



Intensity-sensitive Quality Assessment of Extended Sources in Astronomical Images

Xiaotong Li, Karel Adámek[✉], and Wesley Armour[✉]Oxford e-Research Centre, Department of Engineering Science, University of Oxford, Oxford, UK; wes.armour@oerc.ox.ac.uk

Received 2023 December 21; revised 2024 June 21; accepted 2024 July 30; published 2024 October 1

Abstract

Radio astronomy studies the Universe by observing the radio emissions of celestial bodies. Different methods can be used to recover the sky brightness distribution (SBD), which describes the distribution of celestial sources from recorded data, with the output dependent on the method used. Image quality assessment (IQA) indexes can be used to compare the differences between restored SBDs produced by different image reconstruction techniques to evaluate their effectiveness. However, reconstructed images (for the same SBD) can appear to be very similar, especially when observed by the human visual system (HVS). Hence, current structural similarity methods, inspired by the HVS, are not effective. In the past, we have proposed two methods to assess point-source images, where low amounts of concentrated information are present in larger regions of noise-like data. But for images that include extended source(s), the increase in complexity of the structure makes the IQA methods for point sources oversensitive because the important objects cannot be described by isolated point sources. Therefore, in this article we propose the augmented low-information similarity index (augLISI), an improved version of LISI, to assess images including extended source(s). Experiments have been carried out to illustrate how this new IQA method can help with the development and study of astronomical imaging techniques. Note that although we focus on radio astronomical images herein, these IQA methods are also applicable to other astronomical images and imaging techniques.

Unified Astronomy Thesaurus concepts: [Radio astronomy \(1338\)](#); [Extended radiation sources \(504\)](#); [Astronomy image processing \(2306\)](#)

1. Introduction

Image quality assessment (IQA) of natural images allows not only for qualitative comparisons of images but is also used to optimize image-processing algorithms (Ding et al. 2021) and to assess the quality of image compression, image restoration, and image-denoising algorithms, among many other uses (Bakurov et al. 2022).

Among numerous IQA metrics, structural similarity (SSIM; Wang et al. 2004, 2005; Brunet 2012) and its variants are frequently used (Bakurov et al. 2022). SSIM aims to incorporate characteristics of the human visual system (HVS) and assess the similarity between two input images by combining luminance, contrast, and structure comparisons.

However, radio astronomical (RA) images tend to be low information, where the important information (celestial sources) is contained in a small proportion of the image while the rest of the image is noise-like. When assessing these low-information images using the HVS, which is most sensitive to phase information (Oppenheim & Lim 1981), the differences between images tends not to be very visible. Consequently, metrics such as SSIM, which seek to incorporate HVS characteristics, perform poorly on sparse astronomical images (Li & Armour 2022).

Some current variants of SSIM attach different importance to different information within an image. However, they are all insufficient to focus on the difference in the high-intensity (informative) part of RA images. Specifically, the weighting process in gradient-based SSIMs (G-SSIMs; Chen et al. 2006) is based on edge and smoothness; fixation (F-), percentile (P-),

and PF-SSIMs (Moorthy & Bovik 2009) aim at weighting by visual importance, while the differences in low-information images may be palpable by glancing; information content weighted SSIM (IW-SSIM; Wang & Li 2011) focuses on perceptual information content; and complex wavelet SSIM (CW-SSIM; Sampat et al. 2009) aims to deal with structural distortions in the complex wavelet domain, while the information that we are interested in may not relate to a structural issue.

We have developed two intensity-sensitive IQA indexes in Li & Armour (2022) to assess low-information images, they are the intensity weighted SSIM index (ITW-SSIM) and the low-information similarity index (LISI). In this work, we introduce an augmented version of LISI (augLISI), which further develops one of the two intensity-sensitive IQA indexes that we have developed to assess low-information images (Li & Armour 2022). Our new low-information IQA method, augLISI, improves the performance of LISI on RA images with extended sources. Our method allows us to:

1. quantitatively compare different implementation approaches for the same algorithm, see Section 4.1;
2. quantitatively define the difference between different algorithms, see Section 4.2;
3. jointly characterize the differences in different parts of the source(s) structure without setting an empirical threshold or extracting sources, see Section 4.3.

Although there are different comparison techniques that are suited for RA images, all focus on source extraction, such as the quality assessment tool (Lu et al. 2022) that is used for image quality assessment by the algorithm developers for the Square Kilometre Array (SKA; Dewdney et al. 2009). In contrast, our method adopts a different strategy, evaluating



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

image quality by analyzing the entire image without the need for source extraction. Notably, to compare the object (no matter source or noise) at the same position in the two input images, we assign greater significance to sources based on their intensities, all without relying on empirical thresholds, while also considering background noise.

2. Radio Interferometric Imaging of Extended Sources

Radio emissions of celestial sources in the sky as detected by an interferometer are known as visibilities (Burke et al. 2019). In an interferometer, the vector that connects the phase centers of two antennas is called a baseline vector. The relationship between observed complex visibilities and sky brightness distribution (SBD) associated with the baseline vector is a Fourier transform under certain conditions (Burke et al. 2019) and can be described (Thompson 1999) by

$$V(u, v, w) = \iint \frac{B(l, m)}{\sqrt{1-l^2-m^2}} e^{-i2\pi(ul+vm+w(\sqrt{1-l^2-m^2}-1))} dldm, \quad (1)$$

where V is the observed complex visibility; B is the distribution of sky brightness; l, m , and n are coordinates in the spatial domain; and u, v , and w are coordinates in the frequency domain.

The inversion from visibilities to image is an ill-posed problem due to limited number of antennas (Dabbech et al. 2015; Pratley et al. 2018). Consequently, radio interferometric (RI) imaging methods are developed to restore the SBD. The most pragmatic RI imaging method is CLEAN (Hogbom 1974), which is a nonlinear deconvolution algorithm. The most traditional CLEAN (Hogbom 1974) method is an iterative process used to extract point sources. However, estimating the sky brightness by point sources is not a sensible assumption for extended sources. There are several RI imaging methods that are suitable for extended source imaging, such as multiscale (MS) CLEAN (Cornwell 2008), the maximum entropy method (MEM; Cornwell & Evans 1985), and compressed sensing (CS) based methods, e.g., isotropic undecimated wavelet transform (IUWT; Wiaux et al. 2009; Li et al. 2011a, 2011b; Dabbech et al. 2015).

2.1. MS-CLEAN

MS-CLEAN has been proposed (Cornwell 2008) to deal with the extended sources by multiscale decomposition. MS-CLEAN looks for the maximum peak of all residual images with different scales and then adds the peak to the model image with a certain loop gain. When the noise level of the residual image is under a certain criteria, a model image is generated, estimating the true sky brightness.

There are many data-processing software packages that can be used to implement RI imaging. The most well-known are Common Astronomy Software Applications (CASA; McMullin et al. 2007; The CASA Team et al. 2022), W-Stacking Clean (WSClean; Offringa et al. 2014; Offringa & Smirnov 2017; Van der Tol et al. 2018), and the Radio Astronomy Simulation, Calibration, and Imaging Library.¹ They all have functionality to implement MS-CLEAN. Herein, we use CASA and WSClean to implement MS-CLEAN.

2.2. MEM

MEM is another RI imaging method that is suitable for imaging extended sources (Frieden 1972; Ponsonby 1973; Ables 1974; Gull & Daniell 1978). A function $F(I_M)$ is defined as the entropy of the intensity distribution of the model I_M . One of the best entropy functions (Narayan & Nityananda 1984) is

$$F = -\sum_k I_{M_k} \ln \frac{I_{M_k}}{I_{M_k^0} e}, \quad (2)$$

where I_M is the model image and I_{M^0} is the a priori model image. The sum is taken over every pixel k in the image. In the imaging process, the entropy is maximized subject to the consistency between the model visibility and the measured visibility. The solution is an iterative process (Cornwell & Evans 1985; Gull & Daniell 1978; Narayan & Nityananda 1984; Wernecke & d'Addario 1977).

Herein, we use GPUVMEM (Cárcamo et al. 2018), which is a GPU-accelerated MEM software package, to perform MEM in Section 4.

2.3. IUWT

IUWT is a CS-based RI imaging method which has good performance for imaging extended sources. CS is a robust technique for signal reconstruction. For a sparse image, the image size can be compressed far below the requirement of Nyquist sampling (Thompson et al. 2017). CS fosters sparse regularization methods that can be used to address inverse problems, such as the inverse transformation from visibility to image in RI imaging (Wiaux et al. 2009). The formal requirements (e.g., sparsity and incoherent sampling) of CS can be roughly met in RI imaging (Thompson et al. 2017). The sensing matrix indicates the mathematical relationship between the image to be reconstructed and the measured visibilities. If the sensing matrix obeys the restricted isometry property (Fornasier & Rauhut 2011; Carrillo et al. 2014), then CS is a promising method for reconstructing the SBD (Zhu et al. 2020).

Alternative CS methods have been applied to aperture synthesis (Li et al. 2011b; Garsden et al. 2015; Hardy 2013), e.g., PURIFY (Carrillo et al. 2014), model reconstruction by synthesis-analysis estimators (MORESANE; Dabbech et al. 2015) and SparseRI (Wenger et al. 2010), rotation measure synthesis (Li et al. 2011a; Sun et al. 2015), and CS-ROMER (Cárcamo et al. 2022). Herein, the reweighted version² of IUWT is used in Section 4.

3. Image Quality Assessment Methods

In this section, we start by reviewing a well-known IQA method, the SSIM index, and two IQA methods proposed in Li & Armour (2022): the ITW-SSIMs with different weighting functions and LISI. Herein, we supplement the ITW-SSIMs and LISI with more details of the derivation and maths. ITW-SSIMs and LISI aim to assess low-information images, e.g., RA images with point sources. However, RA images with extended sources include a complicated structure in the nonresidual part. Therefore, in this article, we propose a new variant of LISI, augmented LISI (augLISI), to assess images containing extended sources in RI imaging. ITW-SSIM, LISI, and augLISI are open source and available on our GitHub page³ and Zenodo (Li et al. 2024).

¹ <https://developer.skao.int/projects/rascil/en/latest/>

² <https://code.google.com/archive/p/csra/downloads>

³ <https://github.com/egbdfX/Intensity-sensitive-IQAs>

3.1. SSIM

The equation describing SSIM ($\text{SSIM} \in [-1, 1]$) is given by

$$\text{SSIM}(p, q) = \frac{(2\mu_p\mu_q + C_1)(2\sigma_{pq} + C_2)}{(\mu_p^2 + \mu_q^2 + C_1)(\sigma_p^2 + \sigma_q^2 + C_2)}, \quad (3)$$

where μ_p (or μ_q) is the mean intensity of image p (or q), σ_p (or σ_q) is the standard deviation of the pixel intensities in image p (or q), σ_{pq} is the cross covariance of pixel intensities in image p and image q , and C_1 and C_2 are constant values to avoid instability when the denominator is very small.

Currently, SSIM is widely discussed and used in IQA. However, the result of SSIM is uninformative for assessing low-information images (Li & Armour 2022), such as RA images, where the importance of different objects in the images is different. Celestial sources are the most significant information in RA images. The SSIM results for RA images always approach 1, due to the limited number of sources and the large noise-like regions in the images. Admittedly, some state-of-the-art SSIM also weight the importance of different information, such as multiscale SSIM (Wang et al. 2003), G-SSIM (Chen et al. 2006), and IW-SSIM (Wang & Li 2011). However, they are not suitable for assessing RA images according to the experimental results in Li & Armour (2022).

Therefore, ITW-SSIMs and LISI have been developed in Li & Armour (2022) for low-information RA images.

3.2. IQA for Images that Only Contain Point Source(s)

In ITW-SSIMs (Li & Armour 2022), different weighting functions are applied to weight the importance of each pixel according to its intensity. The mathematical derivation is provided in Appendix A. In addition to ITW-SSIMs, LISI is another IQA index developed to assess low-information images. The amount of information contained in the inputs should be reflected in an effective intensity-sensitive IQA index. The similarity indicated by the index should be higher when the pixels at the same position in the inputs have similar intensities and the intensities are both high. Following this principle, LISI (Li & Armour 2022) (whose range should be $[0, 1]$) is expressed as

$$\text{LISI}(x, y) = D \frac{\sum_{i=1}^N \frac{|x_i + y_i|}{|x_i - y_i| + C_1}}{\max\left(\sum_{i=1}^N x_i, \sum_{i=1}^N y_i\right) + C_2}, \quad (4)$$

where $C_1 < 1$, $C_2 < 1$, and the inputs x and y are normalized.

To set the factor D in a meaningful way, LISI is supposed to be 1 when the two input images are the same ($x_i = y_i$ so that $|x_i - y_i| = 0$) and informative (with high-intensity parts), which can be expressed as

$$\begin{aligned} \text{LISI}(x, y) &= D \frac{\sum_{i=1}^N \frac{|x_i + y_i|}{|x_i - y_i| + C_1}}{\max\left(\sum_{i=1}^N x_i, \sum_{i=1}^N y_i\right) + C_2} \\ &= D \frac{\sum_{i=1}^N \frac{2x_i}{0 + C_1}}{\sum_{i=1}^N x_i + C_2} = \frac{2D}{C_1} \frac{\sum_{i=1}^N x_i}{\sum_{i=1}^N x_i + C_2} \approx \frac{2D}{C_1} = 1. \end{aligned} \quad (5)$$

LISI is supposed to be 0 when the two input images are entirely different, i.e., $\forall i, x_i \neq y_i$ and $|x_i - y_i|$ is very large (to determine D , we make an assumption that $\forall i, x_i > y_i$), which can be expressed as

$$\begin{aligned} \text{LISI}(x, y) &= D \frac{\sum_{i=1}^N \frac{|x_i + y_i|}{|x_i - y_i| + C_1}}{\max\left(\sum_{i=1}^N x_i, \sum_{i=1}^N y_i\right) + C_2} \\ &\approx D \frac{\sum_{i=1}^N \frac{x_i}{x_i + C_1}}{\sum_{i=1}^N x_i + C_2} \approx D \frac{\sum_{i=1}^N 1}{\sum_{i=1}^N x_i + C_2} \\ &= D \frac{N}{N' + C_2} \approx 0, \end{aligned} \quad (6)$$

where the N' is no greater than N .

Therefore, D should be $D = \frac{C_1}{2}$ and very small.

3.3. IQA for Images that Contain Extended Source(s)

One of the advantages of ITW-SSIMs and LISI is that they do not require an empirical threshold to determine whether a certain pixel belongs to noise or a celestial source. According to the expression of the numerator of LISI, besides large differences in very high-intensity parts, LISI is also sensitive to large differences in lower-intensity parts. When an image contains extended source(s), there will be more differences in the lower-intensity part due to the increased complexity of the image. In such cases, LISI may be overly sensitive.

To illustrate this, a simulation can be performed based on Equation (4). The response of LISI to a pair of single-pixel images x and y is shown in Figures 1(a)–(c), where both the pixel values of x and y are between 0 and 1. As we can see from the figure, LISI does emphasize the high-intensity similarity (see the diagonal portion of the figure) but the value of LISI will decrease dramatically when lower-intensity parts are extended and have differences, indicating that the differences mean a lot to the assessment; in other words, LISI is over-sensitive in this case.

To make the index only sensitive to large differences in celestial sources, a new variant of LISI, augmented LISI ($\text{augLISI} \in [0, 1]$), is proposed in this article. When two images are identical, augLISI will be 1.

It is expressed as

$$\text{augLISI}(x, y) = 1 - \frac{\sum_{i=1}^N |x_i + y_i| |x_i - y_i|}{\sum_{i=1}^N x_i + \sum_{i=1}^N y_i + C}, \quad (7)$$

where $C < 1$ and the inputs x and y are normalized.

It is worth noting that in LISI and augLISI , the denominators are used to reduce the effect of different sizes and intensity levels of the inputs.

The numerator of augLISI is designed to be the multiplication of the sum and the difference of the corresponding pixels in the two input images, rather than division as with LISI. This adjustment ensures that the IQA index is less sensitive when the two inputs are very different but both have low intensities because this would indicate that the noise is very different. The denominator of augLISI is designed to be the sum of all pixels in the two input images, rather than using the maximum sum of the two input images. This adjustment is made to avoid the case where a bright image would dominate a dark image in the comparison, a situation that can occur, for example, when assessing the output from different iterations of

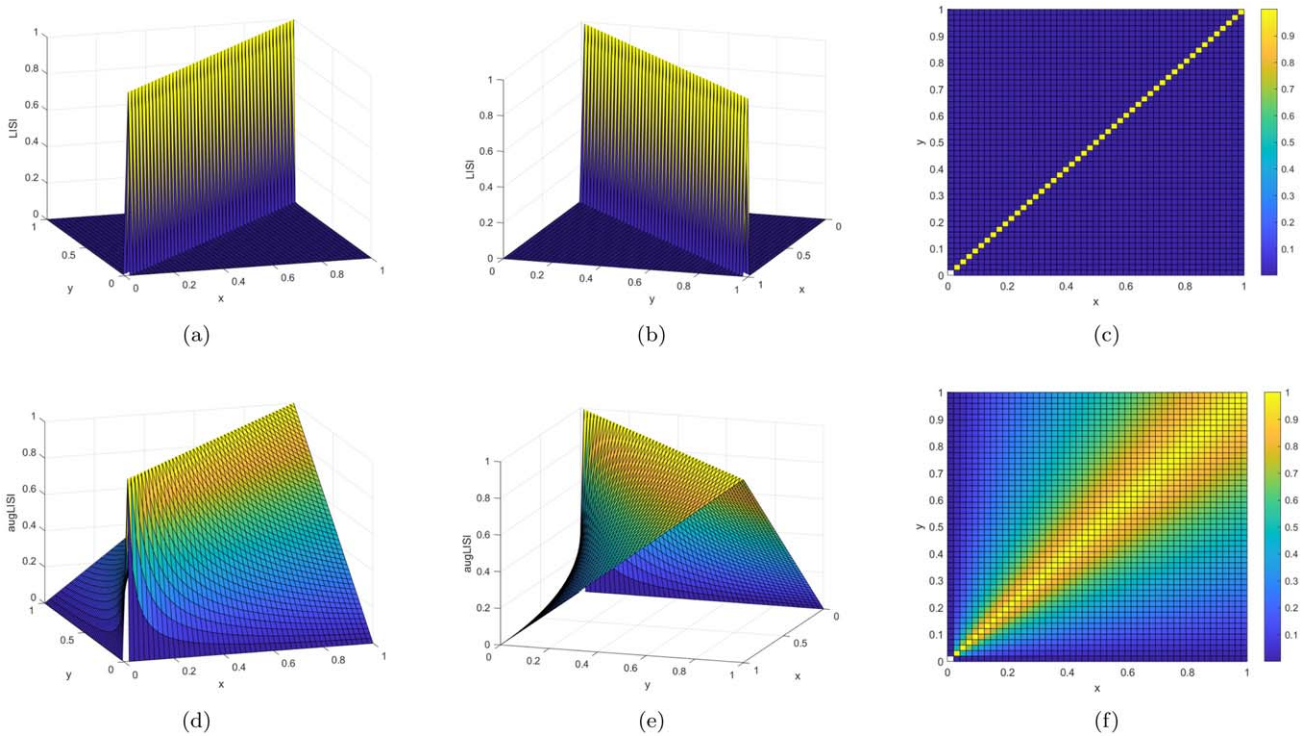


Figure 1. Response of LISI (a), (b), and (c) and augLISI (d), (e), and (f) to a pair of single-pixel images (x, y) , where intensity $x \in [0, 1]$ and $y \in [0, 1]$. To show these 3D plots, three directions are illustrated, separately.

iterative algorithms used in radio astronomy. The response of augLISI is shown in Figures 1(d)–(f), which emphasizes high-intensity differences, but becomes smoother for lower-intensity differences.

Along with the single-pixel simulation, an RA-style data simulation has been done to further support the argument. Oxford’s Square Kilometre Array Radio-telescope simulator (OSKAR; Dulwich 2020) is used to simulate a two-dimensional elliptical Gaussian source having different sizes. In the simulation, 40 pairs of images are generated using the SKA AA2 telescope array layout,⁴ where in each image there is a single source at the center. The size of source is defined by the full width at half maximum (FWHM) Γ ($\Gamma = 0, 1, 2, \dots, 39$ in units of pixels) of the major and minor axes. In each pair of images, the source in one of the images has major and minor axes equal to Γ ; the source in the other image has major and minor axes equal to $\Gamma + 0.1$ and Γ , respectively. Note that “pixel” is used as the unit of source size, rather than using angular resolution, for the reason that the IQA metric is solely based on input images, which should not be affected by the resolution of telescopes. Examples of the simulated Gaussian sources are shown in Figure 2. When using IQA metrics to compare each pair of images with the same Γ value for the minor axis, the difference between the two images in each pair is expected to be small and the IQA should be stable (nearly constant) with the increase in source size because the difference of 0.1 is negligible when comparing the sizes of the sources. The IQA results are shown in Figure 3. According to the result, LISI is more sensitive to the difference in sources; however, it is oversensitive to the difference, especially when the size of the source becomes larger.

LISI is sensitive to large differences in both high- and low-intensity parts of a structure, but will indicate that the inputs are similar only when the high-intensity part of the inputs have small differences. In contrast, augLISI is only sensitive to large differences in the high-intensity parts and will indicate that the inputs are similar in all the other cases.

4. Performance on Data from Extended Sources

In this section, we demonstrate the application of our IQA methods for RI imaging by using two real data sets. The tutorial data⁵ is data measured by the Karl G. Jansky Very Large Array (VLA) of the extended source, supernova remnant G055.7 + 3.4 (referred to as “SNR_G55_10s” hereafter). Along with this, other tutorial data⁶ is data measured by VLA of the extended source, binary black hole system 3C 75 in the Abell 400 cluster of galaxies (referred to as “3C75” hereafter). Sections 4.1–4.3 demonstrate contributions 1–3 presented in Section 1, respectively. It is worth noting that although we used the SKA AA2 telescope array layout to simulate data in Section 3.3, we use data sets from the VLA telescope in this section to demonstrate that augLISI has no bias toward different telescope designs.

4.1. Comparing Different Implementation Approaches for the Same Imaging Algorithm

The development of different imaging algorithms and codes are typically undertaken by disparate research teams often associated with different instruments. Teams formulate distinct implementation approaches for a given algorithm, grounded in

⁴ <https://www.skao.int/en/news/174/green-light-given-construction-worlds-largest-radio-telescope-arrays>

⁵ http://casa.nrao.edu/Data/EVLA/SNRG55/SNR_G55_10s.calib.tar.gz

⁶ https://casa.nrao.edu/Data/EVLA/TDRW0001/TDRW0001_calibrated.ms.tgz

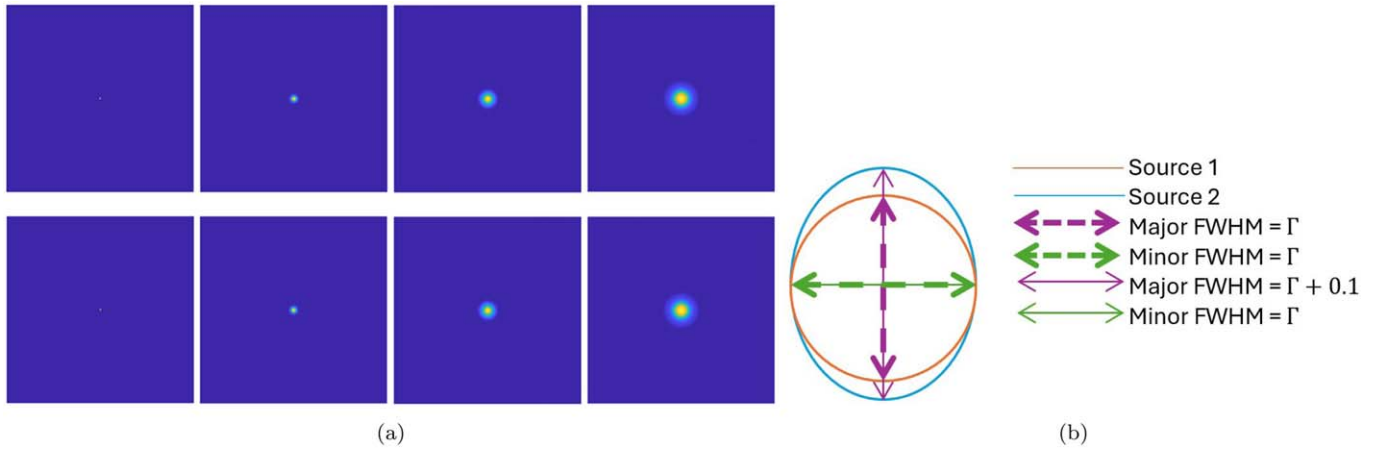


Figure 2. Simulated Gaussian sources with different FWHMs of major and minor axes, where (a) shows four pairs (vertical) of examples, with $\Gamma = 0, 13, 26, 39$ from left to right, and (b) shows a sketch of the two sources in a pair.

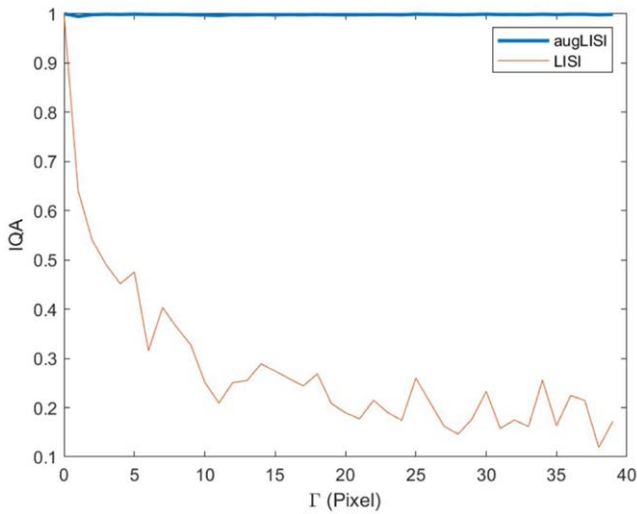


Figure 3. LISI and augLISI results on pairs of images with different sizes of simulated Gaussian sources.

varying development methodologies. For example, both CASA and WSClean have implementations of MS-CLEAN. However, they differ in their implementation, with the most significant difference between them being the process of undertaking gridding. Specifically, WSClean adopts *w*-stacking (Offringa et al. 2014; Ye et al. 2022) for this purpose, while CASA does not incorporate this approach. CASA proffers an array of alternatives for gridding, encompassing standard (prolate spheroid), *w*-projection, faceted, *a*-projection, and *aw*-projection (for more information, refer to “tclean” in CASA.⁷)

Figure 4 shows the MS-CLEAN results for SNR_G55_10s and 3C75 generated by CASA and WSClean. To make the outputs comparable, for the same data set, both implementations use the same image size, cell size, size of CLEAN beam, weighting mode (Briggs weighting; Briggs 1995), and maximum number of iterations. Since WSClean adopts *w*-stacking, which is suitable for wide-field imaging, we set CASA to run *w*-projection to make them as comparable as possible.

It is worth noting that normalization needs to be done in preprocessing of IQA to ensure the intensity levels of the images are comparable. In the preprocessing, we first calculate

the *z*-score $((image - \text{mean}(image))/\text{std}(image))$, where *std* means standard deviation of each image, and then normalize it with the maximum pixel value among all the four images (i.e., images generated by MS-CLEAN in CASA, MS-CLEAN in WSClean, GPUVMEM, and IUWT) so that pixel values of each image are between 0 and 1. The histograms of the images after preprocessing are shown in Figure 5 for SNR_G55_10s and in Figure 6 for 3C75.

The comparisons using SSIM and augLISI are presented in Table 1, where images generated by MS-CLEAN in CASA and WSClean are compared for each data set.

The results of SSIM (0.9846 for SNR_G55_10s and 0.9983 for 3C75, recall a value of 1 indicates the two images are identical) show that the different implementation approaches obtain similar results. SSIM not only reflects the similarity in high-intensity regions (likely to be source) but also reflects the difference in noise pattern.

In contrast, augLISI only reflects the similarity in the high-intensity regions. The results obtained from augLISI confirm that CASA and WSClean produce very similar sources in the output for SNR_G55_10s (0.9916, recall a value of 1 indicates the two images are identical) when using MS-CLEAN, despite the difference in implementation. By visual inspection, the output of CASA is more blurred than the output of WSClean. Since the value of augLISI is larger than that of SSIM, we conclude that the sources in the two images are basically the same and the difference is mainly due to noise or low-intensity structure. For 3C75, the value of augLISI is 0.9928, which is slightly smaller than SSIM (0.9983). But, importantly, the values of both augLISI and SSIM approach 1. In other words, the computational implementation of the same algorithm by two different groups using different programming languages and software engineering practices produces results that generally agree.

Quantitatively, LISI, 0.0479 for SNR_G55_10s and 0.0491 for 3C75 (recall a value of 1 indicates the two images are identical), is oversensitive for extended structures, as explained in Section 3.3.

To illustrate the uncertainty of augLISI and SSIM in Table 1 in the presence of noise, we simulate data sets with a similar noise level as the real data sets but with different noise realizations. The noise level in the real data set is estimated by fitting the histograms of the real and the imaginary parts of the visibilities, which are dominated by noise, with Gaussian

⁷ <https://casa.nrao.edu>

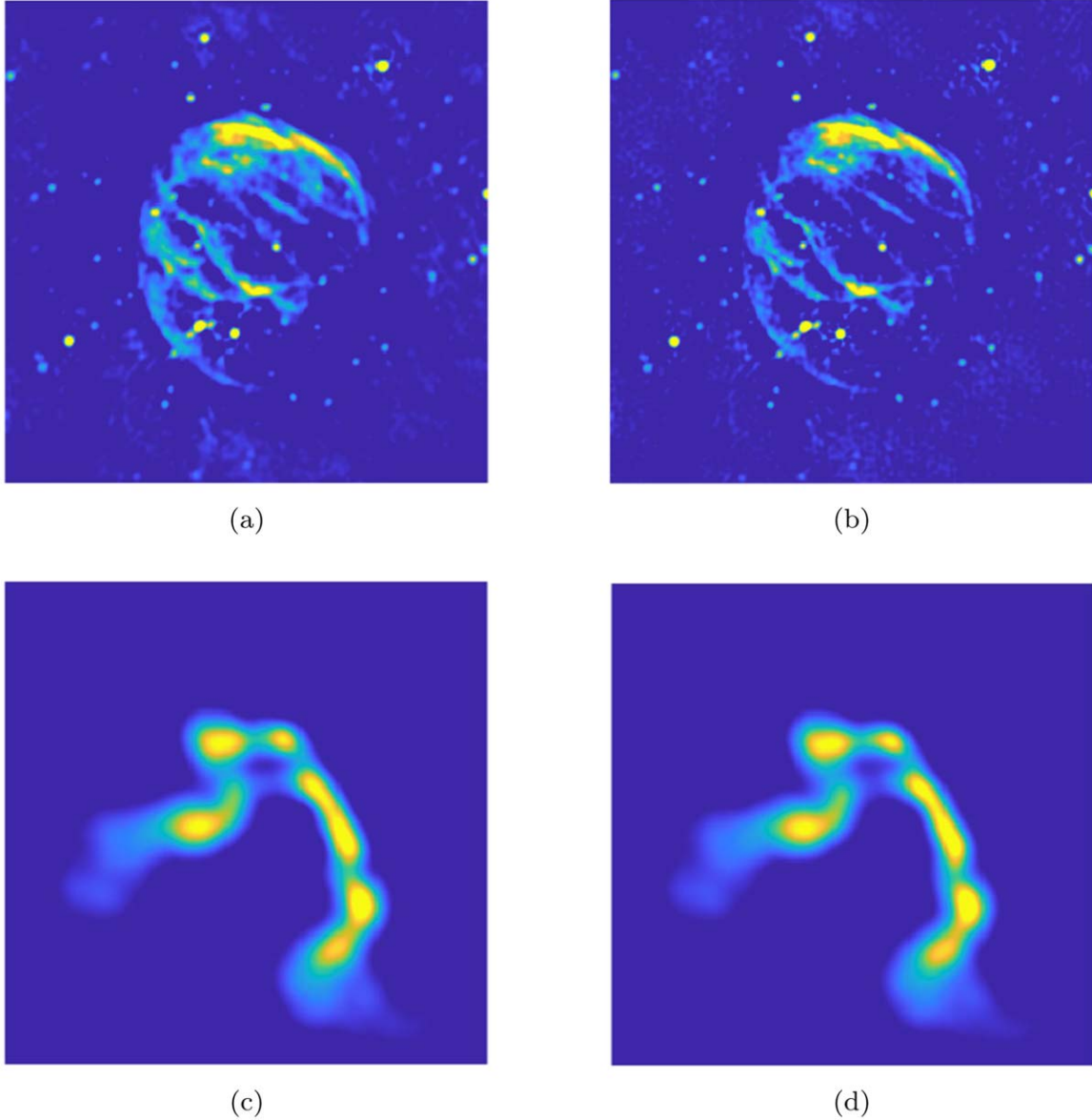


Figure 4. MS-CLEAN results of SNR_G55_10s generated by (a) CASA and (b) WSClean, and MS-CLEAN results of 3C75 generated by (c) CASA and (d) WSClean.

functions using MATLAB’s Open Curve Fitter app. For SNR_G55_10s, the R-square values when fitting to the real and imaginary parts are 0.9967 and 0.9968, respectively; for 3C75, the R-square values when fitting to the real and imaginary parts are 0.9623 and 0.9686, respectively. Different noise realizations are produced by generating different random values drawn from the relevant Gaussian distribution and adding these values back into the original data set. This increases the noise in the original data but ensures that the level of noise in each noise realization is the same. The results of SSIM and augLISI comparing the restored images generated by CASA and WSClean to the simulated data sets are shown in Table 1. The mean values of augLISI and SSIM for the 10 different noise realizations are shown, respectively, in the table along with their uncertainties expressed in standard deviations.

We can see from Table 1 that augLISI is more stable than SSIM in the presence of noise. The IQA results of simulated data sets further support the above analysis of the real data sets. With augLISI, users are able to scientifically prove the

similarity between different implementation approaches for the same algorithm, rather than judging intuitively. It is worth noting that SSIM and augLISI have similar performance in the above examples because both noise pattern and source structure are similar in the two implementation approaches.

To further demonstrate the enhancement of augLISI compared to SSIM and their uncertainties in the presence of noise, we simulate measurement sets with different noise levels. In OSKAR, noise is added to the visibilities by generating random numbers from a Gaussian distribution with a width given by the specified rms value. Those values are then added independently to the real and imaginary components of the visibilities. This method is essentially the same as the one used in the simulation of Table 1. For each noise level, multiple noise realizations are generated by using different random seeds, these noise realizations are then used to statistically assess uncertainties. Take SNR_G55_10s as an example. The simulated data sets adopt the same telescope array as its real measurement set. We generated 100 simulated images of SNR_G55_10s with 10

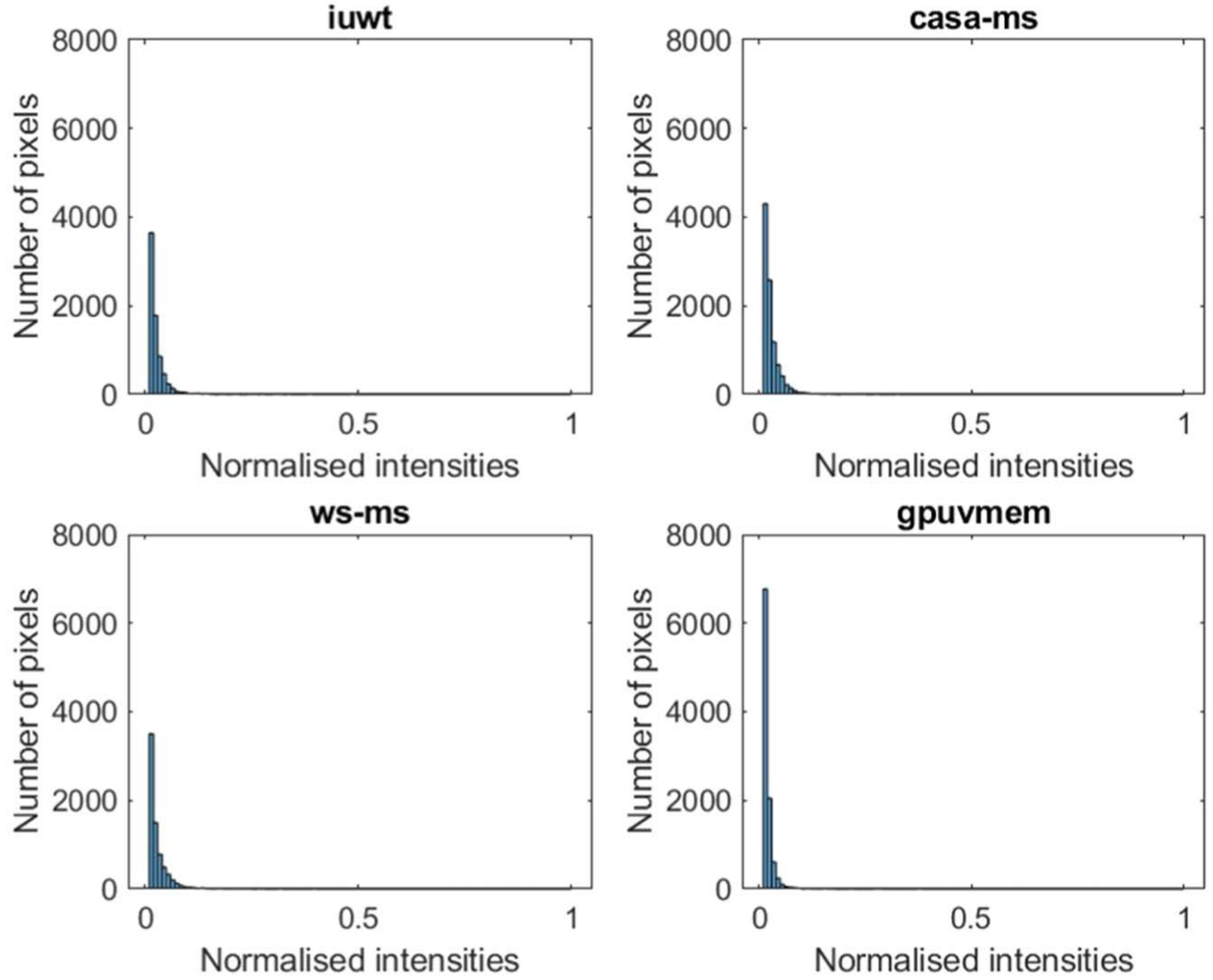


Figure 5. Histograms of SNR_G55_10s for the results of (top left-hand panel) IUWT, (top right-hand panel) MS-CLEAN by CASA, (bottom left-hand panel) MS-CLEAN by WSClean, and (bottom right-hand panel) MEM by GPUVMEM. The distributions are depicted by histograms of the normalized intensities of each pixel. To show the histograms in meaningful ways, all 0 values are ignored when plotting the histograms.

different noise levels, each level having 10 different noise realizations using OSKAR. We define the additive noise levels using the 3σ value from the random noise distribution, called “ 3σ noise”. In the images, the 3σ noise ranges from 0% to 4.42% of the maximum intensity, increasing in increments of approximately 0.5%, resulting in 10 distinct noise levels. The maximum intensity is determined by the highest pixel value in the simulated source without any additive noise, named “0-additive-noise source”. We obtain MS-CLEAN (Briggs weighting) results by using WSClean on each image. A set of images restored by WSClean with different noise levels and different noise realizations are shown in Figure 7 as an example.

As the difference in the level of noise grows between input images (with the sources in those input images remaining the same), the difference between augLISI and SSIM will grow. To illustrate this, Table 2 shows the SSIM and augLISI results comparing the 0-additive-noise image with increasingly noisy images, with the standard deviation (Std) included below the corresponding mean of the IQA. When the noise increases, we still expect the images to be assessed as similar because the sources are same in these data sets. It can be seen from the results that SSIM is affected more by the noise than augLISI.

The gap between the values of SSIM and augLISI becomes larger when the noise increases. Additionally, the standard deviations of SSIM are generally larger than those of augLISI, suggesting that augLISI increases the stability of assessing images in the presence of noise.

The values of SSIM and augLISI can be very different when the differences in noise pattern or source structure are larger. This will be further discussed in Section 4.3.

4.2. Quantitatively Defining the Difference between Different Algorithms

As CLEAN achieves extended imaging by applying the CLEAN beam, both MEM and compressed sensing methods are believed to be superior approaches for imaging extended sources. Our next research question pertains to the extent of improvement achieved by MEM and compressed sensing methods in comparison to CLEAN.

Figure 8 shows (a) the IUWT result and (b) the MEM result of SNR_G55_10s, and (d) the IUWT result and (e) the MEM result of 3C75. To make the images comparable, parameters are set in the implementations to be as close as possible with each other. In IUWT and GPUVMEM, the image size,

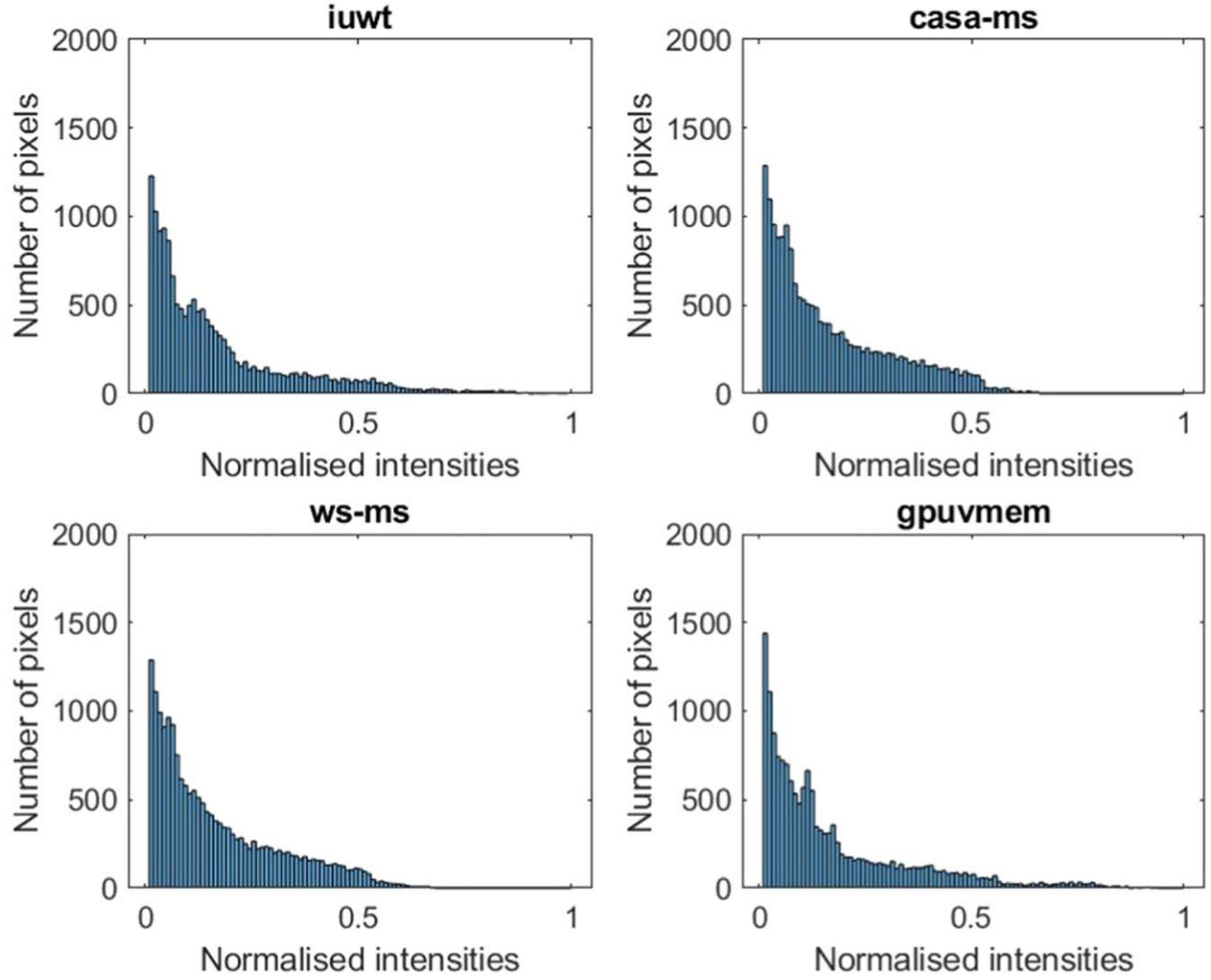


Figure 6. Histograms of 3C75 for the results of (top left-hand panel) IUWT, (top right-hand panel) MS-CLEAN by CASA, (bottom left-hand panel) MS-CLEAN by WSClean, and (bottom right-hand panel) MEM by GPUVMEM. The distributions are depicted by histograms of the normalized intensities of each pixel. To show the histograms in meaningful ways, all 0 values are ignored when plotting the histograms.

Table 1
SSIM and augLISI for Assessing the Results of Different Implementations of MS-CLEAN on Real and Simulated Data Sets

MS-CLEAN	Data Set	SSIM	augLISI
CASA vs WSClean	SNR_G55_10s (real)	0.9846	0.9916
	3C75 (real)	0.9983	0.9928
	SNR_G55_10s (simulated)	0.9793 (± 0.0019)	0.9902 (± 0.0005)
	3C75 (simulated)	0.9983 ($\pm 2.653 \times 10^{-5}$)	0.9928 ($\pm 6.158 \times 10^{-5}$)

cell size, and weighting mode (Briggs weighting) are set by inputting either parameters or an example FITS file (e.g., the dirty image) that includes the information in the header. The termination conditions in IUWT and GPUVMEM are different from the setting of CASA and WSClean. The process of IUWT terminates when the change of residual images between adjacent iterations is small enough, judging by the Frobenius norm, as shown in

$$\epsilon = \frac{\|\Delta \text{Residual}\|_F^2}{\|\text{Residual}\|_F^2}, \quad (8)$$

where ϵ indicates the change, $\Delta \text{Residual}$ is the difference in residual images between adjacent iterations, and $\|\cdot\|_F^2$ is the

square of the Frobenius norm. The termination of GPUVMEM is based on the convergence of conjugate gradient (Cárcamo et al. 2018). Although the setting modes are different in different implementations, we set them to be comparable based on the similarity of noise distribution (Figures 5 and 6, these are the best cases in our experiments to make the noise as similar as possible).

Upon visual inspection, MEM provides better image quality for the extended source compared with the other three approaches. At the same time, IUWT on 3C75 obtains an image with a sharper boundary of the source. To quantitatively measure the extent of difference achieved by these algorithms, augLISI is used. Comparisons using SSIM and augLISI

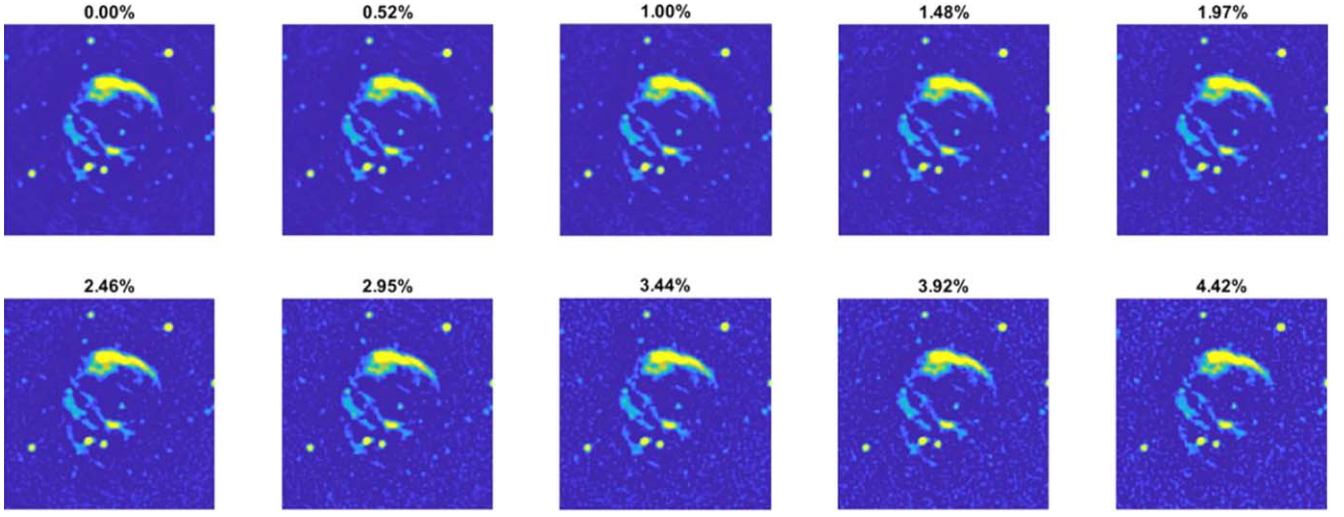


Figure 7. Restored images of simulated SNR_G55_10s generated by WSClean. The 3σ additive noise levels of these figures are set to be a certain percentage (as shown above in each image) of the maximum intensity in the 0-additive-noise celestial source image.

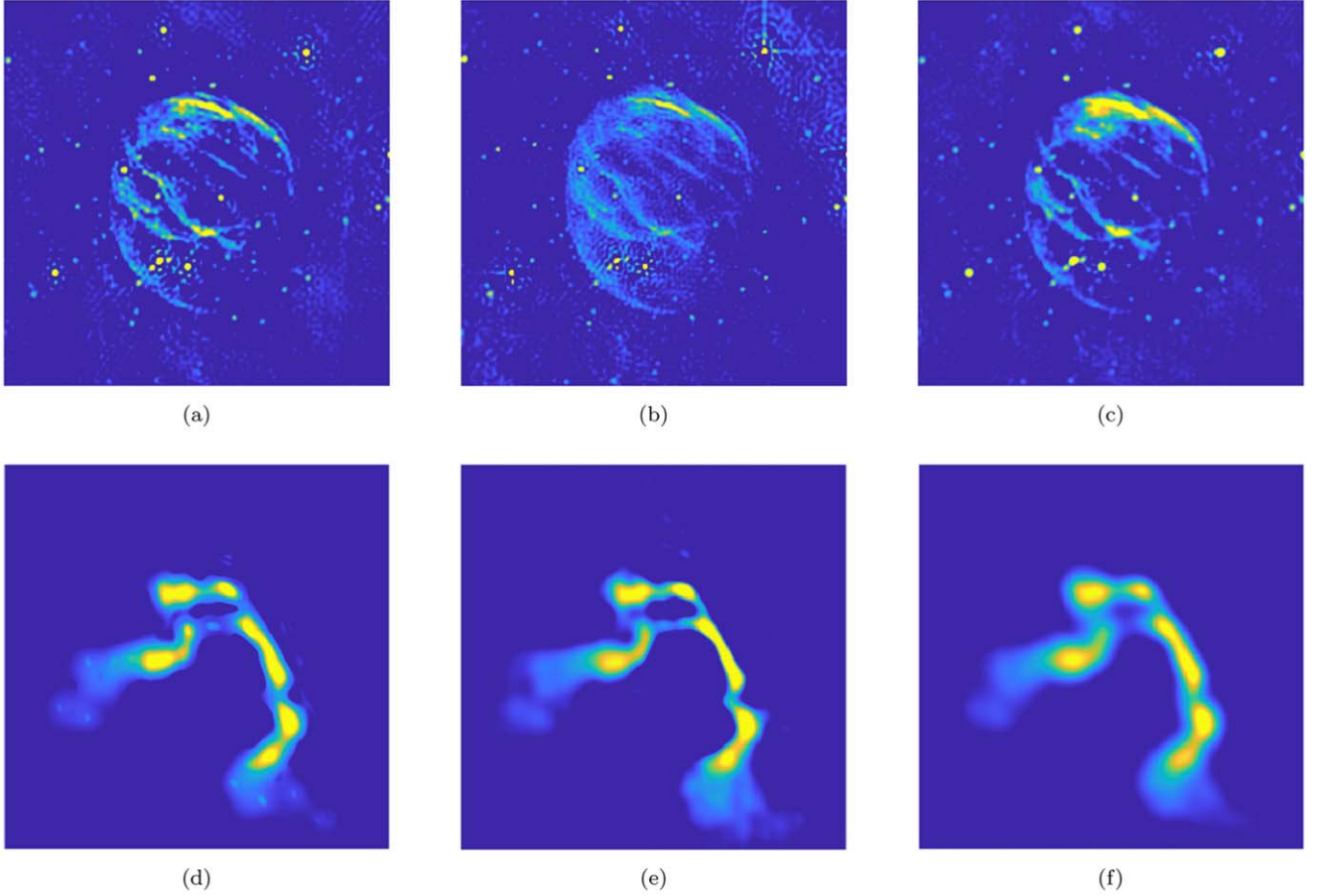


Figure 8. (a) IUWT, (b) MEM, and (c) MS-CLEAN (WSClean, recalled Figure 4) results of SNR_G55_10s, and (d) IUWT, (e) MEM, and (f) MS-CLEAN (WSClean, recalled Figure 4) results of 3C75.

Table 2
SSIM and augLISI of 0-additive-noise Restored Image and Increasingly Noisy Restored Images Generated by Simulated Data Sets

Noise Level (%)		0 vs 0.00	0 vs 0.52	0 vs 1.00	0 vs 1.48	0 vs 1.97	0 vs 2.46	0 vs 2.95	0 vs 3.44	0 vs 3.92	0 vs 4.42
SSIM	Mean	1.0000	0.9965	0.9874	0.9733	0.9553	0.9346	0.9120	0.8887	0.8647	0.8423
	Std	0.0000	0.0002	0.0005	0.0007	0.0011	0.0014	0.0017	0.0022	0.0024	0.0031
augLISI	Mean	1.0000	0.9984	0.9972	0.9960	0.9947	0.9934	0.9921	0.9907	0.9894	0.9879
	Std	0.0000	0.0003	0.0002	0.0002	0.0002	0.0003	0.0003	0.0003	0.0004	0.0004

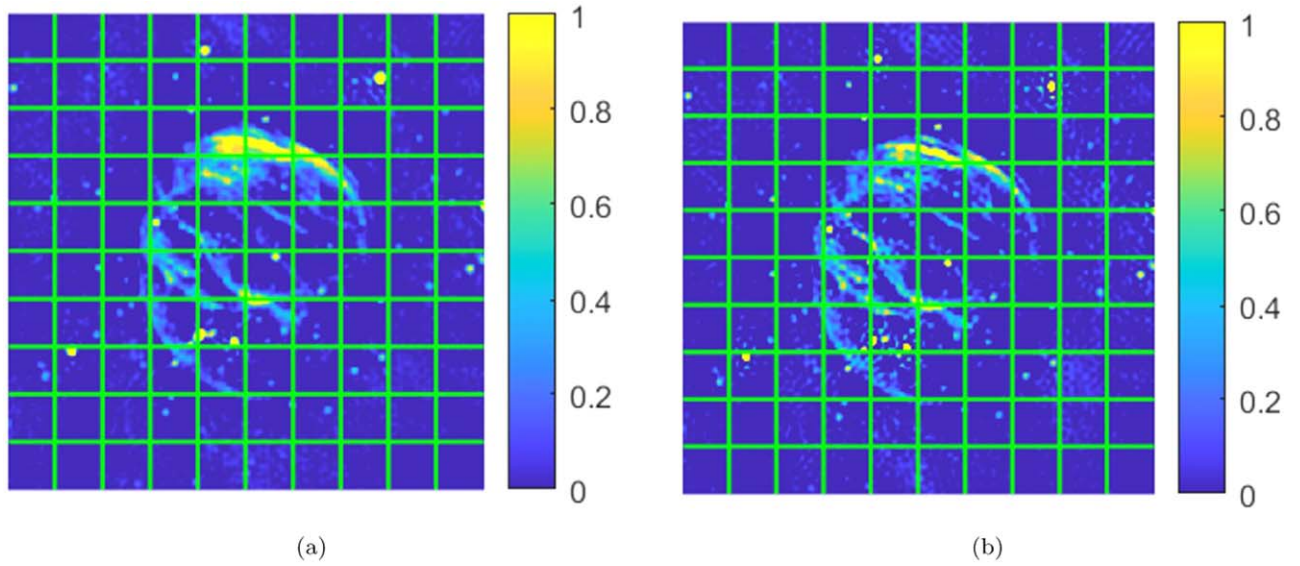


Figure 9. Regular tiles of images generated by (a) MS-CLEAN (in WSClean) and (b) IUWT on SNR_G55_10s. As shown in the color scales, the intensity of each pixel has been normalized by the maximum value of the two images. The tile size can be adjusted based on users' requirement. The tile index (1, 1) starts from the top left-hand corner of each image.

Table 3

SSIM and augLISI for Assessing the Results of Different Imaging Algorithms

Data Set	Algorithms	SSIM	augLISI
SNR_G55_10s	MS-CLEAN vs MEM	0.9554	0.9748
	IUWT vs MEM	0.9604	0.9767
	MS-CLEAN vs IUWT	0.9870	0.9870
3C75	MS-CLEAN vs MEM	0.9155	0.9277
	IUWT vs MEM	0.9326	0.9439
	MS-CLEAN vs IUWT	0.9435	0.9356

between MS-CLEAN (due to the similarity between the MS-CLEAN results generated by CASA and WSClean, we only consider the WSClean result for this comparison), IUWT, and MEM are presented in Table 3. We again see that SSIM (smaller than augLISI) reflects both the difference in residuals and high-intensity source structures in the images, while augLISI provides a more accurate comparison of the extended source restoration. Interestingly, in the comparison of MS-CLEAN and IUWT for SNR_G55_10s, SSIM (0.9870) and augLISI (0.9870) are the same to four decimal places. Since SSIM does not have a preference about the difference in high or low-intensity parts, augLISI only focuses on high-intensity difference and is less sensitive to low-intensity difference. When SSIM and augLISI are same, it means that the difference between the two images concentrates on the high-intensity structure. This aligns with visual inspection.

For the SNR_G55_10s data, the result of augLISI in Table 3 shows that the similarity between IUWT and MS-CLEAN (0.9870) is higher than between MEM and either algorithm (0.9748 for MS-CLEAN and 0.9767 for IUWT). Upon visual inspection, the results of IUWT and MS-CLEAN are more similar, which confirms the augLISI results. Similarly, for the 3C75 data, generally the high-intensity parts of the source are

successfully recovered by IUWT, MEM, and MS-CLEAN. Among the results of the three algorithms, MS-CLEAN produces the result with a diffuse boundary, while the boundary of the result generated by IUWT has a sharper transition. This leads to a smaller augLISI (0.9356) compared to SSIM (0.9435) when comparing the outputs of MS-CLEAN and IUWT. More details will be shown in the next section for analyzing the characteristics of the image using multiple IQA indexes jointly.

4.3. Jointly Analyzing Source and Noise

When IQA metrics indicate disparities, users may still find it challenging to discern whether these differences primarily arise from differences in the source(s) or in background noise. Observed differences in the source(s) could be resolved by adjustment of the imaging algorithm because the source(s) might not be fully resolved. If the difference is in the noise, we can anticipate that the imaging algorithm has restored the source(s) sufficiently. The issue can be more complicated for assessing an extended source(s), as differences can occur in any part of the extended structure.

Analyzing signal and noise together can be useful in signal processing, e.g., in acoustics (Asano & Asoh 2011) and auditory processing (Kaghaz-Garan et al. 2014). For assessing RA images, researchers can analyze the noisy image by extracting faint parts (Teeninga et al. 2016) or extracting contours and textures (Zhu et al. 2019).

Better still, IQA methods can work jointly without extraction of source(s) to analyze where the structure of the extended source(s) differs in the two input images to be compared. Take the analysis of MS-CLEAN vs IUWT of SNR_G55_10s as an example. The first step of the joint analysis should be to split the images into regular tiles, as shown in Figure 9, so that the differences can be localized to each tile, also allowing the following computations on each tile to be performed in parallel. We then calculate SSIM and augLISI on every pair of tiles

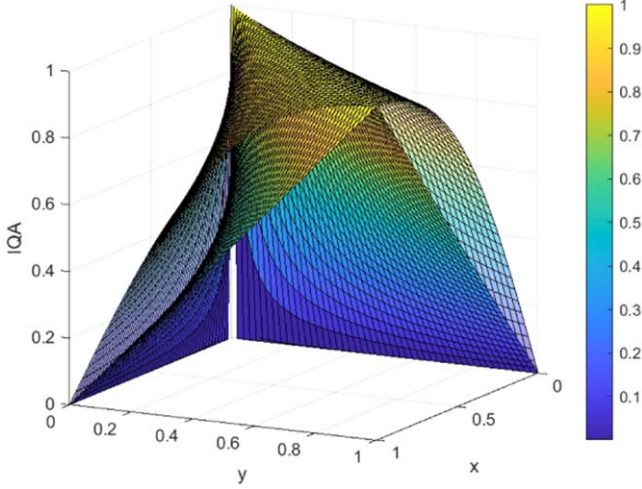


Figure 10. Response of augLISI and SSIM to a pair of single-pixel images (x , y), where intensity $x \in [0, 1]$ and intensity $y \in [0, 1]$. It shows augLISI and SSIM on a grid with a color scale, where the surface with 50% transparency represents the response of SSIM and the other surface represents the response of augLISI.

formed from the same tile position in the two input images to obtain values for the SSIM matrix and augLISI matrix. Note that SSIM is used because it does not have a preference regarding high- or low-intensity differences and augLISI is used as these images contain extended sources.

When comparing the calculated IQA values for each image tile, it is important to assess two quantities. The first is whether the difference between the IQA values is sufficiently small, where a threshold δ is chosen based on requirements, i.e., if

$$|\text{augLISI}(I_1, I_2) - \text{SSIM}(I_1, I_2)| \leq \delta, \quad (9)$$

where I_1 and I_2 are inputs to be compared by the IQA indexes. The second is defining whether augLISI is sufficiently large, where a threshold τ is again chosen based on requirements, i.e., if

$$\text{augLISI}(I_1, I_2) \geq \tau. \quad (10)$$

To find meaningful values of δ and τ , a simulation can be done based on the augLISI equation, i.e., Equation (7). To illustrate the response of augLISI and compare it to the response of SSIM, we generate two input images x and y . Each image consists of a single pixel with an intensity value between 0 and 1. Figure 10 shows the output of augLISI and SSIM for different (x, y) ($x \in [0, 1]$, $y \in [0, 1]$).

To further illustrate the difference between augLISI and SSIM for the same inputs, Figure 11 shows regions where $|\text{augLISI}(x, y) - \text{SSIM}(x, y)| \leq \delta$. In other words, samples of inputs which have similar responses of augLISI and SSIM are shown in this figure, where the “similar” is defined by δ . There are two structures visible for each δ . One is a cluster of IQA points with values approaching 1, where x and y are equal or close to each other. The other is the IQA points with values near 0 located on both x and y -axes, where x and y are very different. In the areas without any points, the augLISI response is lower than the SSIM response for the single-pixel inputs. The key difference between augLISI and SSIM is in the treatment of low-intensity values, as demonstrated by the tapered shape of the two structures shown in Figure 11. The augLISI assigns low-intensity pixels, which are likely to be noise, small

significance, and thus it is less sensitive to their contribution when comparing images.

Smaller significance of low-intensity pixels also implies, as shown by the tip of the tapers in Figure 11, that low-intensity pixels must be very close to each other to be considered the same. In greater detail, this behavior is illustrated in Figure 12, where we see that the augLISI response can be smaller than 1 even for inputs with similar intensities when they are both small (yellow circle). The reason is that in this case x and y are very likely to be noise, while augLISI focuses on the similarity of sources and does not care about the similarity of noise.

We can determine the appropriate value of δ based on the relationship between δ (Figure 11) and $|x-y|$ (Figure 12). The relationship between them is that the maximum difference between x and y for $\delta = 0.02$ is 0.0200, for $\delta = 0.1$ is 0.1062, and for $\delta = 0.2$ is 0.2345. Therefore, δ needs to be set to satisfy the noise sensitivity required in the users’ application. The ranges of noise are different for different δ , as shown Figure 12 (yellow circle). In our example for assessing the images generated by MS-CLEAN (in WSClean) and IUWT (Figure 9), the noise distributions of these two figures are shown in Figure 5 (bottom left-hand panel) and (top left-hand panel). The standard deviations σ of the noise in the images generated by MS-CLEAN (in WSClean) and IUWT are 0.0201 and 0.0153, respectively. Therefore, based on the aforementioned relationship, we set $\delta = 0.02$ (a value with proper fractional precision closest to both standard deviations of noise) for assessing images of SNR_G55_10s.

To establish whether augLISI is sufficiently large, i.e., to determine τ , users need to consider their required similarity between the images. Specifically, Figure 13 depicts the augLISI results for (x, y) when setting different values of τ . When τ is smaller, the tolerance to difference in larger intensities (i.e., large x and y) is higher, reflected in the extended area of intensities shown Figure 13, which means the sensitivity to difference in large intensity is lower. The tolerance/sensitivity to difference in small intensity does not change significantly. To select a proper τ , the sensitivity can be approximated by the proportion of the colored area to the area of the whole normalized region. In our example for assessing the images generated by MS-CLEAN (in WSClean) and IUWT (Figure 9), we select $\tau = 0.85$ to keep the sensitivity of augLISI focused on the highest 15% of intensities.

Applying δ and τ generates four possible scenarios when comparing augLISI and SSIM, which are shown in Table 4 (supported by Figure 14 and Table 5). In the table, $\text{augLISI} > \text{SSIM}$ or $\text{augLISI} < \text{SSIM}$ indicates that the difference between augLISI and SSIM is more than δ ($\delta = 0.02$ in our example); $\text{augLISI} \approx \text{SSIM}$ indicates that the difference between augLISI and SSIM is smaller than δ . The “both large/small” indicates that both augLISI and SSIM are larger/smaller than τ ($\tau = 0.85$ in our example). There is no tile in the example images, where $\text{augLISI} \approx \text{SSIM}$ (both small).

According to the IQAs of each tile in SNR_G55_10s, non-shadowed tiles in Figure 15(a) show that the bright structures in the inputs are similar although the faint structures are very different. Only tile (8,2) (the orange shadowed tile in Figure 15(a)) shows that the bright structures in the inputs are very different. Although there is no gold standard between the images generated by MS-CLEAN (in WSClean) and IUWT,

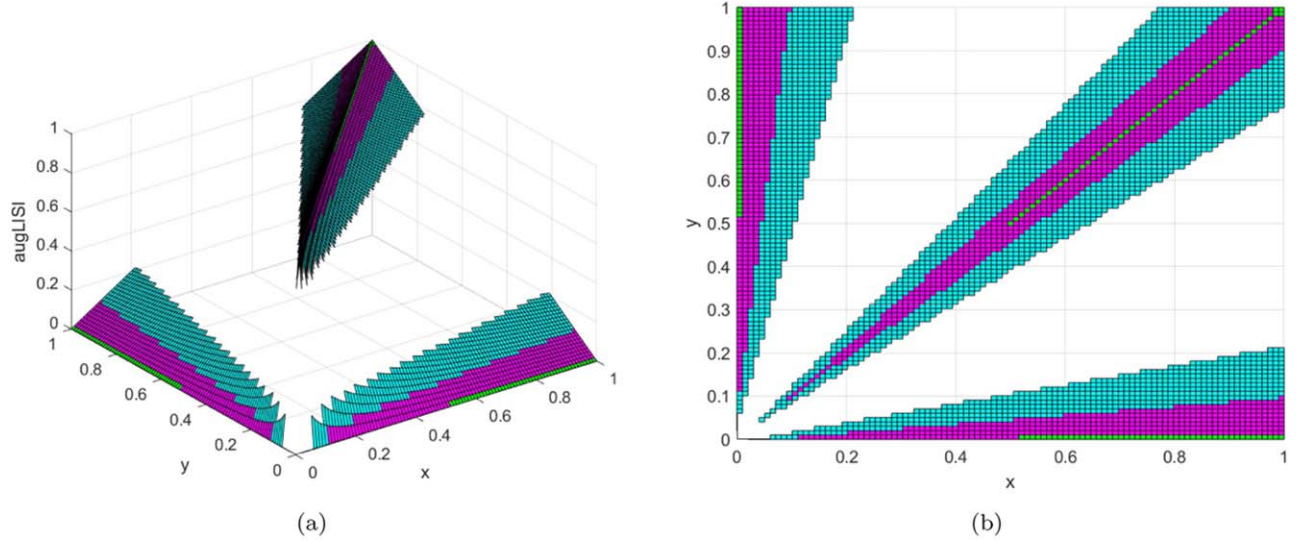


Figure 11. Response of augLISI to pixel intensities for condition $|\text{augLISI}(x, y) - \text{SSIM}(x, y)| \leq \delta$. Different colors represent different values of δ , where green points are for $\delta = 0.02$, magenta points are for $\delta = 0.1$, and cyan points are for $\delta = 0.2$. In this figure, (a) shows the 3D plot of the IQA results and (b) shows the 2D version of it by looking from the top of the 3D plot.

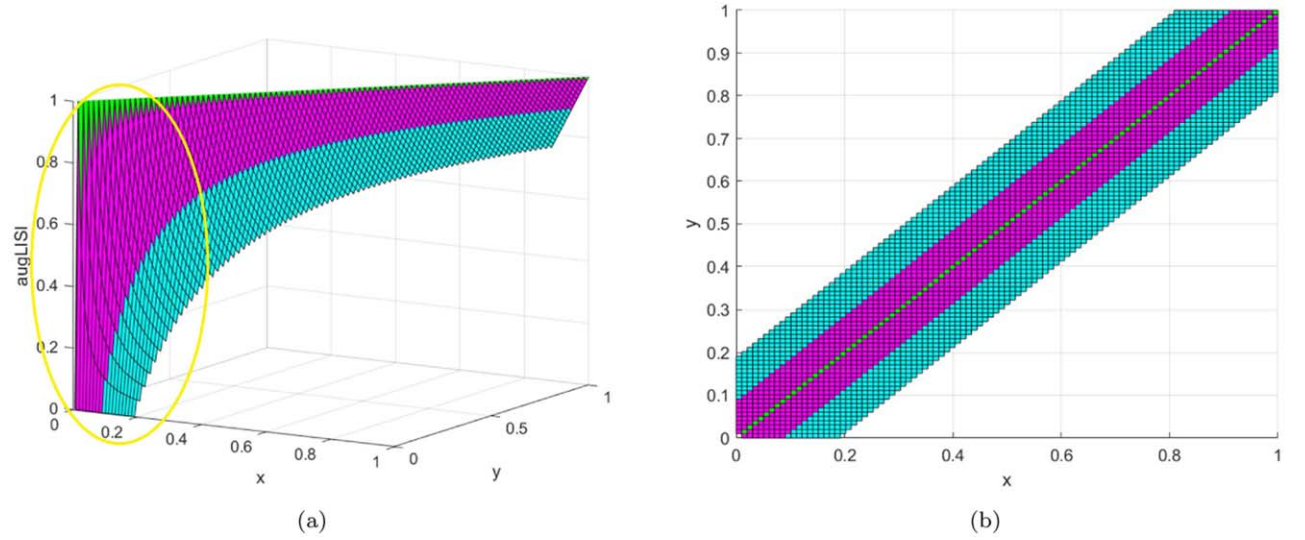


Figure 12. Response of augLISI to pixel intensities x and y when $|x-y|$ is small. Sensitivity to low or near-zero intensities is suppressed in augLISI, so even similar values of (x, y) do not definitely result in a response of 1. Different colors represent different differences between x and y , where $|x-y| \leq 0.02$ for green points, $|x-y| \leq 0.1$ for magenta points, and $|x-y| \leq 0.2$ for cyan points. In this figure, (a) shows the 3D plot of the augLISI results and (b) shows the 2D version of it by looking from the top of the 3D plot. The yellow circle in (a) shows the regions of small intensities (noise).

Table 4
Cases of augLISI and SSIM in Joint Analysis

Case	Description	Example
$\text{augLISI} > \text{SSIM}$	The bright structures in the inputs are similar, although the faint structures are very different.	Figures 14(a) and (d)
$\text{augLISI} < \text{SSIM}$	The bright structures in the inputs are very different, which dominate the IQA result.	Figures 14(b) and (e)
$\text{augLISI} \approx \text{SSIM}$ (both large)	The bright structures in the inputs are similar and the faint structures are similar or slightly different.	Figures 14(c) and (f)
$\text{augLISI} \approx \text{SSIM}$ (both small)	The bright structures in the inputs are very different, while the faint structures are similar or slightly different.	...

Note. This can be evidenced by considering the response to different input values into Equations (3) and (7).

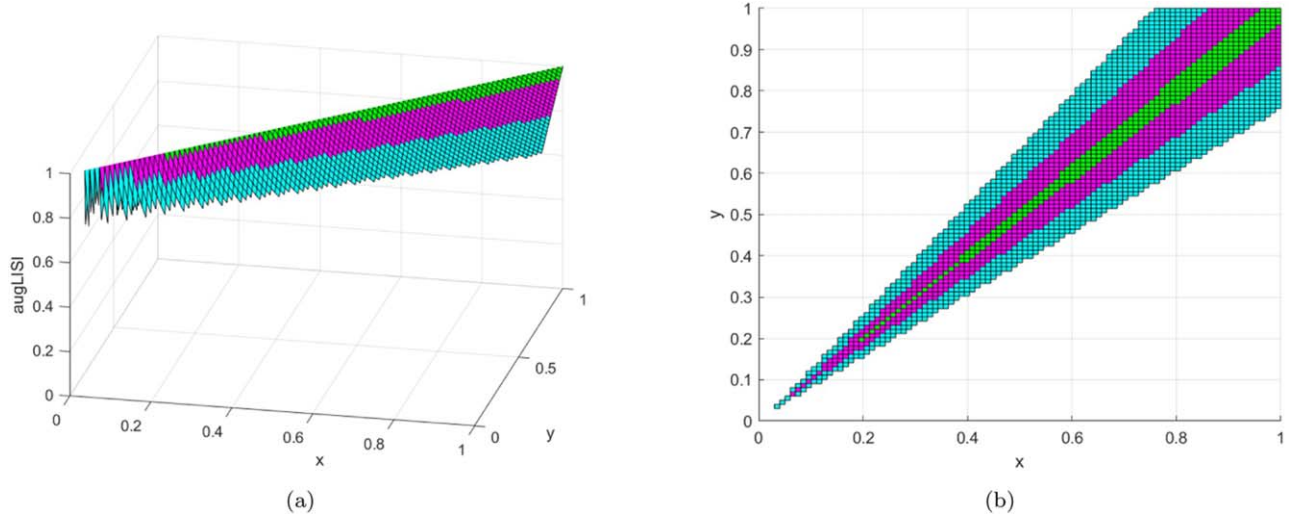


Figure 13. Response of augLISI for (x, y) with different τ , where $\tau = 0.95$ for green points, $\tau = 0.85$ for magenta points, and $\tau = 0.75$ for cyan points., while (a) shows the 3D plot of the augLISI results and (b) shows the 2D version of it by looking from the top of the 3D plot.

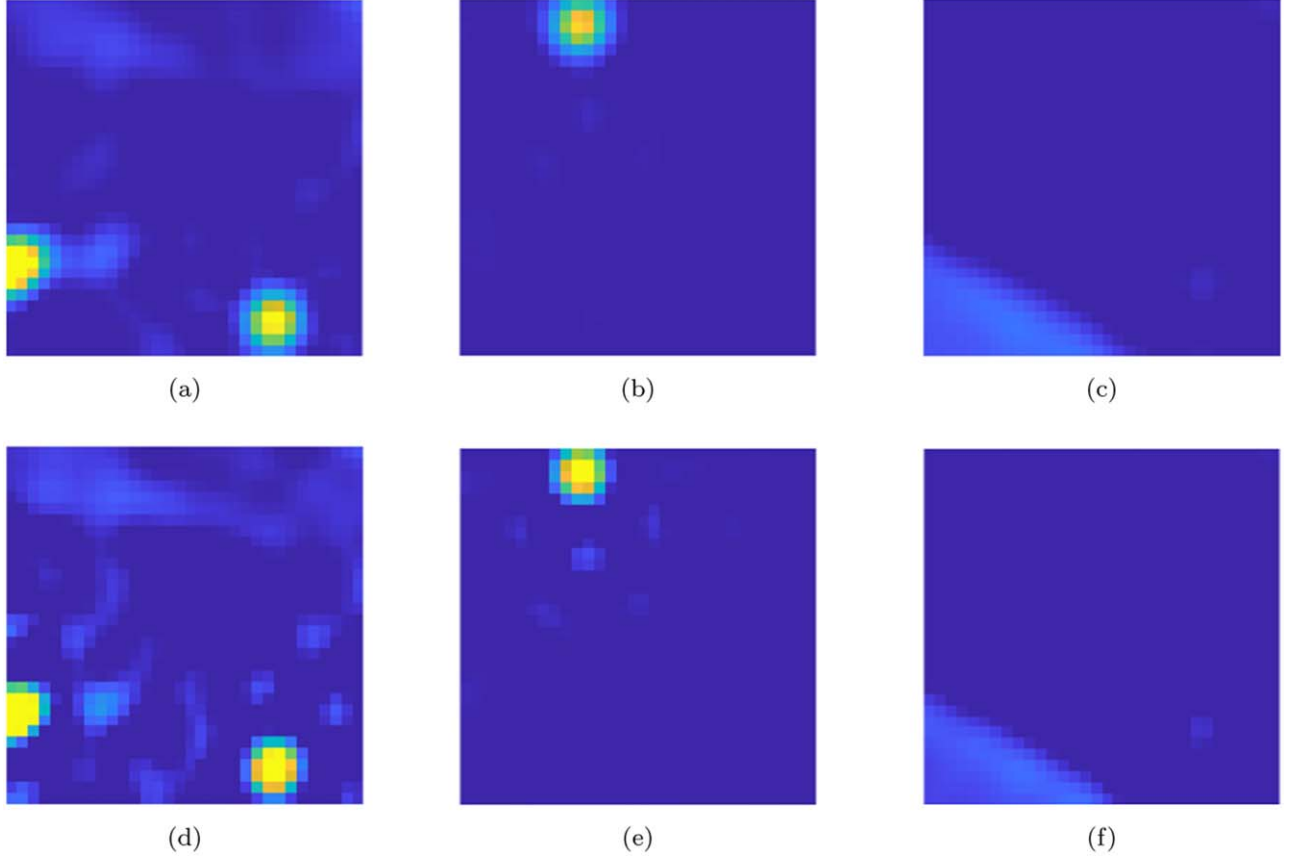


Figure 14. With data of SNR_G55_10s, tile index (7, 5), (8, 2), and (3, 7) in the images generated by MS-CLEAN are illustrated in (a), (b), and (c), respectively; those three tiles in the images generated by IUWT are illustrated in (d), (e), and (f), respectively. The SSIM and augLISI values for each pair of tiles are shown in Table 5.

the joint analysis suggests that further development/improvement of either algorithm should focus on the orange shadowed tile, which, in this example, includes a pointlike source.

Similarly, to assess the images of 3C75 generated by MS-CLEAN (in WSClean) and IUWT, we select $\delta = 0.03$ and

$\tau = 0.85$. As shown in Figure 15(b), the bright structures of the source are correctly restored but the faint structures of the source are very different. Further development/improvement of either algorithm (especially CLEAN) should focus on the fainter regions of this data.

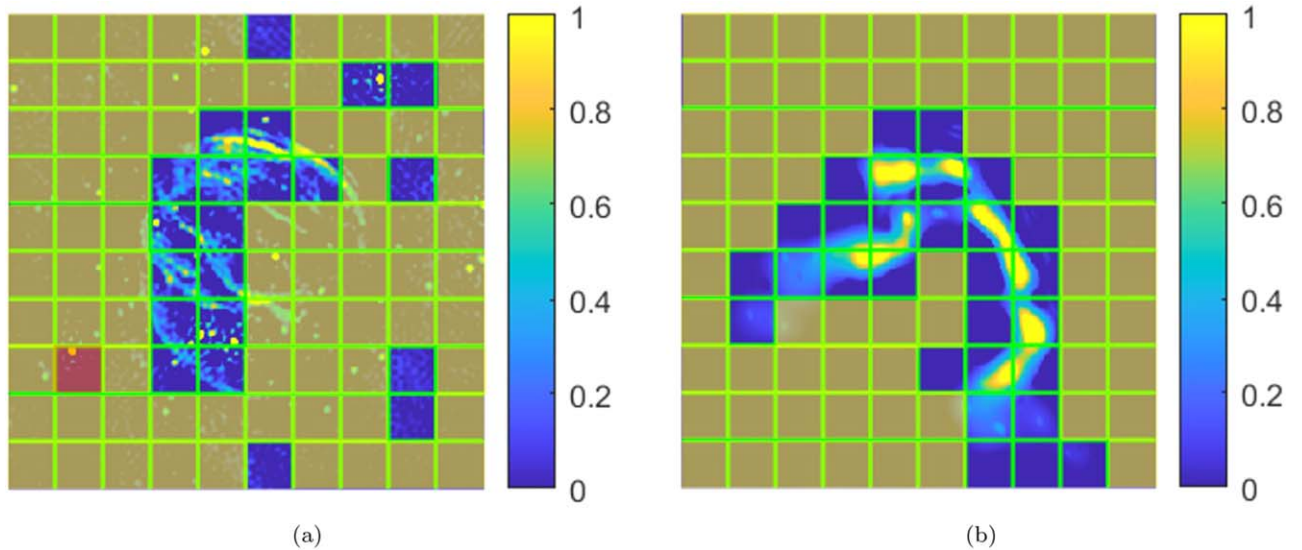


Figure 15. Conclusions of comparing the images generated by MS-CLEAN (in WSClean) and IUWT for (a) SNR_G55_10s and (b) 3C75. In this figure, only the images generated by IUWT are illustrated to show the cases of different tiles. According to the results of the joint analysis, in the yellow shadowed tiles the bright structures in the images generated by MS-CLEAN and IUWT are similar and the faint structures are similar or slightly different; in the nonshadowed tiles the bright structures in the images generated by MS-CLEAN and IUWT are similar, although the faint structures are very different; and in the orange shadowed tile the bright structures in the images generated by MS-CLEAN and IUWT are very different.

Table 5
SSIM and augLISI for Assessing Tile (7, 5), (8, 2), and (3, 7)

Tile Index	SSIM	augLISI
(7,5)	0.8169	0.9312
(8,2)	0.9729	0.9405
(3,7)	0.9878	0.9862

Therefore, SSIM and augLISI can work jointly on RA images with extended sources, to characterize the differences in different parts of the source(s) structure without using source extraction.

5. Conclusion

RA images of the same SBD are known to be very similar, especially when they are observed by the HVS. We propose a new IQA method augLISI in this article to assess the quality of extended source images. The IQA methods are applied to real data sets and simulated data sets in our experiments to illustrate how the IQA method can help radio astronomers with the study of RI imaging techniques.

AugLISI helps astronomers to compare images quantitatively, not only revealing the differences between different implementation approaches for the same imaging algorithm but also assessing the differences between different algorithms. To further characterize the similarity of different sections in the images, SSIM and augLISI can be used to assess the image quality jointly.

In this article, we focus on assessing the quality of RA images. However, our proposed IQA methods could also be applied to other astronomical areas of study. In the future work, our IQA methods may also contribute to automatic comparison of images for quality assessment, e.g., in Continuous Integration/Continuous Delivery (CI/CD) pipeline.⁸ Furthermore,

IQA methods can also help with transient search (Wright et al. 2017), e.g., supernova hunting,⁹ by comparing relevant images to look for transients without source extraction.

Acknowledgments

This work has received support from STFC grant (ST/W001969/1). The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The authors would like to thank Fred Dulwich and Ben Mort for valuable guidance on OSKAR. The authors would also like to thank Dan Walker, Hannah Stacey, Dirk Petry, and Katharina Immer from ALMA Science help center for helpful advice regarding CASA. The authors would also like to thank Miguel Cárcamo for beneficial discussion about GPUVMEM. The authors would like to thank Radostin Stoyanov for his kind help of installing GPUVMEM.

Appendix A ITW-SSIM

We define a weighted mean by weighting each pixel with its importance and summing over all pixels, expressed as

$$\bar{X}_N = \sum_{i=1}^N f(X_i) X_i, \quad (\text{A1})$$

where X is the value of a specific pixel, i is the index of the pixel (where the maximum value of i is N), and $f(X_i)$ is the weighting factor for the pixel at location i .

Also, the weighted variance and weighted cross covariance are expressed mathematically as

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (f(X_i) X_i - \bar{X}_N)^2 \quad (\text{A2})$$

⁸ <https://developer.skao.int/en/latest/tools/ci-cd.html>

⁹ <https://cs.lbl.gov/news-media/news/2015/supernova-hunting-with-supercomputers/>

and

$$\sigma_{XY} = \frac{1}{N-1} \sum_{i=1}^N (f(X_i)NX_i - \bar{X}_N)(f(Y_i)NY_i - \bar{Y}_N), \quad (\text{A3})$$

respectively, where X and Y are the values of a specific pixel at the same position of the two input images. It is known that the degree of freedom is $N-1$ in the original form of covariance.¹⁰ Hence, the degree of freedom should also be $N-1$ in the weighted form, the full derivation and proof are presented in Appendix B.

Reflecting in IQA, X and Y represent intensities of pixels of the two input images to be compared.

Thereby, the ITW-SSIMs (whose range should be $[-1, 1]$) can be expressed by

$$\text{ITW-SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (\text{A4a})$$

$$\begin{cases} \mu_x = \sum_{i=1}^N f(x_i)x_i \\ \sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (f(x_i)Nx_i - \mu_x)^2 \right)^{\frac{1}{2}} \\ \sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (f(x_i)Nx_i - \mu_x)(f(y_i)Ny_i - \mu_y) \end{cases} \quad (\text{A4b})$$

where the inputs x and y are normalized, N is the total number of pixels in the image, μ_y (or σ_y) is similar to μ_x (or σ_x), and the sum of weighting factors over a particular image should be 1, i.e., $\sum_{i=1}^N f(x_i) = 1$ and $\sum_{i=1}^N f(y_i) = 1$. The weighting functions of ITW-SSIMs can be selected by users for different applications. To weight higher intensity with higher importance, three weighting functions are implemented in ITW-SSIMs. They are Gaussian weighting, tanh weighting, and sigmoid weighting.

Appendix B

Proof of the Degree of Freedom in ITW-SSIMs

Proof: To prove the degree of freedom is $N-1$ in weighted variance, we bias of the estimator:

$$\begin{aligned} \text{Bias}(\hat{\sigma}^2) &= E \left[\frac{1}{N-1} \sum_{i=1}^N (f(X_i)NX_i - \bar{X}_N)^2 \right] - \sigma^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N E[(f(X_i)NX_i - \mu) + (\mu - \bar{X}_N)]^2 - \sigma^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N (E[(f(X_i)NX_i - \mu)^2] \\ &\quad + 2E[(f(X_i)NX_i - \mu)(\mu - \bar{X}_N)] \\ &\quad + E[(\mu - \bar{X}_N)^2]) - \sigma^2 \end{aligned}$$

$$\begin{aligned} &= \frac{1}{N-1} \sum_{i=1}^N \left(\sigma^2 - 2E \left[(f(X_i)NX_i - \mu) \left(\sum_{j=1}^N f(X_j)X_j - \mu \right) \right] \right. \\ &\quad \left. + E \left[\left(-N\frac{1}{N}\mu + \sum_{j=1}^N f(X_j)X_j \right)^2 \right] \right) - \sigma^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(\sigma^2 - 2E \left[\left(\sum_{j=1}^N (f(X_j)X_j - \mu + \mu) \right) \right. \right. \\ &\quad \left. \left. \times (f(X_i)NX_i - \mu) - \mu(f(X_i)NX_i - \mu) \right] \right. \\ &\quad \left. + E \left[\left(\sum_{j=1}^N \left(f(X_j)X_j - \frac{\mu}{N} \right) \right)^2 \right] \right) - \sigma^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(\sigma^2 - 2E \left[\sum_{j=1}^N ((f(X_j)X_j - \mu)(f(X_i)NX_i - \mu)) \right. \right. \\ &\quad \left. \left. + (N-1)\mu(f(X_i)NX_i - \mu) \right] + \frac{\sigma^2}{N} \right) - \sigma^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(\sigma^2 - \frac{2}{N} \sum_{j=1}^N E[(f(X_j)NX_j - \mu) \right. \\ &\quad \left. \times (f(X_i)NX_i - \mu) - (N-1)\mu(f(X_i)NX_i - \mu)] \right. \\ &\quad \left. - 2E[(N-1)\mu(f(X_i)NX_i - \mu)] + \frac{\sigma^2}{N} \right) - \sigma^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(\sigma^2 - \frac{2}{N} \sum_{j=1}^N E[(f(X_j)NX_j - \mu)(f(X_i)NX_i - \mu)] \right. \\ &\quad \left. + \frac{2}{N} \sum_{j=1}^N E[(N-1)\mu(f(X_i)NX_i - \mu)] \right. \\ &\quad \left. - 2E[(N-1)\mu(f(X_i)NX_i - \mu)] + \frac{\sigma^2}{N} \right) - \sigma^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(\sigma^2 - \frac{2}{N} E[(f(X_i)NX_i - \mu)^2] + \frac{\sigma^2}{N} \right) - \sigma^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(\sigma^2 - \frac{2}{N} \sigma^2 + \frac{\sigma^2}{N} \right) - \sigma^2 \\ &= \frac{N}{N-1} \frac{N-1}{N} \sigma^2 - \sigma^2 = 0. \end{aligned}$$

Because the bias equals to 0, the degree of freedom of weighted variance should be $N-1$.

The proof for weighted cross covariance is similar.

Q.E.D.

ORCID iDs

Karel Adámek  <https://orcid.org/0000-0003-2797-0595>
Wesley Armour  <https://orcid.org/0000-0003-1756-3064>

References

- Ables, J. 1974, A&AS, **15**, 383
Asano, F., & Asoh, H. 2011, 2011 19th European Signal Processing Conf. (Piscataway, NJ: IEEE), 2009, <https://ieeexplore.ieee.org/document/7073859>
Bakurov, I., Buzzelli, M., Schettini, R., Castelli, M., & Vanneschi, L. 2022, *Expert Syst. Appl.*, 189, 116087

¹⁰ <https://davidmlane.com/hyperstat/A42408.html>

- Briggs, D. S. 1995, PhD thesis, The New Mexico Institute of Mining and Technology
- Brunet, D. 2012, PhD thesis, Univ. of Waterloo, <http://hdl.handle.net/10012/6982>
- Burke, B., Graham-Smith, F., & Wilkinson, P. 2019, *An Introduction to Radio Astronomy* (4th ed.; Cambridge: Cambridge Univ. Press)
- Carrillo, R., McEwen, J., & Wiaux, Y. 2014, *MNRAS*, **439**, 3591
- Chen, G., Yang, C., & Xie, S. 2006, in 2006 Int. Conf. on Image Processing (Piscataway, NJ: IEEE), 2929
- Cornwell, T. 2008, *ISTSP*, **2**, 793
- Cornwell, T., & Evans, K. 1985, *A&A*, **143**, 77
- Cárcamo, M., Román, P., Casassus, S., Moral, V., & Rannou, F. 2018, *A&C*, **22**, 16
- Cárcamo, M., Scaife, A. M. M., Alexander, E. L., & Leahy, J. P. 2022, *MNRAS*, **518**, 1955
- Dabbech, A., Ferrari, C., Mary, D., et al. 2015, *A&A*, **576**, A7
- Dewdney, P., Hall, P., Schilizzi, R., & Lazio, J. 2009, *IEEEP*, **97**, 1482
- Ding, K., Ma, K., Wang, S., & Simoncelli, E. P. 2021, *Int. J. Comput. Vis.*, **129**, 1258
- Dulwich, F., 2020 OSKAR 2.7.6, v2.7.6, Zenodo, doi:10.5281/zenodo.3758491
- Fornasier, M., & Rauhut, H. 2011, in *Handbook of Mathematical Methods in Imaging*, ed. O. Scherzer (Berlin: Springer), 187
- Frieden, B. 1972, *OSAJ*, **62**, 511
- Garsden, H., Girard, J., Starck, J., et al. 2015, *A&A*, **575**, A90
- Gull, S., & Daniell, G. 1978, *Natur*, **272**, 686
- Hardy, S. 2013, *A&A*, **557**, A134
- Hogbom, J. 1974, *A&AS*, **15**, 417
- Kaghaz-Garan, S., Umbarkar, A., & Doboli, A. 2014, in 2014 IEEE Int. Symp. on Robotic and Sensors Environments (ROSE) Proc. (Piscataway, NJ: IEEE), 49
- Li, F., Brown, S., Cornwell, T., & de Hoog, F. 2011a, *A&A*, **531**, A126
- Li, F., Cornwell, T., & de Hoog, F. 2011b, *A&A*, **528**, A31
- Li, X., Adamek, K., & Armour, W., 2024 Intensity Sensitive IQAs, v1.0.0, Zenodo, doi:10.5281/zenodo.10863656
- Li, X., & Armour, W. 2022, in 2022 26th Int. Conf. on Pattern Recognition (ICPR) (Piscataway, NJ: IEEE), 1975
- Lu, Y.-H. C., Hodosan, G., Wang, F., Daley-Yates, S., & Cornwell, T. 2022 in 2022 3rd URSI Atlantic and Asia Pacific Radio Science Meeting (AT-AP-RASC) (Piscataway, NJ: IEEE)
- McMullin, J., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in ASP Conf. Ser. 376, *Astronomical Data Analysis Software and Systems XVI*, ed. R. Shaw, F. Hill, & D. Bell (San Francisco, CA: ASP), 127
- Moorthy, A., & Bovik, A. 2009, *ISTSP*, **3**, 193
- Narayan, R., & Nityananda, R. 1984, in *Indirect Imaging*, ed. J. Roberts (Cambridge: Cambridge University Press), 281
- Offringa, A. R., McKinley, B., Hurley-Walker, et al. 2014, *MNRAS*, **444**, 606
- Offringa, A. R., & Smirnov, O. 2017, *MNRAS*, **471**, 301
- Oppenheim, A., & Lim, J. 1981, *IEEEP*, **69**, 529
- Ponsonby, J. 1973, *MNRAS*, **163**, 369
- Pratley, L., McEwen, J., d'Avezac, M., et al. 2018, *MNRAS*, **473**, 1038
- Sampat, M., Wang, Z., Gupta, S., Bovik, A., & Markey, M. 2009, *ITIP*, **18**, 2385
- Sun, X., Landecker, T., Gaensler, B., et al. 2015, *ApJ*, **811**, 40
- Teeninga, P., Moschini, U., Trager, S. C., & Wilkinson, M. H. 2016, *Math. Morphol. Theory Appl.*, **1**, 100
- The CASA Team, Bean, B., Bhatnagar, S., et al. 2022, *PASP*, **134**, 114501
- Thompson, A. 1999, in ASP Conf. Ser. 180, *Synthesis Imaging in Radio Astronomy II*, ed. G. Taylor, C. Carilli, & R. Perley (San Francisco, CA: ASP), 11
- Thompson, A., Moran, J., & Swenson, G. 2017, *Interferometry and Synthesis in Radio Astronomy* (3rd ed.; Oxford: Wiley)
- Van der Tol, S., Veenboer, B., & Offringa, A. R. 2018, *A&A*, **616**, A27
- Wang, Z., Bovik, A., & Sheikh, H. 2005, in *Digital Video Image Quality and Perceptual Coding*, ed. H. Wu & K. Rao (Boca Raton: CRC Press),
- Wang, Z., Bovik, A., Sheikh, H., & Simoncelli, E. 2004, *ITIP*, **13**, 600
- Wang, Z., & Li, Q. 2011, *ITIP*, **20**, 1185
- Wang, Z., Simoncelli, E., & Bovik, A. 2003, in *The Thirtieth-Seventh Asilomar Conf. on Signals, Systems Computers* (Piscataway, NJ: IEEE), 1398
- Wenger, S., Magnor, M., Pihlstrom, Y., Bhatnagar, S., & Rau, U. 2010, *PASP*, **122**, 1367
- Wernecke, S., & d'Addario, L. 1977, *ITCom*, **C-26**, 351
- Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M. M., Vanderghenst, P., et al. 2009, *MNRAS*, **395**, 1733
- Wright, D. E., Lintott, C. J., Smartt, S. J., et al. 2017, *MNRAS*, **472**, 1315
- Ye, H., Gull, S., Tan, S., & Nikolic, B. 2022, *MNRAS*, **510**, 4110
- Zhu, M., Liu, W., Wang, B. Y., et al. 2019, *AdAst*, 2019, 2397536
- Zhu, Y., Huang, T., Hou, J., Du, S., & Song, S. 2020, in *Big Data in Astronomy*, ed. L. Kong et al. (Amsterdam: Elsevier), 271