(Received 10 December 2024, Accepted 23 January 2025, Scheduled 2 February 2025)

General Array Imaging Network for Near-Field Synthetic Aperture Interferometric Radiometer

Chenggong Zhang[1], Jianfei Chen[1,3,*], Jiahao Yu[1], Yujie Ruan[1], Sheng Zhang[1], Shujin Zhu[2], and Leilei Liu[1]

[1]*College of Electronic and Optical Engineering and the College of Flexible Electronics (Future Technology) Nanjing University of Posts and Telecommunications, Nanjing 210023, China*
[2]*College of Automation and College of Artificial Intelligence Nanjing University of Posts and Telecommunications, Nanjing 210023, China*
[3]*State Key Laboratory of Millimeter Waves, Nanjing 210096, China*

ABSTRACT: Millimeter wave synthetic aperture interferometric radiometer (SAIR) can achieve high-resolution imaging without a large physical aperture antenna and has strong application advantages in the fields of earth remote sensing, astronomical observation, and meteorological monitoring. In order to adapt to various payload platforms and detection needs, the existing SAIR array structures are diverse, but the existing imaging methods are difficult to effectively deal with various arrays and achieve stable high-precision imaging inversion. Thus, this paper proposes a general multi-channel fusion imaging network to achieve SAIR imaging inversion of any array structure. First, with the help of the sensor matrix deduction subnet, a high-precision imaging sensor matrix is deduced according to the position of each array element of the SAIR system, and then high-precision image reconstruction is achieved with the help of the multi-channel fusion imaging subnet. The simulation results show that the network has good adaptability and can achieve high-precision imaging inversion of different SAIR array structures.

1. INTRODUCTION

Millimeter wave imaging technology uses electromagnetic waves with a wavelength in the millimeter range for imaging and plays an important role in remote sensing, radar, security inspection, target detection, and many other fields. However, due to the inherent contradiction between antenna aperture and spatial resolution, traditional real aperture millimeter wave imaging technology is difficult to achieve ideal high-resolution imaging [1]. SAIR uses aperture synthesis technology to form a virtual large aperture antenna using a group of small antennas, thereby achieving high-resolution imaging observation. It is an ideal high-precision imaging detector [2]. Since the advent of the first SAIR (ESTAR), research institutions in various countries have successively developed a variety of SAIR models for various detection applications. For example, Y-array MIRAS and GeoSTAR, circular array GIMS, T-array BHU-2D, and U-array HUT-2D. Most of these traditional arrays are designed based on traditional experience and are designed for specific needs and payload platform limitations. The layout of the SAIR array antenna directly determines the distribution state of the measured visibility function (baseline), which in turn affects the performance of the SAIR imaging detection. For regularly distributed antenna arrays, a large number of redundant baselines are easily generated, which will reduce the SAIR imaging effect to a certain extent [3]. Irregular antenna arrays can effectively increase the longest baseline and improve the spatial resolution through flexible antenna layout. In recent years, such irregular arrays have gradually become a hot topic in SAIR research [4–6].

Although array structures are becoming increasingly diverse, existing imaging methods are unable to effectively cope with various array structures and achieve stable high-precision imaging inversion. Traditional imaging methods are mostly designed for specific arrays, such as MFFT for T-array and U-array, and HFFT for Y-array. However, the lack of periodicity in nonuniform antenna arrays results in irregular distribution of visibility functions, making traditional uniform array imaging methods inapplicable [7]. Currently available methods are mainly divided into two categories: 1) FFT methods based on interpolation-resampling, such as NUFFT and gridding [8] method. When dealing with irregular sampling points, these methods suffer from insufficient accuracy due to errors introduced by interpolation and resampling. 2) Imaging methods based on the G matrix [9]. These methods require the construction of a sensor matrix first and reconstruct the brightness temperature image through an iterative algorithm. While these methods can capture target details more accurately, the reconstruction effect is limited under the influence of complex environmental interference and noise. Under near-field SAIR conditions, the G matrix model has strict accuracy requirements, and slight deviations can lead to serious reconstruction errors. In addition, the iterative process requires significant computing resources and time, which may cause error accumulation and affect the final imaging quality.

* Corresponding author: Jianfei Chen (chenjf@njupt.edu.cn).

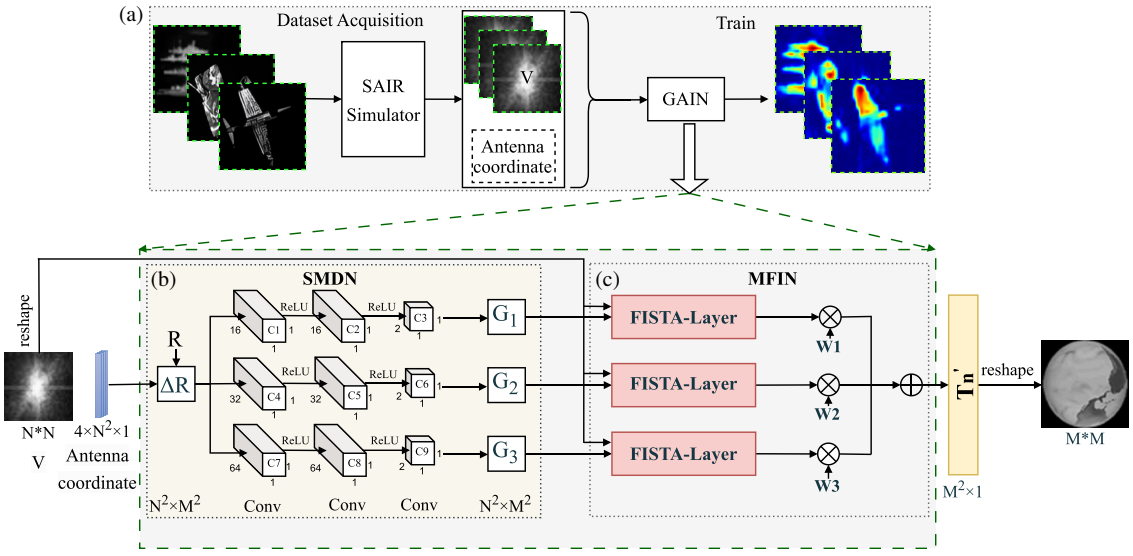

FIGURE 1. GAIN structure diagram. (a) Dataset acquisiton and train. (b) SMDN. (c) MFIN.

In recent years, deep learning technology has shown strong potential in SAIR imaging problems [10–14]. For example, Zhang et al. proposed SAIR-CNN, which combined deep learning with the SAIR imaging process and achieved remarkable results. The IASR-CNN proposed by Xiao et al. provides a high-precision solution for nonuniform antenna arrays. These research results demonstrate the effectiveness of deep learning in SAIR imaging tasks. In view of this, this paper constructs a general array imaging network (GAIN) based on the idea of deep learning to adapt to the SAIR of any array structure and achieve stable high-precision imaging. Through in-depth analysis of the SAIR imaging principle, a sensor matrix deduction subnet based on Conv-ReLU is designed. This subnet uses a multi-layer convolutional network to gradually extract and enhance the correlation between input features and deduce a high-precision imaging sensor matrix based on the position of each array element of the SAIR system. Then, based on the image inversion process of SAIR, a multi-channel fusion imaging subnet is constructed. In the fast iterative shrinkage-thresholding algorithm (FISTA) inversion stage, the residual dense block and attention mechanism are introduced to effectively filter out the noise, and then the multi-channel fusion idea is used to further improve the imaging inversion accuracy. The simulation results show that the proposed GAIN method can adapt to SAIR with any different array structure and achieve higher quality imaging inversion.

2. DESCRIPTION OF GAIN

This paper proposes a GAIN method through a systematic study of the mechanism of synthetic aperture imaging. The structure of this method is shown in Fig. 1, which mainly includes the sensor matrix deduction subnet (SMDN) and multi-channel fusion imaging subnet (MFIN). In this framework, the sensor matrix G of each channel is deduced through the convolution activation operation, and then input into the FISTA-Layer together with the visibility function V to achieve high-precision imaging inversion. In order to cope with the complexity of the SAIR inversion process, the network design adopts a multi-channel fusion strategy to independently adjust each channel matrix G and network parameters, so as to achieve higher quality image reconstruction.

2.1. Data Set Acquisition

The basic unit of the synthetic aperture imaging system is a binary interferometer, which obtains the corresponding visibility function sampling points through complex correlation calculations between two antennas. To simplify the analysis, it is assumed that there is an extended source target S in space, which is simplified to a surface radiation source located on the oxy plane (as shown in Fig. 2). This plane is R away from another parallel OXY plane, on which two antenna array elements A_c and A_l are placed to receive the electromagnetic wave signals radiated by the target. After complex correlation processing, these two signals form a basic binary interferometer.

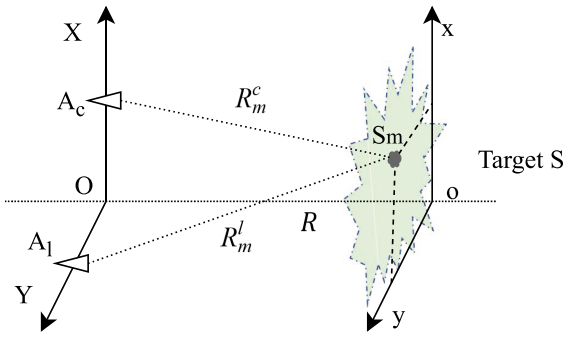

FIGURE 2. Schematic of binary coherence measurement.

The radiation signals of the radiation source S_m at the two antennas are $E(m, A_c, t)$ and $E(m, A_l, t)$. The normalized antenna patterns (voltage transfer function) of antennas A_c and A_l are $F_c(x, y)$ and $F_l(x, y)$, respectively. The output volt-

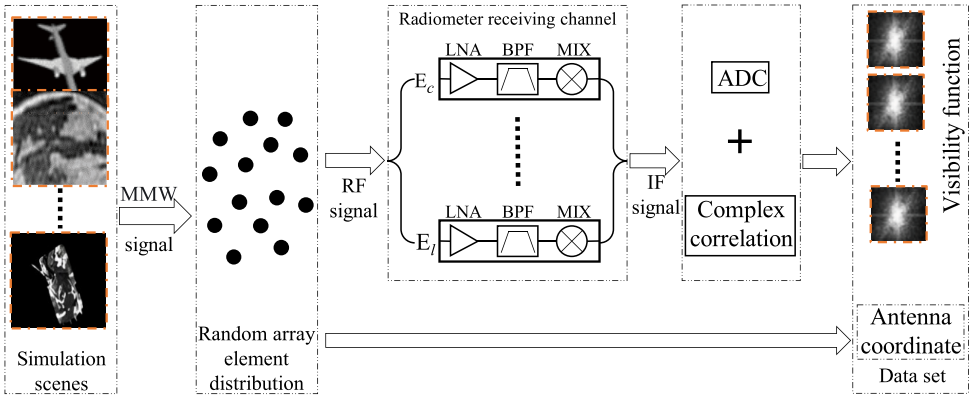

FIGURE 3. SAIR simulator.

age waveforms $U_c(t)$ and $U_l(t)$ of the two antennas can be expressed as:

$$U_c(t) = \sum_{m=1}^{M} F_c(x_m, y_m) E(m, A_c, t) \quad (1)$$

$$U_l(t) = \sum_{m=1}^{M} F_l(x_m, y_m) E(m, A_l, t) \quad (2)$$

where (x_m, y_m) represents the coordinates of the mth point source S_m. According to the principle of coherent measurement, the complex cross-correlation operation of the voltage signals output by the two antennas can yield the cross-correlation function corresponding to the antenna pair (A_c-A_l):

$$V_{c,l} = \langle U_c(t) U_l^*(t) \rangle \quad (3)$$

Substituting the expressions for $U_c(t)$ and $U_l(t)$, we obtain:

$$V_{c,l} = \sum_{m=1}^{M} F_c(x_m, y_m) F_l^*(x_m, y_m) \langle E(m, A_c, t) E^*(m, A_l, t) \rangle \quad (4)$$

where $\langle \cdot \rangle$ represents the time integration operator. According to [15], the relationship between the cross-correlation function (i.e., the visibility function) of the signals received by the two antennas and the brightness temperature distribution of the target scene can be expressed as:

$$V_{c,l} = \sum_{m=1}^{M} T(x_m, y_m) F_c(x_m, y_m) F_l^*(x_m, y_m) \times r_{c,l} e^{-jk\left(R_m^c - R_m^l\right)} \quad (5)$$

where R_m^c and R_m^l are the distances from antennas A_c and A_l to the radiation source S_m; $T(x_m, y_m)$ represents the normalized brightness temperature; k is the wave number $k = \frac{2\pi}{\lambda}$ (where λ is the central wavelength of SAIR); $r_{c,l}$ is the fringe-wash function; and the exponential term represents the critical path difference ΔR in synthetic aperture imaging.

According to the geometric relationship diagram shown in Fig. 2, the exact expressions of the distances R_m^c and R_m^l can be obtained as follows:

$$R_m^c = \sqrt{(x_m - x_c)^2 + (y_m - y_c)^2 + R^2} \quad (6)$$

$$R_m^l = \sqrt{(x_m - x_l)^2 + (y_m - y_l)^2 + R^2} \quad (7)$$

From (6) and (7), we can get the path difference ΔR:

$$\Delta R = R_m^c - R_m^l \quad (8)$$

In far-field imaging, the distances R_m^c and R_m^l can be Taylor approximated. However, for near-field imaging, which requires higher model accuracy, this approximation is not applicable. Refs [16, 17] point out that the path difference ΔR can be directly used to construct the G matrix. Substituting (6) and (7) into (5) yields an accurate synthetic aperture imaging model:

$$V_{N\times 1} = G_{N\times M} T_{M\times 1} \quad (9)$$

$$G(n, m) = F_c(x_m, y_m) F_l^*(x_m, y_m) r_{c,l} e^{j\pi\left(R_m^c - R_m^l\right)/\lambda} \quad (10)$$

where V represents the visibility function, T the brightness temperature distribution vector, and G the precise near-field imaging matrix. (9) describes the relationship between visibility data and brightness temperature image. Regularization can be used to invert (9) to achieve SAIR image reconstruction. The reconstruction model can be described as:

$$\min_T E(T) = \frac{1}{2} \|GT - V\|_2^2 + \sigma P(T) \quad (11)$$

where $\|GT - V\|_2^2$ is the fidelity term, $P(T)$ the regularization term, and σ the regularization parameter, which is used to balance the fidelity term and regularization term.

Based on the above description of the principle of the synthetic aperture imaging system, in this study, we use the SAIR simulator shown in Fig. 3 to generate the data set. The simulator aims to accurately reproduce the signal processing flow in the actual SAIR system. Firstly, the imaging simulator inputs a gray image of 60×60 size, and a random antenna array composed of 30 elements is used to collect the millimeter wave radiation signal of the target scene. The millimeter wave radiation signal is integrated to obtain signals E_c and E_l. According to the linear superposition principle of the radiation

signal, the radio frequency (RF) signal received by each array element is simulated. Then, through a series of signal processing steps, including signal amplification, filtering, and mixing, the intermediate frequency (IF) signal of the radiometer is obtained. Finally, the visibility function V is obtained by Analog-to-Digital Converter (ADC) operation and complex correlation operation. The simulator finally outputs V and antenna coordinates to complete the whole simulation process.

Table 1 lists the specific parameter settings of this simulation in detail. Through the above steps, the simulator can efficiently generate the required visibility function and its corresponding coordinate data. The specific simulation steps used in this paper are designed as follows:

1. Select 5,000 natural scene images of different categories from ImageNet and convert them into 60×60 grayscale images as brightness temperature images T.
2. Use the SAIR simulator to load image T and generate the corresponding simulated visibility function $V(u, v)$ and antenna coordinates.
3. Repeat step 2 to generate 5,000 samples and randomly divide them into training sets (80%) and test sets (20%) for network training and performance evaluation.

TABLE 1. Simulation parameters of SAIR.

Parameters	Wavelength	Antenna spacing
Value	3 mm	1 cm
Parameters	Imaging distance	Source spacing
Value	10 m	7 mm

2.2. General Array Imaging Network

1) Sensor matrix deduction subnet: In the field of synthetic aperture imaging, the accuracy of the sensing matrix G is a key factor affecting the imaging inversion effect of the synthetic aperture interferometric radiometer (SAIR), which is closely related to the position of the antenna array element, imaging distance R, and antenna pattern F. The G constructed by (10) has high description accuracy, but even small parameter errors may cause significant deviations in the images solved by the traditional FISTA method, and even the imaging results cannot be obtained. Therefore, deriving an accurate sensing matrix is a key step in image reconstruction. To this end, we propose a sensor matrix deduction subnet (SMDN) based on convolutional neural network (CNN) to derive an accurate sensor matrix G through network learning.

Network structure design: The input of SMDN is the position information of each array element. Firstly, the wave path difference matrix ΔR is initialized according to (8). On this basis, the complex relationship between ΔR and sensing matrix G is learned. The network gradually extracts and transforms input features through three sets of CNNs to generate an accurate sensing matrix. The convolutional layers in each group of CNN use 1×1 convolution kernels, combined with ReLU activation function, to enhance the nonlinear expression ability of the network. The input of each group of CNN is a single-channel ΔR, and the final output is a dual-channel (G matrix is complex) sensing matrix G. The multi-layer features of CNN contain a lot of information, so the effective use of multi-layer features is crucial to improving feature representation [18, 19].

- *The first group of CNN:* The number of convolutional layer channels responsible for feature extraction is 16. The role of this group of CNNs is to extract basic features from the input data, maintain the details of the information, and respond quickly to changes in the input signal.
- *The second group of CNN:* The number of convolutional layer channels responsible for feature extraction is increased to 32 to capture more features.
- *The third group of CNN:* The number of channels is further increased to 64 to capture more detailed and rich features. Through multiple sets of convolutional layers, the network can better learn the deep relationship between the input and sensing matrix, and further improve the output accuracy.

We take the initial matrix G_0 in (10) as the prior matrix and optimize it during the training process. After many experiments, this initialization can significantly improve the convergence speed of the network and ensure high precision.

2) Multi-channel fusion imaging subnet: After obtaining the accurate sensing matrix G, we design a high-precision imaging subnet, combined with FISTA [20] iterative inversion, residual attention dense block (RADB), and multi-channel fusion strategy [21–23] to achieve high-quality image reconstruction.

FISTA iterative inversion: The core idea of the FISTA algorithm is to update the variables along the negative direction of the gradient and gradually approach the local optimal solution. The specific structure of FISTA-Layer shown in Fig. 1(c) is shown in Fig. 4, which can be divided into the following key steps:

$$s_k = y_k - \alpha A^T(Ay_k - b) \tag{12}$$

$$T_k = \Gamma(s_k) \tag{13}$$

$$y_{k+1} = T_k + \beta(T_k - T_{k-1}) \tag{14}$$

where A is the sensing matrix; T is the target brightness temperature vector to be solved; b is the observed visibility function vector; A^T is the transpose of the sensing matrix; s_k calculates the negative direction of the gradient, minimizing the objective function $\|AT - b\|_2^2$; y_{k+1} is the iterative update rule; $\Gamma(\cdot)$ is the denoising operator; and the residual attention dense block is used to replace the soft threshold function in the traditional FISTA. FISTA-Layer achieves end-to-end training of this process by introducing a deep learning network. The training parameters α (step length) and β (update weight) of the model are automatically learned by the network, thereby accelerating convergence and improving the denoising effect.

Residual attention dense block (RADB): In order to improve the model's ability to focus on key information and suppress irrelevant information, the RADB module combines residual dense block (RDB) [24] and global attention mechanism (GAM) [25]. RDB uses a dense connection strategy to share convolutional layer features and enhance image reconstruction. The GAM module improves the network's ability to

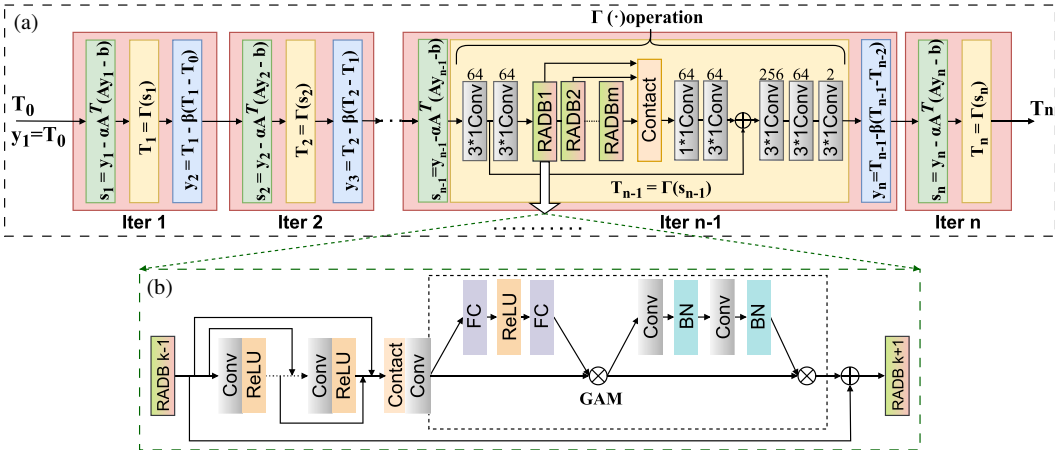

FIGURE 4. The structure of FISTA-Layer. (a) FISTA solving process. (b) RADB.

focus on important areas through channel and spatial attention mechanisms:

- *Channel attention:* Generate channel weights through a fully connected network to help the model adaptively adjust the importance of different channels.
- *Spatial attention:* Spatial weights are generated by a 1D convolution operation and Batch Normalization, so that the network can focus on the key areas of the image.

Multi-channel fusion: We fuse the three independent images obtained by FISTA-Layer inversion through a multi-channel fusion network. Each layer of the network has a trainable weight coefficient W_1, W_2, W_3. These coefficients adaptively adjust the fusion weight of each channel and combine the advantages of different channels to further improve the image quality.

In the GAIN method, considering the complexity and depth of the network, parameter N is set to 30, and M is set to 60, which can avoid excessive calculation while ensuring the extraction of multi-scale features. The number of iterations of FISTA-Layer is set to 3 times. This configuration has been verified by many experiments and can accelerate the training process under the premise of ensuring image quality.

3. EXPERIMENTS AND RESULTS

3.1. Experiments

The input of the GAIN network is the visibility function $V(u, v)$ and the position of each array element of the system. The output is the predicted image T_{pred}, and the learning target is the original brightness temperature scene T_{target}. In image processing, the root mean square error (RMSE) can intuitively measure the degree of pixel difference between the processed image and original image, so we select $\text{Loss}_{\text{RMSE}}$ as the loss function of the network. The GAIN training process is shown in Fig. 5. After multiple iterations, the network can automatically adjust the internal parameters according to the loss function and gradually improve the inversion accuracy, minimizing the RMSE between the predicted image T_{pred} and original brightness temperature scene T_{target}. When the number of training epochs reaches

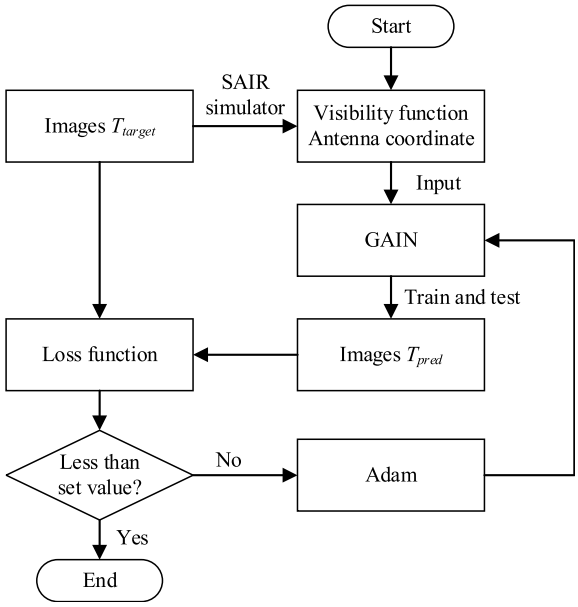

FIGURE 5. Training process of the GAIN.

the preset upper limit, or RMSE tends to converge, the training process terminates. If not, the network will perform back-propagation through the adaptive moment estimation (Adam) optimization algorithm to update the network parameters and further reduce the RMSE. The entire network is built based on the PyTorch framework and trained on a workstation equipped with a Xeon (R) 6130 CPU and Tesla-V100 GPU.

$$\text{LOSS}_{\text{RMSE}} = \sqrt{\frac{1}{HW}\sum_{i=0}^{H-1}\sum_{j=0}^{W-1}\left[T(i,j) - T_0(i,j)\right]^2} \quad (15)$$

where $T(i, j)$ represents the pixel value of the original bright temperature image T at coordinate (i, j), and $T_0(i, j)$ represents the pixel value of the predicted image T_0 at coordinate (i, j). H and W are the sizes of T and T_0.

In order to comprehensively evaluate the effectiveness of the GAIN algorithm, this paper designs imaging simulation experiments for different antenna arrays. Fig. 6(I), Fig. 6(II),

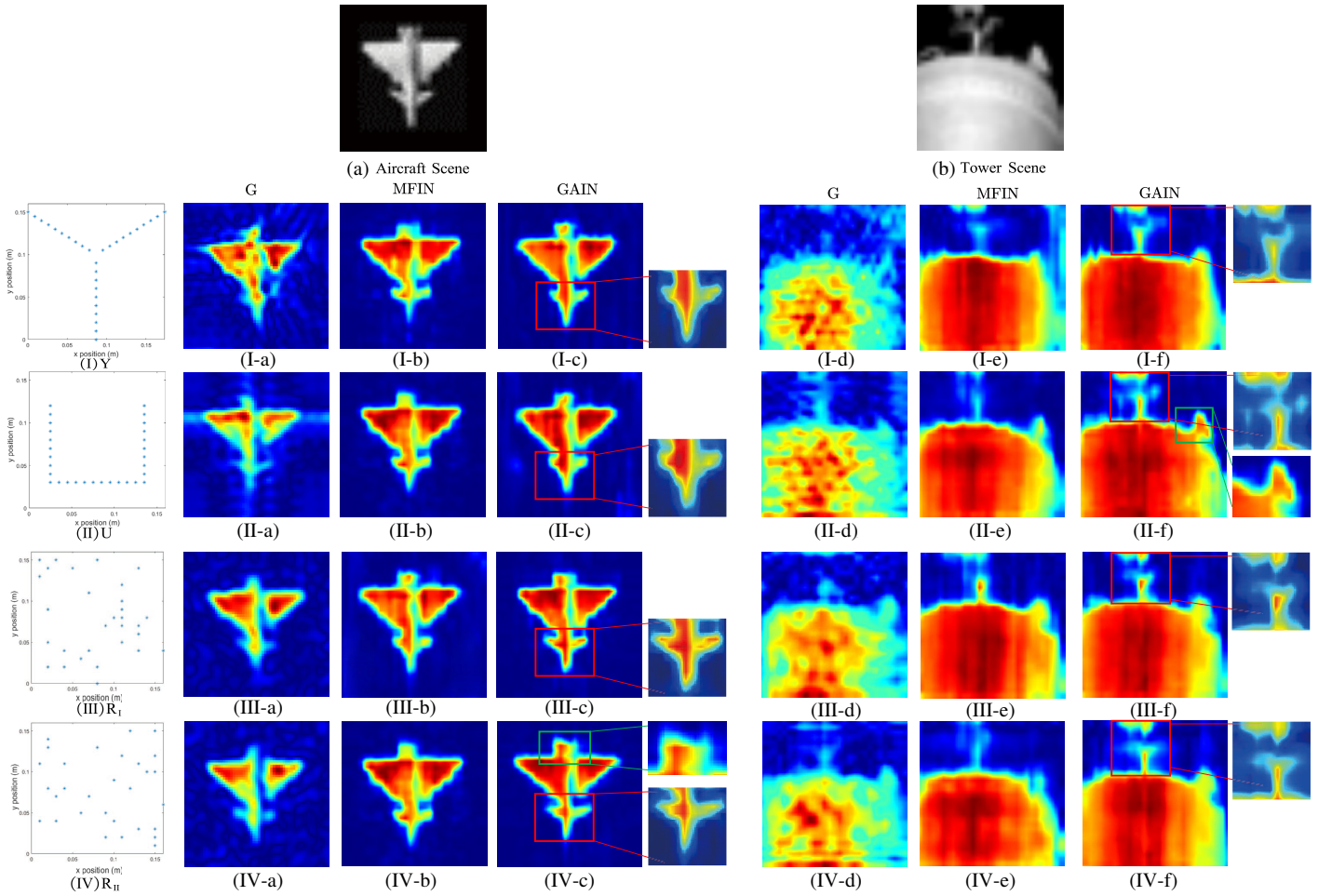

FIGURE 6. 2D scene image reconstruction. (I)–(IV) Antenna array distribution. (a) Aircraft scene. (b) Tower scene.

Fig. 6(III), and Fig. 6(IV) show various antenna arrays, corresponding to Y-shaped array, U-shaped array, random array I (R_I), and random array II (R_{II}), respectively.

In the aircraft scene, the three-component imaging inversion experiment is designed using the simulation parameters shown in Table 1. The first group of experiments uses the traditional G matrix method for imaging. Firstly, according to different array antenna layouts (array element coordinates), the corresponding G matrix (G-Y, G-U, G-RR_I, G-RR_{II}) and visibility function V (V-Y, V-U, V-R_I, V-R_{II}) are calculated, and the brightness temperature image is reconstructed by the regularization iterative algorithm. The second group uses the GAIN method without SMDN (MFIN) for training imaging (that is, the G matrix is not optimized). The G matrix and visibility function V corresponding to each array are directly input into MFIN for iterative inversion imaging. The third group of experiments uses the complete GAIN method proposed by us for imaging inversion. Firstly, with the help of SMDN, the matched high-precision sensor G matrix is adaptively generated according to the antenna coordinate data, and then input into the MFIN with the visibility function V for inversion imaging. The imaging results are shown in Figs. 6(I-a), (I-b), (I-c) to (IV-a), (IV-b), (IV-c).

In order to further verify the versatility of the GAIN algorithm, a three-component image inversion comparison experiment was designed using the simulation parameters shown in Table 1 in a more complex tower scene. The imaging results are shown in Figs. 6(I-d), (I-e), (I-f) to (IV-d), (IV-e), (IV-f).

3.2. Results and Analysis

From the simulation results of (I-a), (II-a), (III-a), (IV-a) and (I-d), (II-d), (III-d), (IV-d) in Fig. 6, it can be seen that the traditional G matrix method can perform imaging inversion for different array forms, but such methods have strict requirements on the accuracy of the model, weak anti-noise ability, and unstable inversion performance. The reconstructed image has large noise interference, unclear outline, and serious loss of target information; Figs. 6(I-b), (II-b), (III-b), (IV-b) and (I-e), (II-e), (III-e), (IV-e) are the imaging results of the MFIN method. This method reduces the background noise interference in the imaging process and improves the image quality by using the denoising operator $\Gamma(\cdot)$ and combining the multi-channel fusion strategy. The details are improved compared with the traditional G matrix method.

In contrast, the GAIN method directly learns the process from the antenna coordinates to the G matrix through the sensor array to deduce the subnet. At the same time, through the continuous feedback optimization of the entire network, a

high-precision imaging sensor matrix is constructed, and the background noise is effectively filtered by means of the multi-channel fusion imaging subnet, and finally a more refined target information reconstruction is achieved. In Figs. 6(I-c), (II-c), (III-c), (IV-c), (I-f), (II-f), (III-f), (IV-f), the head of the aircraft and tower tip in the red frame are more finely restored, and even the tail part of the (IV-c) green frame and the radar area in the (II-f) green frame can be effectively restored. It can be seen that the GAIN method can capture more target information and make the contour and details of the reconstructed image clearer.

From the results shown in Fig. 6, it can be seen that our GAIN method can restore the image details more accurately, and the imaging accuracy is significantly better than the G matrix method and MFIN method. High-quality imaging results can be obtained under various array structures.

In order to evaluate the quality of the reconstructed image more accurately, we calculate two commonly used objective evaluation indicators for the reconstructed image in Fig. 6: peak signal-to-noise ratio (PSNR) and root mean square error (RMSE). The higher the PSNR value is, the smaller the difference is between the reconstructed image and original image, that is, the higher the similarity is between the two, the better the performance of the network is. RMSE is used to quantify the average error between the predicted value and true value, which reflects the accuracy of the model prediction. The lower the RMSE value is, the smaller the difference is between the predicted image and target image, and the higher the quality of image restoration is.

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}_T^2}{\frac{1}{\mathrm{HW}}\sum_{i=0}^{H-1}\sum_{j=0}^{W-1}\left[T(i,j)-T_0(i,j)\right]^2}\right) \tag{16}$$

where T and T_0 represent the original bright temperature image and predicted image.

Table 2 and Table 3 list the imaging quality comparison results of traditional G matrix method, MFIN method, and GAIN method under different antenna arrays in aircraft scene and tower scene. The results show that the GAIN method significantly improves PSNR and reduces RMSE under all array conditions, showing its significant advantages in imaging quality. This is demonstrated in our experimental results, which fully verifies the effectiveness and versatility of our proposed algorithm.

TABLE 2. Comparison of image quality of three methods under different antenna arrays (Aircraft scene).

Array	PSNR			RMSE		
	G	MFIN	GAIN	G	MFIN	GAIN
Y	18.76	22.63	23.62	29.42	18.79	16.86
U	18.24	23.12	24.61	31.24	17.47	15.61
R_I	19.15	23.26	24.70	28.13	17.52	14.84
R_{II}	18.21	22.53	24.07	31.33	18.62	15.59

In the imaging process, the inversion time is a crucial consideration. Although in the early stage, the preparation of the data set and training steps often take a lot of time, once the network training is completed, the image can be generated only by loading the network model weight in the inversion stage, so the imaging inversion speed of GAIN is extremely fast. In contrast to the traditional G matrix method, in the inversion process, the G matrix must be constructed first, and then a large number of iterative calculations need to be carried out. Usually, thousands of iterations are required to obtain the results. This process is extremely time-consuming. Table 4 lists the time required for imaging inversion of the two methods in detail. It can be clearly observed that the G matrix method takes a long time, while the GAIN method takes a significantly shorter time. This result fully shows that the GAIN method shows excellent advantages in real-time performance and can meet the requirements of imaging speed in practical applications.

TABLE 3. Comparison of image quality of three methods under different antenna array (Tower scene).

Array	PSNR			RMSE		
	G	MFIN	GAIN	G	MFIN	GAIN
Y	11.84	19.17	21.53	63.67	28.05	19.05
U	12.28	21.20	23.07	62.23	22.21	16.14
R_I	12.59	21.39	23.13	62.19	22.04	16.13
R_{II}	12.14	20.27	22.94	63.00	22.71	18.39

TABLE 4. Image reconstruction time comparison.

Scene	G	GAIN
Aircraft/Tower	35s	4s

4. CONCLUSIONS

In order to achieve stable and high-precision SAIR imaging of any array structure, a general SAIR imaging network GAIN is proposed in this paper. The network solves the problem of low robustness and poor stability of traditional G matrix by sensor matrix deduction subnet and can deduce high-precision sensor matrix. At the same time, multi-channel fusion imaging subnet is used to effectively filter out noise and obtain high-precision reconstructed image. The simulation results show that GAIN not only improves the imaging quality, but also shows good adaptability to various antenna array structures, which is suitable for SAIR imaging requirements of any array layout and for high-precision imaging inversion of any SAIR array structure.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China under Grant No. 62371255 and No. 61601237, in part by the China Postdoctoral Foundation under Grant No. 2017M621585, in part by the Millimeter-wave State Key Laboratory Open project grant No. K202526.

REFERENCES

[1] Martàn-Neira, M., D. M. LeVine, Y. Kerr, N. Skou, M. Peichl, A. Camps, I. Corbella, M. Hallikainen, J. Font, J. Wu, S. Mecklenburg, and M. Drusch, "Microwave interferometric radiometry in remote sensing: An invited historical review," *Radio Science*, Vol. 49, No. 6, 415–449, 2014.

[2] LeVine, D. M., "Synthetic aperture radiometer systems," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, No. 12, 2228–2236, Dec. 1999.

[3] Zhou, X., H. Sun, J. He, and X. Lu, "NUFFT-based iterative reconstruction algorithm for synthetic aperture imaging radiometers," *IEEE Geoscience and Remote Sensing Letters*, Vol. 6, No. 2, 273–276, 2009.

[4] Chen, L., W. Ma, Y. Wang, and H. Zhou, "Array factor forming with regularization for aperture synthesis radiometric imaging with an irregularly distributed array," *IEEE Geoscience and Remote Sensing Letters*, Vol. 17, No. 1, 97–101, 2020.

[5] Zhang, Y., X. Qi, Y. Jiang, W. Liao, and Y. Du, "Reconstruction algorithm for staggered synthetic aperture radar with modified second-order keystone transform," *Journal of Applied Remote Sensing*, Vol. 15, No. 2, 026511, 2021.

[6] Li, H., C. Xiao, J. Dong, W. Wang, Y. Li, L. Chen, and H. Dou, "Image reconstruction method of 1-D nonuniform aperture synthesis radiometers based on convolutional neural network," in *2024 International Conference on Microwave and Millimeter Wave Technology (ICMMT)*, Vol. 1, 1–3, Bejing, China, May 2024.

[7] Feng, L., Q. Li, K. Chen, Y. Li, X. Tong, X. Wang, H. Lu, and Y. Li, "The gridding method for image reconstruction of nonuniform aperture synthesis radiometers," *IEEE Geoscience and Remote Sensing Letters*, Vol. 12, No. 2, 274–278, 2015.

[8] Andersson, F., R. Moses, and F. Natterer, "Fast fourier methods for synthetic aperture radar imaging," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 48, No. 1, 215–229, 2012.

[9] Wu, Y., Y. Li, G. Song, H. Dou, D. Wen, P. Li, X. Yang, R. Lv, and H. Li, "A near-field imaging method based on the near-field distance for an aperture synthesis radiometer," *Remote Sensing*, Vol. 16, No. 5, 767, 2024.

[10] Zhang, Y., Y. Ren, W. Miao, Z. Lin, H. Gao, and S. Shi, "Microwave SAIR imaging approach based on deep convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 57, No. 12, 10 376–10 389, 2019.

[11] Xiao, C., X. Wang, H. Dou, H. Li, R. Lv, Y. Wu, G. Song, W. Wang, and R. Zhai, "Non-uniform synthetic aperture radiometer image reconstruction based on deep convolutional neural network," *Remote Sensing*, Vol. 14, No. 10, 2359, 2022.

[12] Faucheron, R., E. Anterrieu, L. Yu, A. Khazaal, and N. J. Rodríguez-Fernández, "Deep learning based approach in imaging radiometry by aperture synthesis: An alias-free method," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 17, 6693–6711, 2024.

[13] Dou, H., C. Xiao, H. Li, Y. Li, P. Dang, R. Lv, P. Li, G. Song, Y. Wu, X. Yang, *et al.*, "Deep learning imaging for 1-D aperture synthesis radiometers," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 61, 1–16, 2023.

[14] Xiao, C., H. Dou, H. Li, R. Jin, R. Zhai, W. Wang, R. Lv, and Y. Li, "Image reconstruction of synthetic aperture radiometer by transformer," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 61, 1–15, 2023.

[15] Corbella, I., N. Duffo, M. Vall-llossera, A. Camps, and F. Torres, "The visibility function in interferometric aperture synthesis radiometry," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42, No. 8, 1677–1682, 2004.

[16] Chen, J., Y. Li, J. Wang, Y. Li, and Y. Zhang, "Regularization imaging algorithm with accurate G matrix for near-field MMW synthetic aperture imaging radiometer," *Progress In Electromagnetics Research B*, Vol. 58, 193–203, 2014.

[17] Wu, Y., Y. Li, G. Song, H. Dou, D. Wen, P. Li, X. Yang, R. Lv, and H. Li, "A near-field imaging method based on the near-field distance for an aperture synthesis radiometer," *Remote Sensing*, Vol. 16, No. 5, 767, 2024.

[18] Feng, X., X. Li, and J. Li, "Multi-scale fractal residual network for image super-resolution," *Applied Intelligence*, Vol. 51, No. 4, 1845–1856, 2021.

[19] Ma, Y., Y. Xu, Y. Liu, F. Yan, Q. Zhang, Q. Li, and Q. Liu, "Multi-scale cross-attention fusion network based on image super-resolution," *Applied Sciences*, Vol. 14, No. 6, 2634, 2024.

[20] Xiang, J., Y. Dong, and Y. Yang, "FISTA-Net: Learning a fast iterative shrinkage thresholding network for inverse problems in imaging," *IEEE Transactions on Medical Imaging*, Vol. 40, No. 5, 1329–1339, 2021.

[21] Tang, H., W. Zhang, H. Zhu, and K. Zhao, "Self-supervised real-world image denoising based on multi-scale feature enhancement and attention fusion," *IEEE Access*, Vol. 12, 49 720–49 734, 2024.

[22] Cao, R., L. Fang, T. Lu, and N. He, "Self-attention-based deep feature fusion for remote sensing scene classification," *IEEE Geoscience and Remote Sensing Letters*, Vol. 18, No. 1, 43–47, 2021.

[23] Chen, Y., R. Xia, K. Yang, and K. Zou, "MFFN: Image super-resolution via multi-level features fusion network," *The Visual Computer*, Vol. 40, No. 2, 489–504, 2024.

[24] Zhang, Y., Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vol. 43, No. 7, 2480–2495, Salt Lake City, UT, USA, Jun. 2018.

[25] Liu, Y., Z. Shao, and N. Hoffmann, "Global attention mechanism: Retain information to enhance channel-spatial interactions," *arXiv preprint arXiv:2112.05561*, 2021.