Progress In Electromagnetics Research C, Vol. 52, 39–52, 2014

A Multi-Feature Visibility Processing Algorithm for Radio Interferometric Imaging on Next-Generation Telescopes

Mu-Min Chiou[1], Jean-Fu Kiang[1, *], and Raj Mittra[2]

Abstract—The visibility distribution, which is related to the configuration of stations, can be categorized into different features, each having different levels of data number density. A computationally efficient multi-feature image reconstruction algorithm, well adapted for next-generation telescopes, is proposed based on this observation, which is more flexible to handle massive amount of visibility data expected in the future. In reconstructing the M87 image with the visibility data simulated on the Low-Frequency Array (LOFAR), this algorithm turns out to be a few hundreds to one thousand times faster and is more resilient to noises than the conventional algorithms.

1. INTRODUCTION

Radio interferometry techniques have been used to reconstruct images of astrophysical objects using incomplete and noisy visibility data. By augmenting more stations into an existing telescope, higher angular resolution and sensitivity can be achieved [1, 2].

In general, the image reconstruction process with radio interferometry can be formulated as a linear inverse problem, in which the available measurement data are used to reconstruct the source image, subject to certain constraints [3–6]. The measured visibility data, $\bar{v}^{\rm mea}$, are related to the image, $\bar{x}$, via a measurement equation, $\bar{v}^{\rm mea} = \bar{\bar{\Phi}} \cdot \bar{x} + \Delta\bar{v}^{\rm mea}$, where $\bar{\bar{\Phi}}$ is the measurement matrix, and $\Delta\bar{v}^{\rm mea}$ is the measurement error [7, 8]. The inverse problem is to reconstruct $\bar{x}$ from $\bar{v}^{\rm mea}$, usually accompanied by certain regularization.

The CLEAN algorithm has been widely used in this area [9], which is a variant of the steepest descent algorithms to minimize an object function, for example, $\|\bar{v}^{\rm mea} - \bar{\bar{\Phi}} \cdot \bar{x}\|_2^2$, under a regularization [6]. Each iteration of the CLEAN algorithm can be divided into a major cycle and a minor cycle. The major cycle estimates the increments of image, and the minor cycle modifies these increments before adding them to the previous version of reconstructed image.

The CLEAN algorithm works well for multiple point-like sources, but has a poor convergence rate when processing an extended source. A multi-scale CLEAN (MS-CLEAN) algorithm improves over its predecessor by processing different scales simultaneously [10], which speeds up the convergence rate when processing an extended source.

Several compressive sensing (CS) algorithms have been proposed for radio-interferometric imaging to leverage the sparsity regularization [11–13]. The image is assumed to be sparse when expressed in an orthonormal basis or a redundant dictionary (overcomplete dictionary), $\bar{\bar{\Psi}}$, as $\bar{x} = \bar{\bar{\Psi}} \cdot \bar{\alpha}$, where $\bar{\alpha}$ is called the decomposition of the signal.

If the measurement matrix has the restricted isometry property (RIP), the image can be recovered by solving a convex problem, $\min \|\bar{\alpha}\|_1$ subject to $\|\bar{v}^{\rm mea} - \bar{\bar{\Phi}} \cdot \bar{\bar{\Psi}} \cdot \bar{\alpha}\|_2 \le \varepsilon$, where ε is an upper bound on the L_2 norm of the noise, and $\|\bar{\alpha}\|_1$ is the L_1 norm of $\bar{\alpha}$. In [12], a sparsity averaging reweighted

Received 15 May 2014, Accepted 13 July 2014, Scheduled 15 July 2014

* Corresponding author: Jean-Fu Kiang (jfkiang@ntu.edu.tw).

[1] Department of Electrical Engineering and the Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan. [2] EMC Laboratory, the Pennsylvania State University, University Park, PA 16802, USA.

analysis (SARA) has been proposed, which adopts a reweighted L_1 minimization scheme and redundant dictionaries to further enhance the image fidelity.

CS-based techniques are claimed to have better image fidelity, flexibility and computation speed over the conventional methods. A CS-based technique starts from an idealized random discrete visibility coverage or variable density sampling patterns, which mimic generic sampling patterns of interferometric measurements.

The CS technique has been extended to process continuous visibility distributions with high scalability and speed enhancement, in which the simultaneous-direction method of multipliers (SDMM) is applied to solve sparse imaging problems [14]. The SDMM was developed to deal with non-differentiable object functions. The CS theory assumes a prior that the image is sparse when expressed in a chosen dictionary, and small amount of measurement data inherently, which is not the case for next-generation telescopes [11].

Next-generation radio telescopes, such as Square Kilometre Array Pathfinder (ASKAP) [15], Murchinson Widefield Array (MWA) [16], Meerkat [17], Square kilometre Array (SKA) [18, 19], and Low-Frequency Array (LOFAR) [20], are expected to achieve a wider dynamic range and higher angular resolution than current instruments.

These telescopes will acquire a massive amount of data. Taking LOFAR for example, there are over forty telescope stations and the number is likely to increase [20]. If the sampling time of 2 s with the high-band antennas (HBA's) is taken, over 15,000,000 visibility data will be collected within 6 hours of operation [21, 22]. The number of measurement data will be increased further if the visibility data are taken at multiple frequencies [6].

As a comparison, the amount of visibility data processed with conventional methods is relatively small. For example, 11,469 data were simulated with a probabilistic image algorithm [23], 505,440 data points were simulated with the MS-CLEAN algorithm [24], and 131,072 data points were simulated with the PURIFY algorithm [14]. More efficient imaging techniques are required to process the vast amount of measurement data collected by using the next-generation instruments for radio interferometry.

A regridding technique has been proposed to process the visibility data [25], which takes much less computational time than the MS-CLEAN algorithm. The regridded visibility data are obtained by weighting the measured visibility data first, then applying the convolution of a sampling function to the weighted visibility data. Both processes involve the choice of empirical parameters, which may lead to different images. Moreover, the regridded data are correlated and result in a poor signal-to-noise ratio (SNR) in the recovered image [26].

In this work, a multi-feature algorithm is proposed to process four different features of the visibility data separately, with the data simulated on the LOFAR configuration. This method is driven by the measurement data, without the need of a prior function for regularization. It takes much shorter computational time and is more resilient to noises than the conventional methods.

This paper is organized as follows. The reconstruction theory and the configuration of stations in next-generation telescopes are briefly reviewed in Section 2, the multi-feature visibility distribution of the M87 image [27] simulated on the LOFAR configuration is presented in Section 3, the multi-feature algorithm is presented in Section 4, followed by simulations and discussions in Section 5. Finally, some conclusions are drawn in Section 6.

2. RECONSTRUCTION THEORY AND CONFIGURATION OF STATIONS

Figure 1 shows the coordinate systems for the image and the visibility distributions, where $\hat{\nu}_u$ and $\hat{\nu}_v$ are the unit vectors parallel to the geographic east-west and north-south directions, respectively; $\hat{\theta}_u$ and $\hat{\theta}_v$ are the unit vectors parallel to $\hat{\nu}_u$ and $\hat{\nu}_v$, respectively. The Earth-centered Earth-fixed coordinates, (x, y, z), are used to describe the station positions in a telescope.

At a small field of view (FOV), the image distribution, $x(\theta_u, \theta_v)$, and the visibility distribution, $v(\nu_u, \nu_v)$, form a Fourier transform pair as [1]

$$x(\theta_u, \theta_v) = \iint v(\nu_u, \nu_v) e^{-j2\pi(\nu_u\theta_u + \nu_v\theta_v)} d\nu_u d\nu_v \tag{1}$$

$$v(\nu_u, \nu_v) = \iint x(\theta_u, \theta_v) e^{j2\pi(\nu_u\theta_u + \nu_v\theta_v)} d\theta_u d\theta_v \tag{2}$$

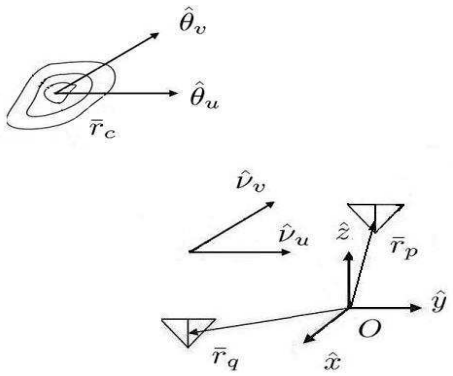


Figure 1. Coordinate systems for image and visibility distributions, the Earth-centered Earth-fixed coordinate is used to describe station positions.

where λ is the operating wavelength; θ_u and θ_v are the angular position, measured from the center, $\bar{r}_c$. Similarly, ν_u and ν_v are the components of a baseline, measured in wavelength. The position difference between stations p and q is $\bar{r}_{p,q}$, and its projection onto the visibility plane is called a baseline, determined as $\bar{B} = \hat{\nu}_u B_u + \hat{\nu}_v B_v = (\hat{\nu}_u \hat{\nu}_u + \hat{\nu}_v \hat{\nu}_v) \cdot \bar{r}_{p,q}$.

2.1. Review of LOFAR Configuration

Figure 2 shows the distribution of LOFAR stations, including low-band antennas (LBA's) operating in the 10–90 MHz band and high-band antennas (HBA's) operating in the 110–240 MHz band. The shortest baseline is about 70 m [20–22]. There are 24 core stations (CS's), located within a core area of 2 km in radius. At the center of the core area, a Superterp consisting of six stations resides within an area of 320 m in diameter. Outside the core area, 16 remote stations (RS's) are located in Netherlands, and 8 international stations are distributed over Germany, France, the United Kingdom and Sweden [16]. In this work, only core and remote stations are considered in the simulations without loss of generality.

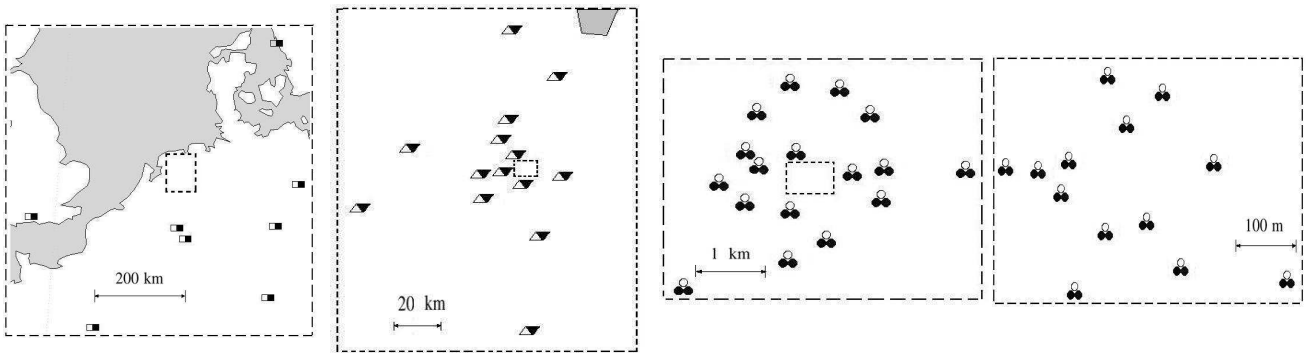


Figure 2. Distribution of LOFAR stations [20]. LBA's and HBA's are marked by hollow and solid symbols, respectively; core, remote and international stations are marked by circles, triangles and squares, respectively; inlets (b), (c) and (d) are the magnified picture of the dashed rectangular regions in inlets (a), (b) and (c), respectively.

2.2. Brief Review of Other Next-generation Telescopes

In the SKA configuration, 50% of the data are collected within the central area with radius of 2.5 km, another 25% are collected within 180 km off the center, and the other 25% are attributed to the other stations [19].

In the Meerkat, a large fraction of data are collected within the central area with 700 m in diameter, to reconstruct images of wide angular extension with a wide dynamic range, and wide-field transients can be detected at a lower data rate [17]. Some stations are placed 10 km off the center to acquire higher image resolution and to locate point-like sources.

In the MWA, 50 tiles are uniformly distributed in the core area with 100 m in diameter, surrounded by 62 tiles distributed over an area with 1.5 km in diameter [16]. Another 16 tiles are placed as far as 3 km off the center to achieve the highest angular resolution of images [16].

In the ASKAP, three station configurations are available, a very compact configuration with maximum baseline of 400 m, a medium compact configuration with maximum baseline of 2 km, and an extended configuration with maximum baseline of 8 km [15].

The next-generation radio telescopes will face the challenges of having to deal with tremendous volume of data over a wide high dynamic range and wide bandwidth, many different radio interferences, as well as serious instrumentation and analysis difficulties [23].

3. MULTI-FEATURE VISIBILITY DISTRIBUTION

Figure 3 shows the image intensity of M87 with 256×256 pixels, which is located at the declination of $12°23'28''$ [27]. The image is taken at 140 MHz, with the resolutions of $\Delta\theta_u = 3.4''$ and $\Delta\theta_v = 3.1''$. The image intensity is normalized as

$$\sum_{n=1}^{N} x_n = 1 \tag{3}$$

where N is the total number of image pixels.

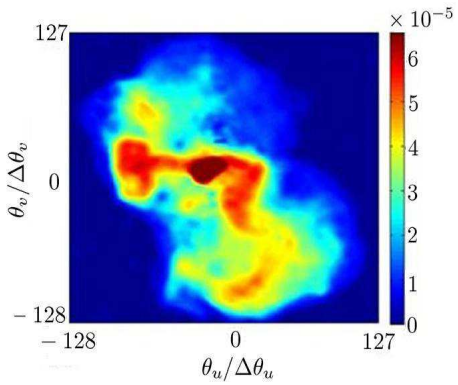


Figure 3. LOFAR image of M87 near the center of Virgo cluster [27]. The resolution is $\Delta\theta_u = 3.4''$ and $\Delta\theta_v = 3.1''$.

Figure 4 shows the visibility distributions simulated at different declinations of 90°, 65°, 40° and 10°, respectively. The sampling interval is 60 seconds, and the total observation time is 6 hours. The visibility data collected from the CS-CS links concentrate around the origin, which bear low-resolution information of the image. The visibility data collected from the CS-RS links usually take the shape of ribbons, and those from the RS-RS links take the string shape. Both bear high-resolution information.

Figure 5(a) displays the number density of simulated visibility data when observing the M87 at the declination of $12°23'28''$, and Fig. 5(b) displays a zoom-in near the central region. The number density is defined as the number of visibility data within a visibility cell. Each visibility cell has the size of $\Delta\nu_{u,\mathrm{def}} \times \Delta\nu_{v,\mathrm{def}}$, which are related to the image resolution as

$$\Delta\nu_{u,\mathrm{def}} = \frac{1}{N_g \Delta\theta_u}, \qquad \Delta\nu_{v,\mathrm{def}} = \frac{1}{N_g \Delta\theta_v}$$

where $N_g = \sqrt{N}$ is number of pixels in one row or one column of the image.

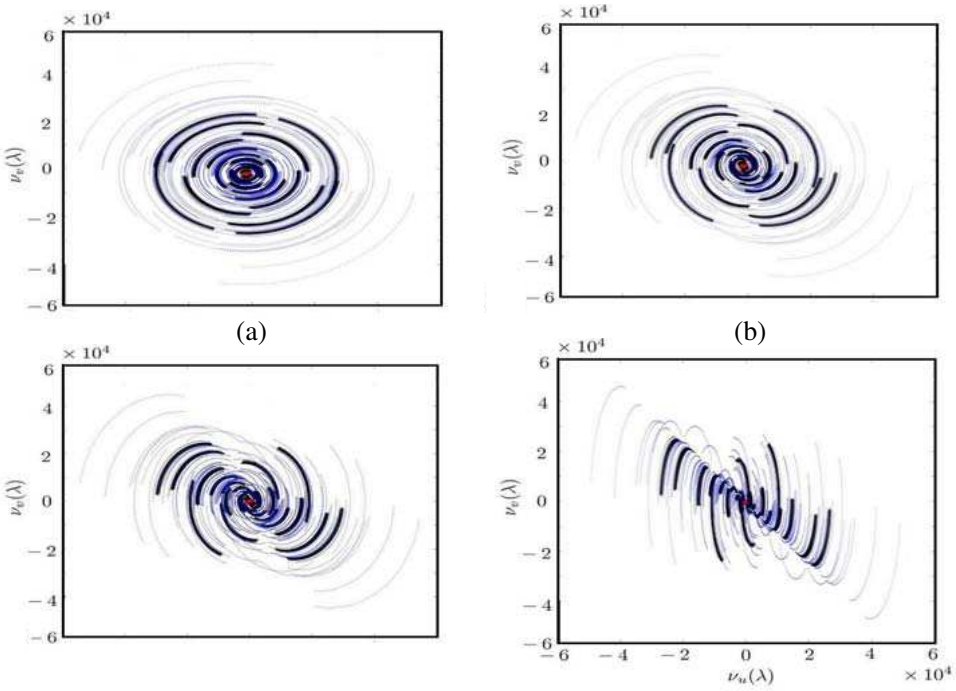


Figure 4. Visibility distribution at declination of (a) 90°, (b) 65°, (c) 40° and (d) 10°; red: data from CS-CS links, black: data from CS-RS links, blue: data from RS-RS links.

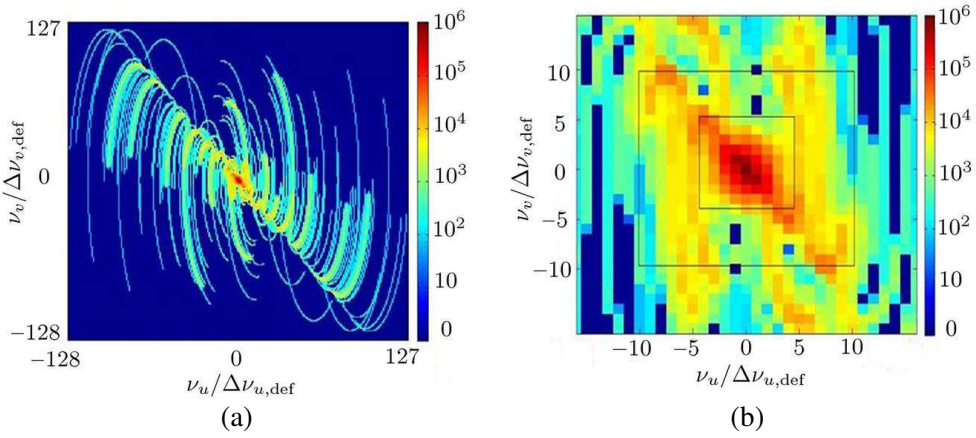


Figure 5. Number density of visibility data in the visibility plane, (a) whole region of interest, (b) zoom in of the central region.

The number density varies drastically near the origin. The most crowded region is roughly in the inner patch region, $[-W_u, W_u] \times [-W_v, W_v]$, with

$$W_u = \max_{B_u \in \text{CS-CS links}} |B_u|/\lambda, \qquad W_v = \max_{B_v \in \text{CS-CS links}} |B_v|/\lambda$$

The next crowded region lies in the outer patch, $[-2W_u, 2W_u] \times [-2W_v, 2W_v]$ excluding the inner patch. The number density in the inner patch is on the order of $10^5 \sim 10^6$, while that in the outer patch is on the order of $10^2 \sim 10^4$. Several ribbon regions are observed in Fig. 5(b), with their width near W_u or W_v. The rest of the visibility data take the string form.

4. MULTI-FEATURE ALGORITHM

The multi-feature algorithm processes the four different types of feature separately and reconstruct the image as

$$\bar{x} = \bar{\bar{F}}^{-1} \cdot \bar{v}_i + \bar{\bar{F}}^{-1} \cdot \bar{v}_o + \bar{\bar{F}}^{-1} \cdot \bar{v}_r + \bar{\bar{F}}^{-1} \cdot \bar{v}_s \tag{4}$$

where $\bar{\bar{F}}^{-1}$ stands for the inverse Fourier transform, $\bar{v}_i$, $\bar{v}_o$, $\bar{v}_r$ and $\bar{v}_s$ represent the visibility data in the inner patch, outer patch, ribbons and strings, respectively.

By observation, the data associated with a specific remote station form a ribbon. Some of the visibility data associated with the CS-RS links fall in the inner patch or the outer patch, and are counted into the inner patch or the outer patch. Some data fall in the intersection of two or more ribbons, and are counted into only one of them. Most of the visibility data associated with the RS-RS links take the string form. The part of a string falling in a patch or a ribbon is counted into the patch or the ribbon.

The multi-feature algorithm is a data-driven process. It can be scaled up to include additional stations, like the international stations in the LOFAR. The multi-feature algorithm can also be adapted for parallel computation, which is more flexible to deal with the vast amount of data the next-generation telescopes will collect.

4.1. Grid Model of Visibility

In each visibility region, a grid with an optimized cell size is chosen to derive a set of modeled visibility data, $v[m_u, m_v] = v(m_u\Delta\nu_u, m_v\Delta\nu_v)$, which are stored in a one-dimension vector, $v[m]$. The modeled visibility data on the grid are obtained as

$$\bar{v}^{\text{opt}} = \arg\min_{\bar{v}\in\Omega} f_{\text{data}}(\bar{v}) \tag{5}$$

where Ω contains all the grid points in the subject region, each having at least one measurement data within its four surrounding cells,

$$f_{\text{data}}(\bar{v}) = \sum_{n_d} w_{n_d} \left| v_{n_d}^{\text{mea}} - v_{n_d}^{\text{mod}} \right|^2 \tag{6}$$

is the object function defined to determine $v_{n_d}^{\text{mod}}$'s, the modeled visibility data at $\bar{\nu}_{n_d}$'s. The weighting coefficients, w_{n_d}'s, are estimated as

$$w_{n_d} = \text{var}\left\{\text{Re}\left\{v_{n_d}^{\text{mea}}\right\}\right\}^{-1} = \text{var}\left\{\text{Im}\left\{v_{n_d}^{\text{mea}}\right\}\right\}^{-1} \tag{7}$$

where var $\{\alpha\}$ stands for the variance of the random variable α. Eq. (5) can be solved using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [28], which is a variant of the quasi-Newton's method.

As shown in Fig. 6, the modeled visibility data corresponding to a measurement data is related to the visibility data at its four surrounding grid points via bilinear interpolation as

$$v_{n'}^{\text{mod}} = \sum_{i=1}^{4} R_{n',m_i'} v[m_i'] \tag{8}$$

where all the elements in the n'th row of $\bar{\bar{R}}$ are reset to zero, except

$$R_{n',m_1'} = (1-\xi_v')(1-\xi_u'), \qquad R_{n',m_2'} = \xi_u'(1-\xi_v'), \qquad R_{n',m_3'} = \xi_u'\xi_v', \qquad R_{n',m_4'} = \xi_v'(1-\xi_u')$$

m_1', m_2', m_3' and m_4' are the one-dimensional indices of (m_u', m_v'), $(m_u'+1, m_v')$, $(m_u'+1, m_v'+1)$ and $(m_u', m_v'+1)$, respectively, with

$$m_u' = \left\lfloor \frac{\nu_u'}{\Delta\nu_u} \right\rfloor, \qquad m_v' = \left\lfloor \frac{\nu_v'}{\Delta\nu_v} \right\rfloor, \qquad \xi_u' = \frac{\nu_u' - m_u'\Delta\nu_u}{\Delta\nu_u}, \qquad \xi_v' = \frac{\nu_v' - m_v'\Delta\nu_v}{\Delta\nu_v}$$

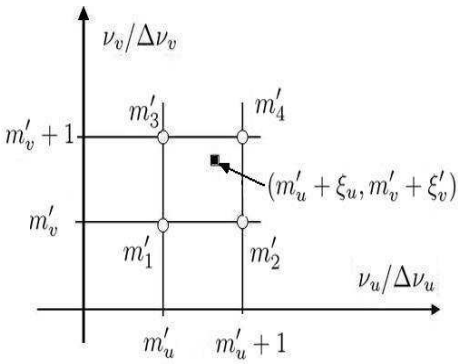


Figure 6. Modeled visibility data at the baseline, $\bar{\nu}_{n'} = (\nu'_u, \nu'_v)$, to be interpolated from four surrounding visibility data, $v[m'_1]$, $v[m'_2]$, $v[m'_3]$ and $v[m'_4]$.

In the CLEAN and CS-based algorithms, the modeled visibility data in each iteration are derived by transforming the most updated image via $v_{n'}^{\text{mod}} = \sum_{n=1}^{N} \Phi_{n',n} x_n$. As a comparison, applying (8) to calculate one modeled visibility data takes only four multiplications. In addition, the real value of image intensity implies $v(-\nu_u, -\nu_v) = v^*(\nu_u, \nu_v)$, which can be used to reduce the memory size by half.

Each type of visibility feature is processed separately as follows, which can be implemented in parallel.

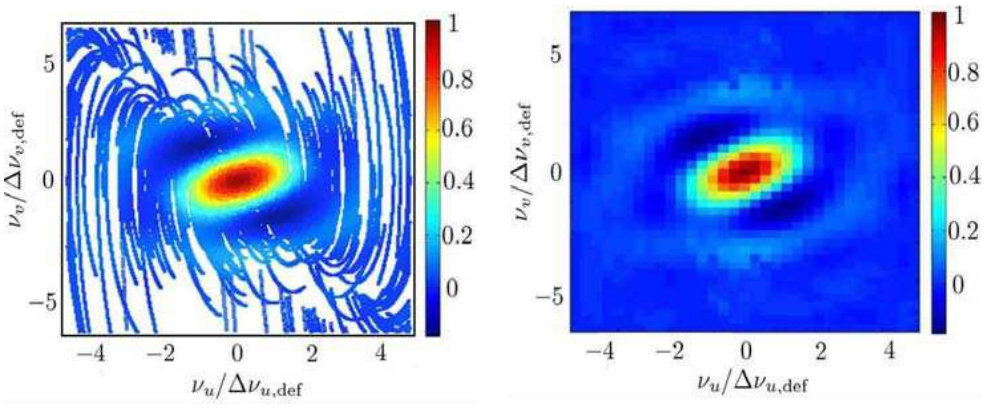
4.2. Inner Patch

Figure 7 shows the real part of the measured and the modeled visibility data, respectively, in the inner patch. The visibility resolutions, $\Delta\nu_u$ and $\Delta\nu_v$, are related to the image resolutions, $\Delta\theta_u$ and $\Delta\theta_v$, as

$$\Delta\nu_v = \Delta\nu_u \frac{\Delta\theta_u}{\Delta\theta_v}$$

To choose the optimum resolution in a given visibility region, a mean-squared error (MSE) is defined as

$$\text{MSE} = \frac{1}{N} \left\| \bar{x} - \bar{x}^{\text{ref}} \right\|_2^2 \tag{9}$$



(a) (b)

Figure 7. Real part of visibility data in the inner-patch region, (a) measured and (b) modeled, $\Delta\nu_u = 54\lambda$, $\Delta\nu_v = 59.225\lambda$.

where $\bar{x}^{\text{ref}}$ is the reference image (true image without noise), and $\|\bar{x}\|_2$ denotes the L_2 norm of the image, $\bar{x}$.

Figure 8 shows the MSE_i defined in (9), with $\bar{x} = \bar{x}_i$, derived from the visibility data in the inner patch. Choosing $\Delta\nu_{u,i} = 65\lambda$ ($\Delta\nu_{v,i} = 71.29\lambda$) results in the lowest MSE_i. The optimal reconstructed image using only the inner-patch visibility data is obtained as

$$\bar{x}_i^{\text{opt}} = \bar{\bar{F}}^{-1} \cdot \bar{v}_i$$

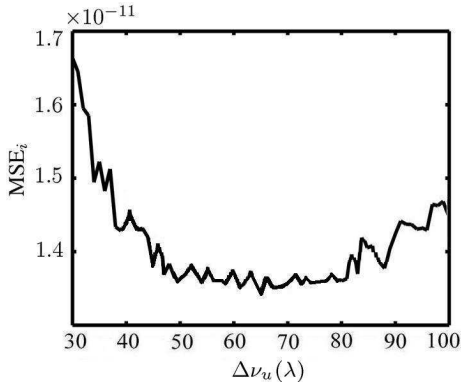


Figure 8. MSE_i as a function of $\Delta\nu_u(\lambda)$ in the inner patch.

4.3. Outer Patch

Figure 9 shows the real part of the measured and the modeled visibility distributions, respectively, in the outer patch. Fig. 10 shows the MSE_o with different $\Delta\nu_u$'s, where MSE_o is defined in (9), with $\bar{x} = \bar{x}_o + \bar{x}_i^{\text{opt}}$, and $\bar{x}_o$ is the reconstructed image using only the visibility data in the outer patch. Choosing $\Delta\nu_{u,o} = 142\lambda$ ($\Delta\nu_{v,i} = 155.742\lambda$) results in the lowest MSE_o. The optimal reconstructed image using only the outer-patch visibility data is obtained as

$$\bar{x}_o^{\text{opt}} = \bar{\bar{F}}^{-1} \cdot \bar{v}_o$$

4.4. Ribbons

Figure 11 shows the configuration of 14 ribbons in the positive-ν_v half-plane, which are contributed by the CS-RS links, and the symmetry property.

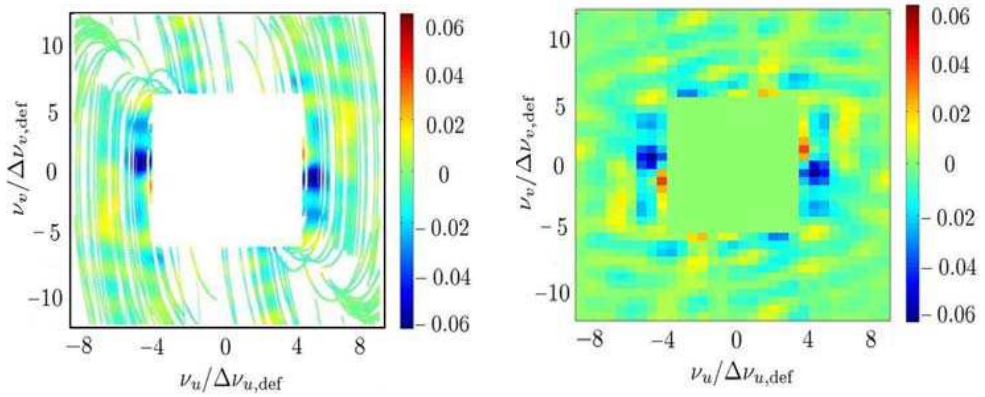


Figure 9. Real part of visibility data in the outer patch, (a) measured and (b) modeled, $\Delta\nu_u = 142\lambda$, $\Delta\nu_v = 155.742\lambda$.

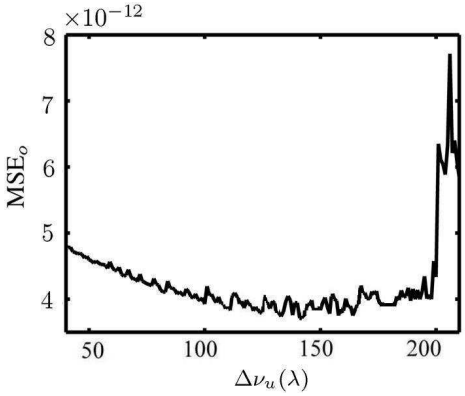


Figure 10. MSE_o as a function of $\Delta\nu_u(\lambda)$ in the outer patch.

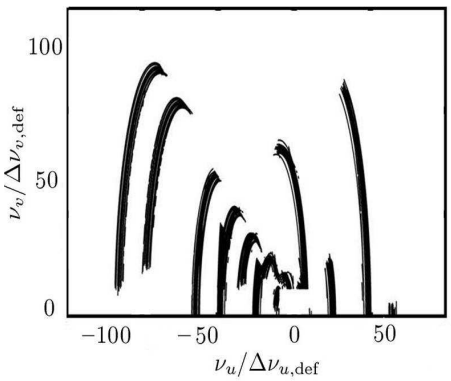


Figure 11. Distributions of 14 ribbons in the positive ν_v half-plane.

The image component reconstructed with these ribbons can be calculated as

$$\begin{aligned} x[n_u, n_v] &= \frac{u_{\mathrm{sc}} v_{\mathrm{sc}}}{N_g^2} \sum_{m_u=1}^{M_u} \sum_{m_v=1}^{M_v} v[m_u, m_v] e^{-j2\pi(n_u m_u u_{\mathrm{sc}} + n_v m_v v_{\mathrm{sc}})/N_g} \\ &\quad + \frac{u_{\mathrm{sc}} v_{\mathrm{sc}}}{N_g^2} \sum_{m_u=1}^{M_u} \sum_{m_v=1}^{M_v} v[-m_u, -m_v] e^{-j2\pi(-n_u m_u u_{\mathrm{sc}} - n_v m_v v_{\mathrm{sc}})/N_g} \\ &= 2\frac{u_{\mathrm{sc}} v_{\mathrm{sc}}}{N_g^2} \mathrm{Re} \sum_{n_u=1}^{N_u} \sum_{m_v=1}^{N_v} v[m_u, m_v] e^{-j2\pi(n_u m_u u_{\mathrm{sc}} + n_v m_v v_{\mathrm{sc}})/N_g} \end{aligned}$$

where M_u and M_v are the number of pixels in the ν_u and ν_v direction, respectively;

$$u_{\mathrm{sc}} = \frac{\Delta\nu_u}{\Delta\nu_{u,\mathrm{def}}}, \qquad v_{\mathrm{sc}} = \frac{\Delta\nu_v}{\Delta\nu_{v,\mathrm{def}}}$$

are the scaling factors in the ν_u and ν_v direction, respectively.

Figure 12 shows the real part of the measured and the modeled visibility distributions, respectively, in one of the ribbons. Fig. 13 shows the MSE_r as a function of $\Delta\nu_u$, where MSE_r is defined as in (9), with $\bar{x} = \bar{x}_r + \bar{x}_i^{\mathrm{opt}} + \bar{x}_o^{\mathrm{opt}}$, and $\bar{x}_r$ is the image component reconstructed with only the visibility data in the ribbons. Choosing $\Delta\nu_{u,r} = 112\lambda$ ($\Delta\nu_{v,r} = 122.8387\lambda$) results in the lowest MSE_r. The optimal reconstructed image using only the ribbon visibility data is obtained as

$$\bar{x}_r^{\mathrm{opt}} = \bar{\bar{F}}^{-1} \cdot \bar{v}_r$$

4.5. Strings

Figure 14 shows the real part of the measured and the modeled visibility distributions, respectively, in one of the strings. Fig. 15 shows the MSE_s as a function of $\Delta\nu_u$, where MSE_s is defined as in (9), with $\bar{x} = \bar{x}_s + \bar{x}_i^{\mathrm{opt}} + \bar{x}_o^{\mathrm{opt}} + \bar{x}_r^{\mathrm{opt}}$, and $\bar{x}_s$ is the reconstructed image component with only the visibility data in the strings. Choosing $\Delta\nu_{u,s} = 119\lambda$ ($\Delta\nu_{v,r} = 130.516\lambda$) results in the lowest MSE_s. The optimal reconstructed image using only the string visibility data is

$$\bar{x}_s^{\mathrm{opt}} = \bar{\bar{F}}^{-1} \cdot \bar{v}_s$$

Finally, the reconstructed image becomes

$$\bar{x}^{\mathrm{opt}} = \bar{x}_i^{\mathrm{opt}} + \bar{x}_o^{\mathrm{opt}} + \bar{x}_r^{\mathrm{opt}} + \bar{x}_s^{\mathrm{opt}}$$

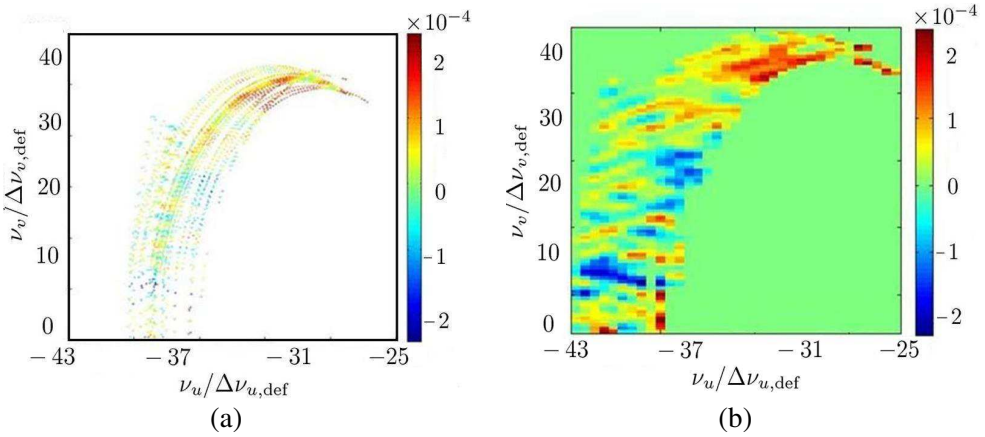


Figure 12. Real part of visibility data in one of the ribbon regions, (a) measured and (b) modeled, $\Delta\nu_u = 102\lambda$, $\Delta\nu_v = 111.87\lambda$.

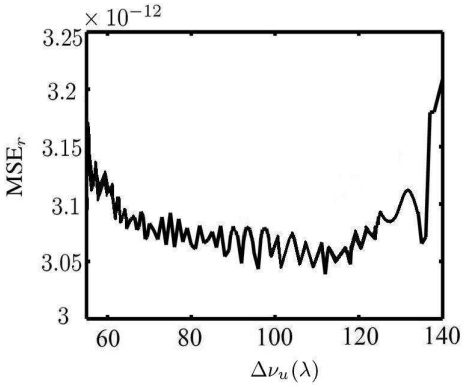


Figure 13. MSE_r as a function of $\Delta\nu_u(\lambda)$ in the ribbon regions.

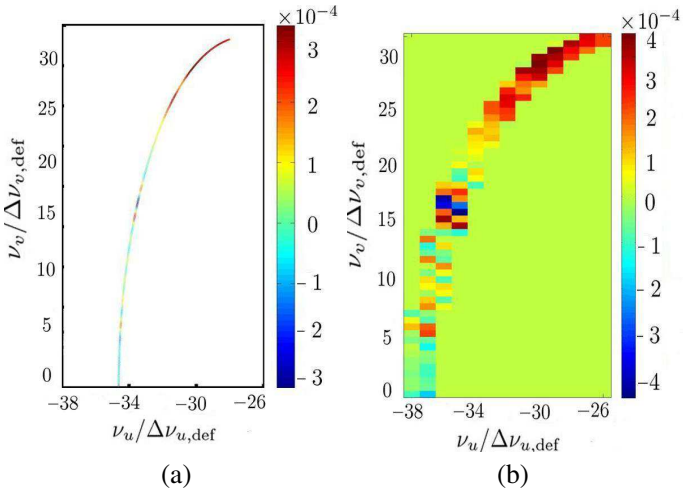


Figure 14. Real part of visibility data in one of the strings, (a) measured and (b) modeled, $\Delta\nu_u = 121$, $\Delta\nu_v = 132.7$.

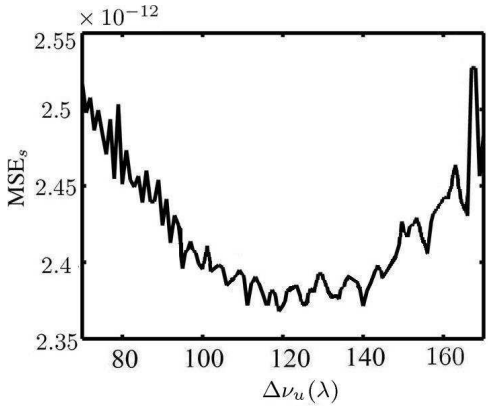


Figure 15. MSE_s as a function of $\Delta\nu_u(\lambda)$ in the string regions.

5. SIMULATIONS AND DISCUSSIONS

The effects of noise can be analyzed by defining an input signal-to-noise ratio (ISNR) [13]

$$\mathrm{ISNR} = 20\log_{10}\frac{\left\|\bar{v}^{\mathrm{ref}}\right\|_2}{\left\|\bar{v}^{\mathrm{mea}} - \bar{v}^{\mathrm{ref}}\right\|_2} \tag{10}$$

where $\bar{v}^{\mathrm{ref}}$ is the Fourier transform of the reference image without noise. In radio interferometry, the amplitude and phase of the measured visibility data can be described as [29]

$$\begin{aligned} |v_{n_d}^{\mathrm{mea}}| &= |v_{n_d}^{\mathrm{ref}}| + \Delta v_{\mathrm{amp},n_d} \\ \phi_{n_d}^{\mathrm{mea}} &= \arg\{v_{n_d}^{\mathrm{ref}}\} + \Delta\phi_{n_d} \end{aligned}$$

where $v_{n_d}^{\mathrm{ref}}$ is the n_dth reference visibility data without noise, and $\Delta v_{\mathrm{amp},n_d}$ and $\Delta\phi_{n_d}$ are the additive zero-mean Gaussian noises to amplitude and phase, respectively. In the simulations, the standard deviation of amplitude is set proportional to $\left|v_{n_d}^{\mathrm{ref}}\right|$.

The quality of reconstruction can be evaluated in terms of SNR, defined as [12]

$$\mathrm{SNR} = 20\log_{10}\frac{\left\|\bar{x}^{\mathrm{ref}}\right\|_2}{\left\|\bar{x}^{\mathrm{ref}} - \bar{x}\right\|_2} \tag{11}$$

When $\bar{x}$ is closer to $\bar{x}^{\mathrm{ref}}$, the SNR becomes higher.

Table 1. Performance of multi-feature algorithm.

visibility data	optimal $\Delta\nu_u(\lambda)$ of added feature	SNR (dB)
$\bar{v}_i$ only	65	15.884
add $\bar{v}_o$	142	21.468
add $\bar{v}_r$	112	22.331
add $\bar{v}_s$	119	23.414

Table 1 lists the optimal cell size and SNR performance in different visibility regions, based on simulations. Table 2 lists the feasible range of baseline in different visibility regions, corresponding to the optimal cell sizes in Table 1.

Table 2. Estimated cell sizes.

region	ΔB_u (m)	ΔB_v (m)	sampling interval (s)	total observation time (hr)
Inner patch	$107 \sim 171$	$117 \sim 187$	60	6
Outer patch	$214 \sim 321$	$234 \sim 352$	60	6
Ribbons	$150 \sim 278$	$164 \sim 305$	60	6
Strings	$223 \sim 300$	$245 \sim 330$	60	6

Figure 16 shows the images reconstructed with the multi-feature algorithm and the conventional MS-CLEAN algorithm, respectively [10]. The corresponding SNR values are 23.414 dB and 19.417 dB, respectively.

A desktop PC (ASUS AMD 750, Core i73.4 GHz, 32 Gbyte of RAM) is used to run the simulations. If 6 hours of measurement is considered, at the sampling interval of 60 seconds, there will be 504,720 visibility data points. The multi-feature algorithm takes 856 seconds of CPU time, and the MS-CLEAN algorithm takes 4,573 seconds.

If the sampling interval is reduced to 2 seconds, there will be 15,141,600 visibility data points. In this case, the multi-feature algorithm takes 25,144 seconds of CPU time. The SNR of the reconstructed image is 23.8 dB, which is slightly higher than that using 504,720 visibility data points.

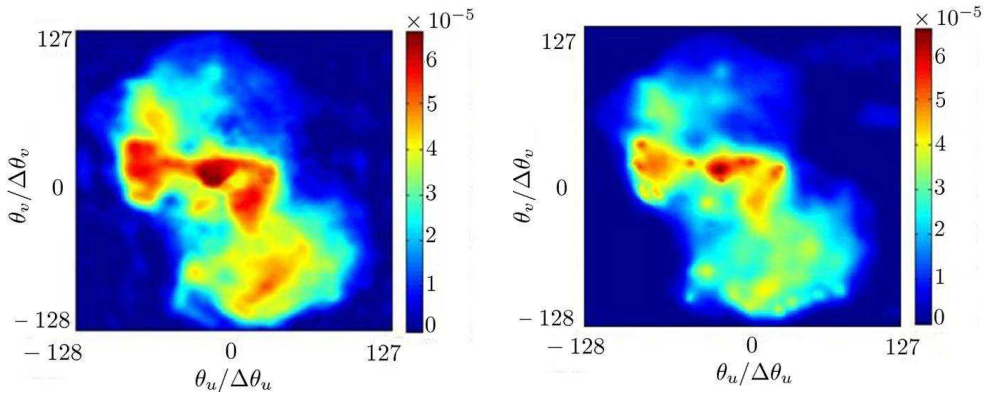


Figure 16. Reconstructed images, (a) with multi-feature algorithm and (b) with MS-CLEAN algorithm [10].

Table 3 lists the estimated numbers of multiplication/division operations needed in the MS-CLEAN, the PURIFY and the multi-feature algorithm, respectively; where M/D stands for multiplication/division and A/S stands for addition/subtraction; $T_{\rm MS}$, $T_{\rm P}$ and $T_{\rm MF}$ are the total number of iterations in MS-CLEAN, PURIFY and multi-feature algorithm, respectively; M is the total number of grid points in the visibility distribution used in the multi-feature algorithm; $T_{\rm CG}$ and $T_{\rm LS}$ are the number of iterations in the conjugate gradient method used in the PURIFY and that in the line search method used in the BFGS algorithm; $B_{\rm MF}$ is the number of surrounding grid points used in the interpolation, which is 4 in the multi-feature algorithm. By simulations, $T_{\rm MB}$ is larger than 2,000, $T_{\rm MF}$ is less than 250, and $T_{\rm LS}$ is less than 10. The computation of $\bar{\bar{\Phi}}^{\dagger}\cdot\bar{\bar{\Phi}}$ is dependent on the interpolation scheme and the measured visibility distribution, which is estimated as $N_v \times N$.

In this work, the parameters are chosen as $T_{\rm MS} = 10,000$, $N_v = 10,000,000$, $N = 256^2$, $T_{\rm CG} = 256$, $T_{\rm P} = 100$, $B_{\rm MF} = 4$, $T_{\rm MF} = 250$, $T_{\rm LS} = 10$, and $M = 20,000$. The numbers of M/D operations required in the MS-CLEAN, the PURIFY and the multi-feature algorithms are 4.426×10^{13}, 2.206×10^{14} and

Table 3. Number of multiplication/division operations.

MS-CLEAN [10]		PURIFY [14]		Multi-Feature	
Process	M/D	Process	M/D	Process	M/D
Dirty image	$N \times N_v$	$\bar{\bar{\Phi}}^{\dagger}\cdot\bar{\bar{\Phi}}$	$N_v \times N$	$\nabla_{\bar{v}} f$ (per iteration)	$B_{\rm MF}N_v + B_{\rm MF}N_v$
Dirty beam	$N \times N_v$	$\bar{\bar{\Phi}}\cdot\bar{x}$ (per iteration)	$N_v \log N$	Line search (per iteration)	$T_{\rm LS}(B_{\rm MF}+2)N_v$
Convolution (per iteration)	$N \times N$	Conjugate gradient (per iteration)	$T_{\rm CG}(2N^2+6N)$	$\bar{\bar{F}}^{-1}$	$M \times N$
Total	$T_{\rm MS}N^2+2N_vN$	Total	N_vN+ $T_{\rm P}T_{\rm CG}(2N^2+6N)$ $+T_{\rm P}N_v\log N$	Total	$2T_{\rm MF}N_vB_{\rm MF}$ $+T_{\rm MF}N_vT_{\rm LS}$ $(B_{\rm MF}+2)+MN$

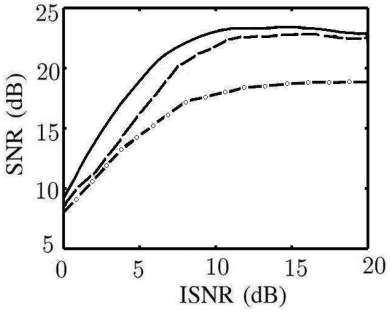


Figure 17. Average SNR as a function of ISNR over 100 simulations, ———: multi-feature algorithm on 15,141,600 visibility data, − − −: multi-feature algorithm on 504,720 visibility data, −∘−: MS-CLEAN algorithm [10] on 504,720 visibility data.

1.713×10^{11}, respectively; leading to the ratio of 258.4 : 1287.6 : 1. This ratio increases if the number of visibility measurement data or the image size is increased.

Figure 17 shows the average SNR of the reconstructed image, as a function of ISNR over 100 realizations of Monte-Carlo simulation. With 504,720 visibility data points, the multi-feature algorithm achieves higher SNR than the MS-CLEAN algorithm. With the visibility data points increased to 15,141,600, the SNR achieved with the multi-feature algorithm gets even higher. The SNR values with both the multi-feature algorithm and the MS-CLEAN algorithm saturate at high ISNR values, and the SNR value with the former is higher than that with the latter.

6. CONCLUSION

A multi-feature algorithm is proposed to reconstruct the visibility distribution instead of the image itself. The visibility data are categorized into four different feature regions, depending on the configuration of stations and the target to be observed. The cell sizes in different regions are optimized separately, and up to 15,000,000 visibility data points have been processed in the simulations, based on the LOFAR configuration. Compared with PURIFY and MS-CLEAN algorithms, the computational time can be reduced by a few hundreds to one thousand times, and expected to be even more if the amount of data is increased. The multi-feature algorithm renders better image fidelity and is more resilient to noises. This algorithm can be scaled up to accommodate vast amount of measurement data anticipated on the next-generation telescopes.

ACKNOWLEDGMENT

This work was sponsored by the National Science Council, Taiwan, under contract NSC 102-2221-E-002-043; and the Ministry of Education, Taiwan, under Aim for Top University Project 103R3401-1.

REFERENCES

1. Thompson, A. R., J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*, 2nd Edition, Wiley-VCH, 2004.
2. Baars, J. W. M., L. R. D'Addario, and A. R. Thompson, "Radio astronomy in the early twenty-first century," *Proc. IEEE*, Vol. 97, No. 8, 1377–1381, Aug. 2009.
3. Gull, S. F. and J. Skilling, "Maximum entropy method in image processing," *IEE Proc. Commun. Radar Signal Process.*, Vol. 131, No. 6, 646–659, 1984.
4. Molina, R., J. Nunez, F. J. Cortijo, and J. Mateos, "Image restoration in astronomy: A Bayesian perspective," *IEEE Signal Process. Mag.*, Vol. 18, No. 2, 11–29, 2001.
5. Chen, B.-D. and L. Amir, "Parametric high resolution techniques for radio astronomical imaging," *IEEE J. Select. Topics Signal Process.*, Vol. 2, No. 5, 670–684, 2008.

6. Rau, U., S. Bhatnagar, M. A. Voronkov, and T. J. Cornwell, "Advances in calibration and imaging techniques in radio interferometry," *Proc. IEEE*, Vol. 97, No. 8, 1472–1481, Aug. 2009.
7. Levanda, R. and A. Leshem, "Synthetic aperture radio telescopes," *IEEE Trans. Signal Process.*, Vol. 27, No. 1, 14–29, Jan. 2010.
8. Kundur, D. and D. Hztzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, Vol. 13, No. 3, 43–64, May 1996.
9. Högbom, J. A., "Aperture synthesis with a non-regular distribution of interferometers baselines," *A & AS*, Vol. 15, 417–426, 1974.
10. Cornwell, T. J., "Multiscale CLEAN deconvolution of radio synthesis images," *IEEE J. Select. Topics Signal Process.*, Vol. 2, No. 5, 793–801, Oct. 2008.
11. Bobin, J., J.-L. Starck, and R. Ottensamer, "Compressed sensing in astronomy," *IEEE J. Select. Topics Signal Process.*, Vol. 2, No. 5, 718–726, Oct. 2008.
12. Carrillo, R. E., J. D. McEwen, and Y. Wiaux, "Sparsity averaging reweighted analysis (SARA): A novel algorithm for radio-interferometric imaging," *Mon. Not. Roy. Astronom. Soc.*, 591–594, Jun. 2012.
13. Carrillo, R. E., J. D. McEwen, D. van de Ville, J.-P. Thiran, and Y. Wiaux, "Sparsity averaging for compressive imaging," *IEEE Signal Process. Lett.*, Vol. 20, No. 6, 591–594, Jun. 2013.
14. Carrillo, R. E., J. D. McEwen, and Y. Wiaux, "PURIFY: A new approach to radio-interferometric imaging," *Mon. Not. Roy. Astronom. Soc.*, Vol. 439, No. 4, 3591–3604, Feb. 2014.
15. Johnston, S., M. Bailes, N. Bartel, C. Baugh, et al., "Science with the Astralian square kilometre array pathfinder," *Pub. Astronom. Soc. Australia*, Vol. 24, 174–188, Dec. 2007.
16. Tingay, S. J., R. Goeke, J. D. Bowman, et al., "The Murchison widefield array: The square kilometre array precursor at low radio frequencies," *Pub. Astronom. Soc. Australia*, Vol. 30, e007, 2013.
17. Jonas, J. L., "MeerKAT — The South African array with composite dishes and wide-band single pixel feeds," *Proc. IEEE*, Vol. 97, No. 8, 1522–1530, Aug. 2009.
18. Fomalont, E. and M. Reid, "Microarcsecond astrometry using the SKA," *New Astronomy Rev.*, Vol. 48, 1473–1482, Sep. 2004.
19. Dewdney, P. E., P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio, "The square kilometre array," *Proc. IEEE*, Vol. 97, No. 8, 1482–1496, Aug. 2009.
20. Van Haarlem, M. P., M. W. Wise, et al., "LOFAR: The low-frequency array," *Astron. Astrophys.*, Vol. 556, A2, Jul. 16, 2013.
21. Marco, D. V., W. G. Andre, and N. Ronald, "The LOFAR telescope: System architecture and signal processing," *Proc. IEEE*, Vol. 97, No. 8, 1431–1437, Aug. 2009.
22. Stefan, J. W., "In situ antenna performance evaluation of the LOFAR phased array radio telescope," *IEEE Trans. Antennas Propagat.*, Vol. 59, No. 6, 1981–1989, Jun. 2011.
23. Sutter, P. M., B. D. Wandelt, et al., "Probabilistic image reconstruction for radio interferometers," *Mon. Not. Roy. Astronom. Soc.*, Vol. 438, No. 1, 768–778, Sep. 2013.
24. Cornwell, T. J., K. Golap, and S. Bhatnagar, "The non-coplanar baselines effect in radio interferometry: The W-projection algorithm," *IEEE J. Select. Topics Signal Process.*, Vol. 2, No. 5, 647–657, Oct. 2008.
25. Thompson, A. and R. Bracewell, "Interpolation and Fourier transformation of fringe visibilities," *Astron. J.*, Vol. 79, No. 1, 11V24, 1974.
26. Thiebaut, E. and J. F. Giovannelli, "Image reconstruction in optical interferometry using a general framework to formally describe and compare different methods," *IEEE Signal Process. Mag.*, Vol. 27, No. 1, 97–109, Jan. 2010.
27. De Gasperin, F., et al. "M87 at metre wavelengths: The LOFAR picture," *Astron. Astrophys.*, Vol. 547, A56, Oct. 2012.
28. Griva, I., G. N. Stephen, and S. Ariela, *Linear and Nonlinear Optimization*, Ch. 13, Soc. Indust. Applied Math., 2009.
29. Serge, M. and M. M. Laurent, "Convex approximation to the likelihood criterion for aperture synthesis imaging," *J. Opt. Soc. Am. A*, Vol. 22, 2348–2356, Nov. 2005.