

where $S = \sum_{p=1}^P \sum_{q=1}^Q |I(p, q, a)|^2$.

Image contrast is defined as

$$IC = \sigma / \mu \quad (13)$$

where

$$\mu = \frac{1}{PQ} \sum_{p=1}^P \sum_{q=1}^Q |I(p, q, a)| \quad (14)$$

$$\sigma = \sqrt{\frac{1}{PQ} \sum_{p=1}^P \sum_{q=1}^Q \left(|I(p, q, a)| - \mu \right)^2} \quad (15)$$

Experiments have shown that IE-based methods get better ISAR images with longer time. While IC-based methods obtain slightly inferior quality images with much less time.

To get better imaging results with less time, a new objective function is proposed in the paper, which is defined as

$$\begin{aligned} \hat{a} &= \arg \min f(a) \\ &= \arg \min \alpha f_{IC}(a) + (1 - \alpha) f_{EN}(a) \end{aligned} \quad (16)$$

where $f_{EN}(a)$ and $f_{IC}(a)$ are image entropy and the reciprocal of image contrast of ISAR image after translational motion compensation with polynomial coefficient $a = [a_1, a_2, \dots, a_L]^T$, and $\hat{a}$ is the estimate of a . It is well known that if the image entropy is smaller, the image quality is better. On the contrary, if the image contrast is bigger, the image quality is better. Therefore, to combine image entropy and image contrast in an objective function, the reciprocal of image contrast rather than image contrast is used in (16). $\alpha = (T_{\max} - t) / T_{\max}$ is the weight, where t is current number of iterations and $T_{\max}$ is maximum number of iterations. Setting weight in this way makes the objective function is mainly based on IC during the initial iterations and mainly based on IE during the later iterations.

3.2 Adam algorithm

Many intelligent optimization algorithms such as PSO and GA have been used to estimate the polynomial coefficients of translational motion. However, they are too time-consuming to get desired results. Compared with intelligent optimization algorithms, gradient-based optimization algorithms are more efficient, which search for the global optimal point based on the gradient of objective function. Among them, stochastic gradient descent (SGD) algorithm is an excellent one [10, 11]. However, it is difficult for SGD to jump out of local optimal points as SGD updates parameters only based on current gradient of objective function. To overcome the shortcoming, Adam was proposed, which updates parameters based on current gradient and previous gradients. Therefore, the first moment (the mean) and the second moment (the uncentered variance) of the gradient are computed in Adam using exponential moving average (EMA) method. Now, Adam is widely used in deep learning.

The pseudo code of Adam is shown in Algorithm 1. First, the gradient of objective function with respect to parameter vector θ_{t-1} is computed. Then, exponential moving average (EMA) of gradient and gradient's square, i.e., $\mathbf{m}_t$ and $\mathbf{v}_t$ are updated, where parameters $\beta_1, \beta_2 \in [0, 1)$ control the exponential decay rates of the moving averages. In fact, $\mathbf{m}_t$ and $\mathbf{v}_t$ are estimates of the mean and the uncentered variance of the gradient. As

$\mathbf{m}_t$ and $\mathbf{v}_t$ are initialized as zero vectors, they are biased towards zeros during the initial iterations. So, they are corrected to $\hat{\mathbf{m}}_t$ and $\hat{\mathbf{v}}_t$ to counteract the initialization bias. Finally, parameter vector θ_{t-1} is updated to θ_t .

Algorithm 1 Adam algorithm	
Input:	Initial parameter vector θ_0 , step size α , regularization constant ε , moment decay parameters $\{\beta_1, \beta_2\}$, objective function $f(\theta)$
Initial:	$\mathbf{m}_0 = 0, \mathbf{v}_0 = 0, t = 0$
Repeat:	while θ_t not converged do $t = t + 1$ $\mathbf{g}_t = \nabla f_t(\theta_{t-1})$ $\mathbf{m}_t = \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \mathbf{g}_t$ $\mathbf{v}_t = \beta_2 \mathbf{v}_{t-1} + (1 - \beta_2) \mathbf{g}_t^2$ $\hat{\mathbf{m}}_t = \mathbf{m}_t / (1 - \beta_1^t)$ $\hat{\mathbf{v}}_t = \mathbf{v}_t / (1 - \beta_2^t)$ $\eta_t = \alpha / (\sqrt{\hat{\mathbf{v}}_t} + \varepsilon)$ $\theta_t = \theta_{t-1} - \eta_t \hat{\mathbf{m}}_t$ end while

Adam has been used in deep learning successfully. Inspired by this, the paper proposes using Adam algorithm to estimate the polynomial coefficients $\hat{\mathbf{a}}$ in problem (16). The iteration is terminated when the following condition is satisfied

$$\|\hat{\mathbf{a}}_t - \hat{\mathbf{a}}_{t-1}\|_2 \leq \delta \quad (17)$$

where δ is a small positive constant which is related to the estimation accuracy.

3.3 Madam based method

Although Adam is widely used in deep learning, it has some disadvantages.

3.3.1 New updating strategy for weights

First, Adam cannot converge fast near global optimal point as exponential moving average (EMA) is used.

The explanation is as follows. In Adam, EMA of the gradient can be written as

$$\mathbf{m}_t = \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \mathbf{g}_t \quad (18)$$

where $\mathbf{g}_t$ is the gradient of objective function, $\mathbf{m}_{t-1}$ and $\mathbf{m}_t$ are estimates of the mean of gradient at last and current iterations, and $0 \leq \beta_1 < 1$ controls the exponential decay rate of the moving average. As $\mathbf{m}_{t-1} = \beta_1 \mathbf{m}_{t-2} + (1 - \beta_1) \mathbf{g}_{t-1}$, (18) can be written as

$$\mathbf{m}_t = (1 - \beta_1) \mathbf{g}_t + (1 - \beta_1) \beta_1 \mathbf{g}_{t-1} + \beta_1^2 \mathbf{m}_{t-2} \quad (19)$$

Similarly, $\mathbf{m}_{t-2}$ in (19) can be replaced by $\mathbf{m}_{t-2} = \beta_1 \mathbf{m}_{t-3} + (1 - \beta_1) \mathbf{g}_{t-2}$. Repeat the above process, then (18) can be expressed as

$$\begin{aligned} \mathbf{m}_t &= (1 - \beta_1) \mathbf{g}_t + (1 - \beta_1) \beta_1 \mathbf{g}_{t-1} + \dots + (1 - \beta_1) \beta_1^{t-1} \mathbf{g}_1 \\ &= (1 - \beta_1) \sum_{i=1}^t \beta_1^{i-1} \mathbf{g}_{t-i+1} \end{aligned} \quad (20)$$

It can be seen that $\mathbf{m}_t$ is not only influenced by current gradient $\mathbf{g}_t$ but also by past gradients $\mathbf{g}_{t-1}, \mathbf{g}_{t-2}, \dots, \mathbf{g}_1$. So, when approaching global optimal point, although the absolute value of current gradient $\mathbf{g}_t$ becomes small, the absolute value of $\mathbf{m}_t$ is still not small for it is also influenced by past gradients. The update formula of parameters θ_t is

$$\theta_t = \theta_{t-1} - \eta_t \hat{\mathbf{m}}_t \quad (21)$$

where $\hat{\mathbf{m}}_t$ is bias-corrected estimate of $\mathbf{m}_t$. So, when approaching global optimal point, it is easy to skip global optimal point due to the large absolute value of $\hat{\mathbf{m}}_t$, which results in more iterations around global optimal point. Therefore, EMA slows down the convergence speed when approaching global optimal point.

An experiment is used to illustrate. A function with noise can be written as

$$f(x) = (\sin(x) - x + 1)(\cos(x) + 2x \sin(x)) + n \quad (22)$$

where n is white Gaussian noise. Adam is used to find x that minimizes $f(x)$. To see the iteration path more clearly, n is set to be zero. The iteration path is shown in Fig. 1, where $\beta_1 = 0.9$, $\beta_2 = 0.99$, $\alpha = 1$ and $\varepsilon = 10^{-8}$. It can be seen Adam has iterated many times around global optimal point before converging to global optimal point.

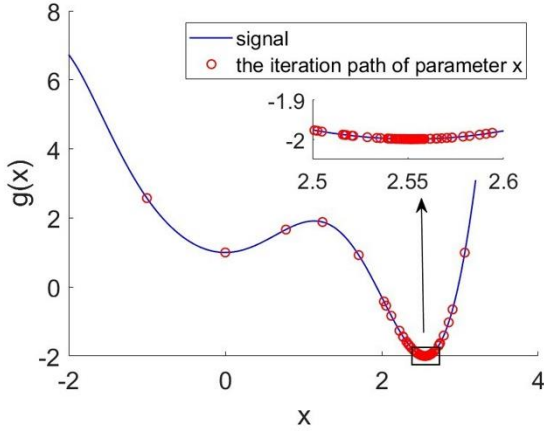


Fig. 1. The iteration path of x based on Adam

To accelerate the convergence around global optimal point, the influence of past gradients should be decreased near global optimal point, i.e., the weights of past gradients in (20) should be decreased when approaching global optimal point. The weights are determined by β_1 , so β_1 is modified in the paper. Here, β_1 is set to be a value that varies with the number of iterations, which is

$$\beta_{1,t} = \beta_{1,t-1} - \delta_1 e^{\max\{0, t/10-1\}} \quad (23)$$

where δ_1 is a very small positive constant, t and $t-1$ are the number of iterations. The new Adam algorithm based on the modification (23) is named improved Adam. Based on (23), in the early stages of iterations, as the decrease of $\beta_{1,t}$ is very small, it maintains the merits of Adam. In the later stages of iterations, as the attenuation of $\beta_{1,t}$ is significant, it makes the updating of $\mathbf{m}_t$ is more affected by current gradient, which is beneficial for fast convergence near global optimal point. In this way, improved Adam accelerates the convergence while maintaining the advantages of Adam.

Improved Adam is also used to find x that minimizes $f(x)$ in (22). The iteration path is shown in Fig. 2, where $\beta_{1,0} = 0.9$, $\beta_2 = 0.99$, $\delta_1 = 0.001$, $\alpha = 1$ and $\varepsilon = 10^{-8}$. It can be easily seen that improved Adam significantly reduces the number of iterations when approaching global optimal point.

3.3.2 New updating strategy for learning rate

The second disadvantage of Adam is that as the learning rate increases when approaching global optimal point, Adam may iterate a lot of times around global optimal point or not converge.

Here is an explanation. When approaching global optimal point, the absolute value of gradient decreases

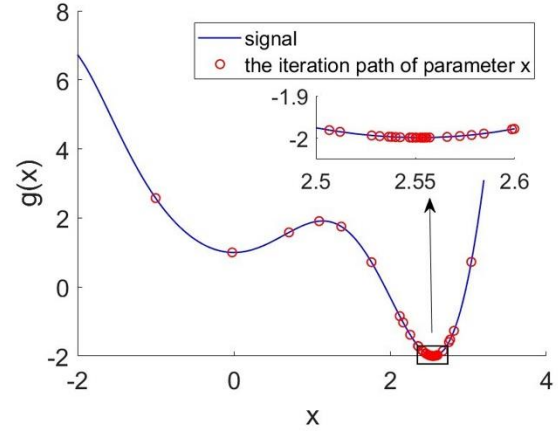


Fig. 2. The iteration path of x based on improved Adam.

as the gradient at global optimal point is zero. As second moment of gradient $\mathbf{v}_t$ is EMA of gradient's square, $\mathbf{v}_t$ decreases too. And learning rate η_t is

$$\eta_t = \alpha / (\sqrt{\hat{\mathbf{v}}_t} + \varepsilon) \quad (24)$$

where $\hat{\mathbf{v}}_t$ is the bias-corrected estimate of $\mathbf{v}_t$. Therefore, the learning rate η_t increases near global optimal point. According to (21), the increase of learning rate η_t near optimal point makes Adam easy to skip global optimal point. Thus, Adam may iterate many times near global optimal point or not converge.

To overcome the shortcoming, the learning rate should be modified to ensure it does not increase near global optimal point. According Algorithm 1, $\hat{\mathbf{v}}_t$ in (24) is computed as

$$\hat{\mathbf{v}}_t = \mathbf{v}_t / (1 - \beta_2^t) \quad (25)$$

where $\beta_2 \in [0, 1)$ controls the exponential decay rate of the moving average.

Substituting $\mathbf{v}_t = \beta_2 \mathbf{v}_{t-1} + (1 - \beta_2) \mathbf{g}_t^2$ into (25), the following is obtained

$$\hat{\mathbf{v}}_t = \beta_2 \mathbf{v}_{t-1} / (1 - \beta_2^t) + \mathbf{g}_t^2 / (1 + \beta_2 + \beta_2^2 + \dots + \beta_2^{t-1}) \quad (26)$$

When approaching global optimal point, to reduce the impact of $\mathbf{g}_t^2$ to get bigger $\hat{\mathbf{v}}_t$, β_2 should be increased. Therefore, β_2 is set to be a value that increases with iterations in the paper, which is

$$\beta_{2,t} = \min\{0.999, \beta_{2,t-1} + \delta_2 e^{\max\{0, t/10-1\}}\} \quad (27)$$

where δ_2 is a very small positive constant. Similarly, using parameter $\beta_{2,t}$ in (27) can not only maintain the advantages of Adam but also accelerate the convergence near global optimal point. To further ensure that $\hat{\mathbf{v}}_t$ does

not decrease near global optimal point, when approaching global optimal point, $\hat{\mathbf{v}}_t$ is modified as

$$\hat{\mathbf{v}}_t = \max(\hat{\mathbf{v}}_{t-1}, \hat{\mathbf{v}}_t) \quad (28)$$

Based on the above modifications, the learning rate η_t doesn't increase any longer near global optimal point. The Adam algorithm with the above modifications for β_1 and β_2 is called modified Adam (Madam) algorithm and its pseudo code is shown in Algorithm 2.

Algorithm 2 Madam algorithm	
Input:	Initial parameter vector θ_0 , step size α , regularization constant ε , moment decay parameters $\{\beta_{1,0}, \beta_{2,0}\}$, objective function $f(\theta)$, small positive constant $\{\delta_1, \delta_2\}$, iterations near optimal point T
Initial:	$\mathbf{m}_0 = \mathbf{0}$, $\mathbf{v}_0 = \mathbf{0}$, $t = 0$
Repeat:	while θ_t not converged do $t = t + 1$ $\beta_{1,t} = \beta_{1,t-1} - \delta_1 e^{\max\{0, t/10-1\}}$ $\beta_{2,t} = \min\{0.999, \beta_{2,t-1} + \delta_2 e^{\max\{0, t/10-1\}}\}$ $\mathbf{g}_t = \nabla f_t(\theta_{t-1})$ $\mathbf{m}_t = \beta_{1,t} \mathbf{m}_{t-1} + (1 - \beta_{1,t}) \mathbf{g}_t$ $\mathbf{v}_t = \beta_{2,t} \mathbf{v}_{t-1} + (1 - \beta_{2,t}) \mathbf{g}_t^2$ $\hat{\mathbf{m}}_t = \mathbf{m}_t / (1 - \beta_{1,t}^t)$ $\hat{\mathbf{v}}_t = \mathbf{v}_t / (1 - \beta_{2,t}^t)$ if $t \geq T$ $\hat{\mathbf{v}}_t = \max(\hat{\mathbf{v}}_{t-1}, \hat{\mathbf{v}}_t)$ endif $\eta_t = \alpha / (\sqrt{\hat{\mathbf{v}}_t} + \varepsilon)$ $\theta_t = \theta_{t-1} - \eta_t \hat{\mathbf{m}}_t$ end while

3.3.3 Translational motion parameters estimation

Madam has significantly accelerated the convergence of Adam, so it can be used to estimate the polynomial coefficients of translational motion efficiently. And the steps of the proposed parametric translational motion compensation method are as follows.

- (1) To simulate complex translational motion and low SNR scenes, translational motion modelled by fourth-order polynomial and white Gaussian noise are added to measured ISAR echo.
- (2) Based on the ISAR echo, Madam algorithm is used to optimize objective function (16), where the initial value of $\hat{\mathbf{a}}$ is set to be zero vector.

- (3) Compensation function (10) is constructed using the estimate of $\mathbf{a}$ and the translational motion is compensated as (11).
- (4) A focus ISAR image is obtained by using inverse Fourier transform along range direction and Fourier transform along azimuth direction.

4 Experimental results and discussion

To verify the effectiveness of the proposed translational motion parameters estimation method, it is applied into measured Yak-42 aircraft ISAR data collected by a ground-based radar. The radar parameters are shown in Tab. 1.

Table 1. Parameters of ISAR

Parameter	Value
Carrier frequency	5.52 GHz
Bandwidth	400 MHz
Pulse width	25.6 μ s
Pulse repetition frequency	400 Hz
Sampling frequency	10 MHz

In the following experiments, translational motion and complex white Gaussian noise are added to the Yak-42 data. Translational motion is usually modelled as three-order polynomial. However, the cubic model may not accurately describe the target's actual movement. To break through this limitation, fourth-order polynomial is employed in the paper, which is a reasonable trade-off between computational complexity and estimation accuracy [12].

First, imaging performance based on different objective functions is compared. For Adam-E (Adam based on image entropy), Adam-C (Adam based on image contrast) and Adam based on the proposed new objective function, the average performance indicators of 100 experiments from -5 dB to 5 dB are shown in Figs. 3 to 5, respectively. It is evident that compared with Adam-C, the method based on new objective function obtains better imaging quality by using the similar time. Compared with Adam-E, the new method takes about half the time to get the similar imaging quality. Therefore, the proposed new objective function is used as objective function in the following.

Second, imaging performance of Madam is compared with that of conventional methods. To assess the estimation accuracy of motion parameters, mean-squared error (MSE) is used, which is defined as

$$MSE = 1/I \cdot \sum_{i=1}^I \|\hat{\mathbf{a}}_i - \mathbf{a}\|_2^2 \quad (29)$$

where I is the number of experiments at each SNR and set to be 20 here, $\mathbf{a}$ and $\hat{\mathbf{a}}_i$ are the set translational motion polynomial coefficient vector and its estimation in experiment i respectively. The average results are

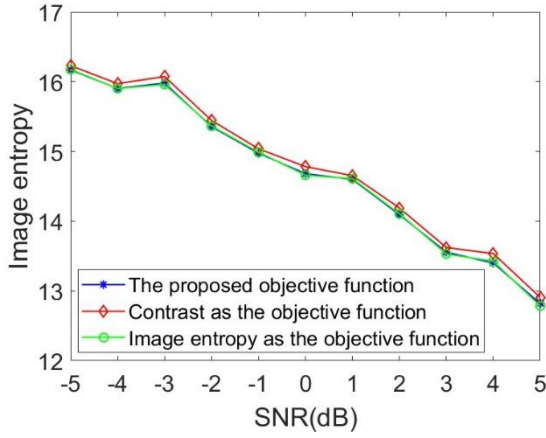


Fig. 3. The comparison of image entropy

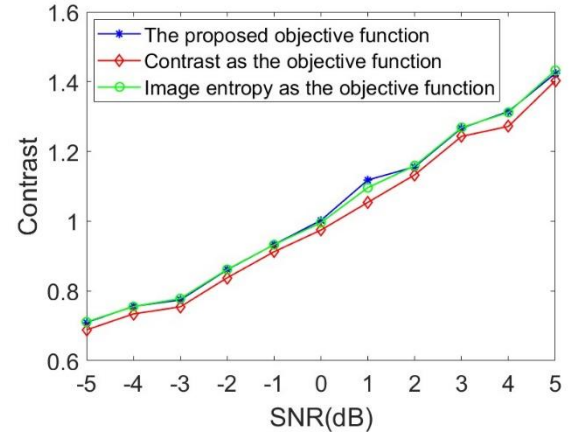


Fig. 4. The comparison of image contrast

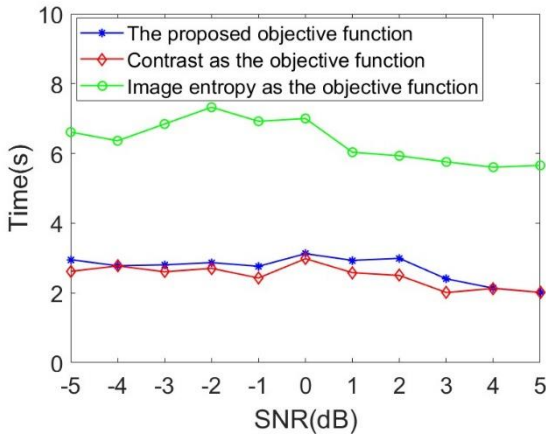


Fig. 5. The comparison of imaging time

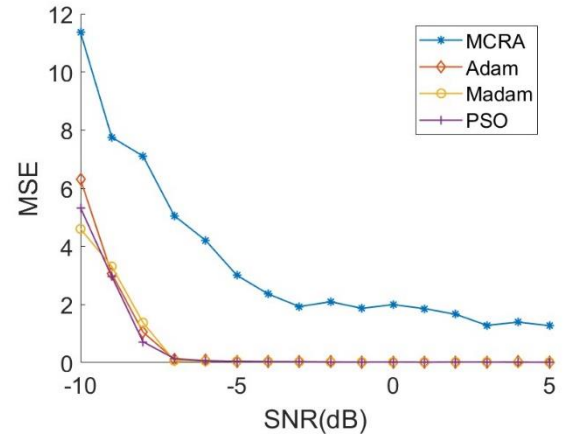


Fig. 6. MSE of different methods

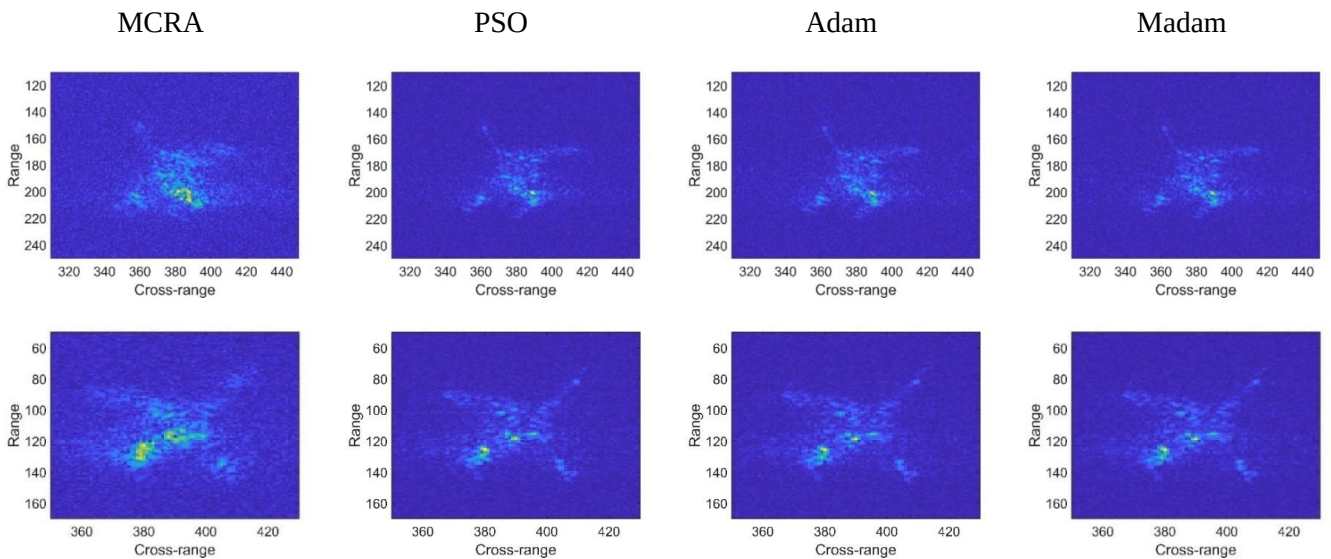


Fig. 7. Imaging results based on different methods at SNR of -5 dB

shown in Fig. 6. It can be seen MCRA cannot estimate the motion parameters well under low SNR. However, PSO, Adam and Madam all estimate the motion parameters accurately when SNR is higher than -7 dB.

For MCRA, PSO, Adam and Madam, at different time periods, the imaging results of Yak-42 at $SNR = -5dB$ are shown in Fig. 7. It can be seen the images based on MCRA defocus as MCRA cannot accurately estimate the motion parameters under low SNR. However, PSO, Adam and Madam all get focused images because they can estimate the motion parameters very well. The image entropy, image contrast and

imaging time of PSO, Adam and Madam are presented in Tab. 2. The image quality based on the three methods is comparable. However, Madam has significant advantage in imaging time. Compared with Adam and PSO, Madam has improved imaging efficiency about 2 times and 10 times respectively. Hence, Madam can achieve well-focused ISAR images with high imaging efficiency at low SNR.

Table 2. Comparison of Yak-42 imaging results at SNR of $-5dB$

Yak-42 data	Indicators	PSO	Adam	Madam
Time period 1	Entropy	16.1807	16.1979	16.1996
	Contrast	0.6929	0.6919	0.6923
	Time(s)	17.7971	3.1036	1.9030
Time period 2	Entropy	15.7598	15.7702	15.7640
	Contrast	0.7556	0.7530	0.7532
	Time(s)	15.8380	2.5474	1.1156

To further illustrate the advantages of Madam, 100 experiments at each low SNR are done and the average results are shown in Fig. 8. It can be clearly seen Madam only takes about one tenth of PSO's time or half of Adam's time.

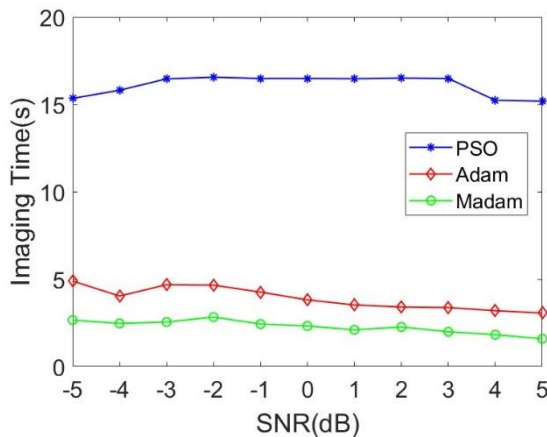


Fig. 8. Imaging time comparison of three methods

5 Conclusion

A fast ISAR translational motion compensation method is proposed for low SNR scenes. Madam algorithm is proposed to estimate motion polynomial coefficients, which modified the updating strategy of weights and learning rate of Adam to significantly accelerate the convergence. And a new objective function combining image contrast and image entropy is presented to improve both imaging efficiency and image quality. Imaging results of measured data show that the

proposed method can get well-focused images with much higher imaging efficiency than PSO and Adam.

Acknowledgement

The authors wish to thank Harbin Institute of Technology for providing Yak-42 aircraft measured data for this research project.

References

- [1] X. Wei, J. Yang, M. Lv, W. Chen, and X. Ma, "A reweighted alternating direction method of multipliers for joint phase adjustment and ISAR imaging", *Remote Sensing Letters*, vol. 13, no. 10, pp. 1020-1028, 2022.
- [2] S. Zhang, Y. Liu and X. Li, "Fast entropy minimization based autofocusing technique for ISAR imaging", *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3425-3434, 2015.
- [3] S. Han, X. Yang and W. Sheng, "An improved minimum entropy range alignment algorithm for airborne ISAR imaging of aerial targets", *2023 IEEE 6th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE)*, Shenyang, China, pp. 451-456, 2023.
- [4] Y. Zhang, W. Wang, Z. Guo, and N. Li, "Enhanced PGA for dual-polarized ISAR imaging by exploiting cloude-pottier decomposition", *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [5] F. Liu, D. Huang, X. Guo, and C. Feng, "A robust translational motion compensation method based on weighted optimization framework—second-order drift particle swarm optimization", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16, 2022.
- [6] F. Liu, D. Huang, X. Guo, and C. Feng, "Translational motion compensation for maneuvering target echoes with sparse aperture based on dimension compressed optimization", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16, 2022.

- [7] D. Kingma and J. Ba, "Adam: a method for stochastic optimization", *2015 3th International Conference on Learning Representations*, San Diego, CA, USA, pp. 1-15, 2015.
- [8] Y. Gao, M. Xing, Y. Li, W. Sun, and Z. Zhang, "Joint translational motion compensation method for ISAR imagery under low SNR condition using dynamic image sharpness metric optimization", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-15, 2022.
- [9] T. Yang, H. Shi, J. Guo, and M. Liu, "FMCW ISAR autofocus imaging algorithm for high-speed maneuvering targets based on image contrast-based autofocus and phase retrieval", *IEEE Sensors Journal*, vol. 20, no. 3, pp. 1259-1267, 2020.
- [10] S. Vlaski and A. Sayed, "Second-order guarantees of stochastic gradient descent in nonconvex optimization", *IEEE Transactions on Automatic Control*, vol. 67, no. 12, pp. 6489-6504, 2022.
- [11] V. Gandikota, D. Kane, R. Maity, and A. Mazumdar, "vqSGD: vector quantized stochastic gradient descent", *IEEE Transactions on Information Theory*, vol. 68, no. 7, pp. 4573-4587, 2022.
- [12] K. Abratkiewicz, J. Drozdowicz and P. Samczyński, "Vertical synchrosqueezing for high-resolution radar imaging", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-13, 2023.

Ping Cheng was born in Yancheng, Jiangsu, China, in 1976. She received the B.S. degree in electronic engineering, the M.S. degree in communication and information processing, and the Ph.D. degree in electronic and communication system from Harbin Institute of Technology, Harbin, China, in 2000, 2002 and 2007, respectively. Her research interests include radar imaging and compressed sensing.

Xinkai Yang was born in Tianshui, Gansu, China in 2000. He received the B.S. degree in electronic information engineering from Hohai University, Nanjing, China, in 2022, and is currently pursuing the M.S. degree in signal and information processing in this university. His research interests include radar imaging and compressed sensing.

Weiyang Chen was born in Suqian, Jiangsu, China in 1996. He received the B.S. degree in electronic information engineering from Jiangsu University of Technology, Changzhou, China, in 2020 and the M.S. degree in electronic science and technology in Hohai University, Nanjing, China, in 2024. His research interests include radar imaging and compressed sensing.

Jiaqun Zhao was born in Qiqihar, Heilongjiang, China in 1976. He received the B.S. degree in physics from Northeast Forestry University in 1999, and the M.S. degree in optics and the Ph.D. degrees in physical electronics from Harbin Institute of Technology in 2005 and 2011. His research interests include solid state laser and nonlinear optical frequency conversion.

Zhiyu Qu received the Ph.D. degree in communication and information systems from Harbin Engineering University, Heilongjiang, China, in 2008. Her current research interests include radar signal processing and recognition.

Received 13 August 2024
