Draft version July 31, 2018
Typeset using LaTeX **manuscript** style in AASTeX62

RICA - Radio Imaging Combination Analyzer

Miles Lucas[1,2]

[1] *Iowa State University, Ames, Iowa*

[2] *National Radio Astronomy Observatory, Socorro, New Mexico*

ABSTRACT

1. INTRODUCTION

In radio synthesis imaging a common problem arising from interferometry is the lack of zero-spacing data. Without this data, images lack total power information. Single dish radio telescopes retain this total power information but lack the angular resolution capabilities of radio interferometers. To solve this problem, astronomers combine the interferometry data with the total power data. This report seeks to characterize the effectiveness of different combination methods and parameters.

The parameters analyzed were multiscale deconvolution versus normal Hogbom deconvolution, the effect of single-dish size on the combination, masking, and number of total iterations (i.e. depth of clean). One of the previous assumptions is that if the single-dish UV-coverage does not adequately complement the interferometer UV-coverage, feathering will not be effective. Another assumption is that simple point structures should not see much discrepancy in parameters or methods. Therefore, more complex, fluffy structure ought to show more distinction in results.

There are four methods of combination that were tested. The first is CASA's *feather* task. The process of feathering starts with regridding the lower resolution to the higher resolution image. Each is then Fourier transformed and gridded, while the lower resolution data is scaled by the ratio of the clean beams (high resolution/low resolution). The high-resolution grid is then added to this, scaled by a weighting that starts at 0 and increases to 1 as the wavelength increases. This final combination is then inverse Fourier transformed back into the image plane.

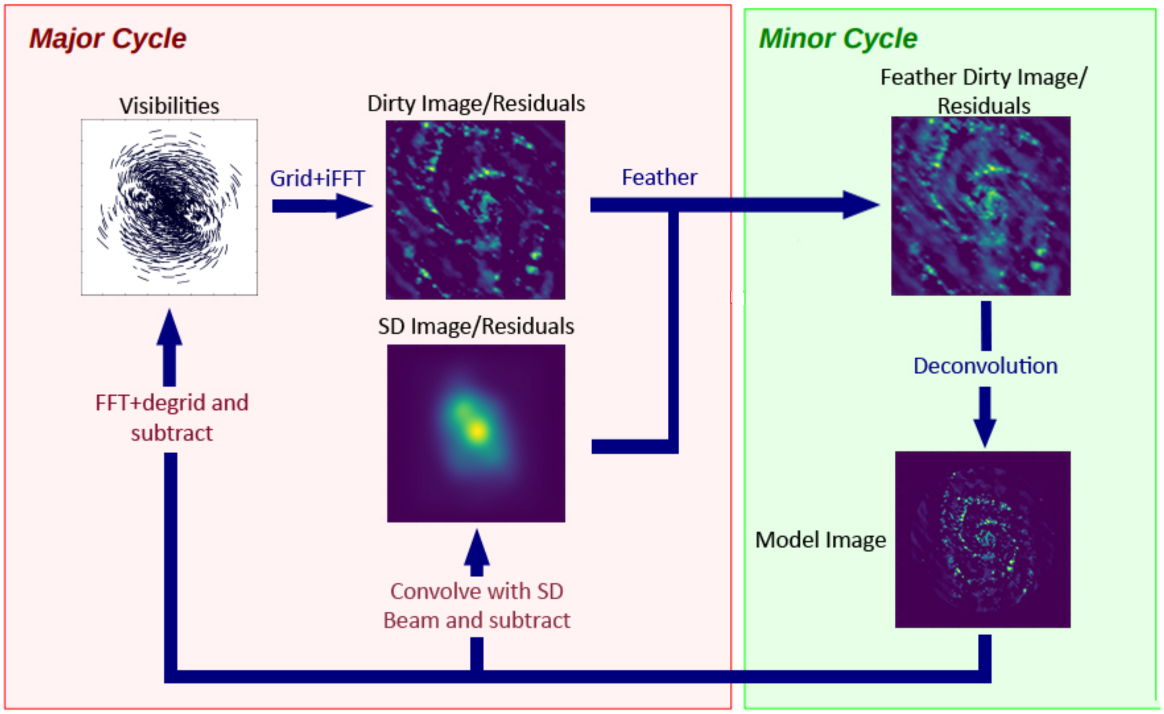

Figure 1. A flow chart describing the joint-deconvolution approach to combining interferometer and single dish data. Adapted from Rau, U. & Naik, N. 2018 (in prep).

The next method is using the total power image as the starting model in CASA's *tclean* task. In the Cotton-Schwab CLEAN algorithm (CSCLEAN), each major-cycle begins with a blank model image and constructs an image on it by using minor-cycle iterations. By using a starting model, this model is no longer blank, but is passed as a parameter. By using the total power image as the starting model, we can hope to retain the total power information through the CLEAN algorithm.

Another modification to the Cotton-Schwab cycle in order to combine images is by doing a modified joint deconvolution. In each major Cotton-Schwab cycle, the visibilities are gridded and inverse Fourier transformed to create an image residual, which is then feathered with the total power residual before deconvolving. The deconvolved model is then convolved with the single-dish beam and subtracted from the previous single-dish residual. The deconvolved model is also Fourier transformed, degridded, and subtracted from the visibilities to create a new residual image These residuals are then degridded and transformed to finish the major cycle. See Figure 1 for a visual flowchart of the process.

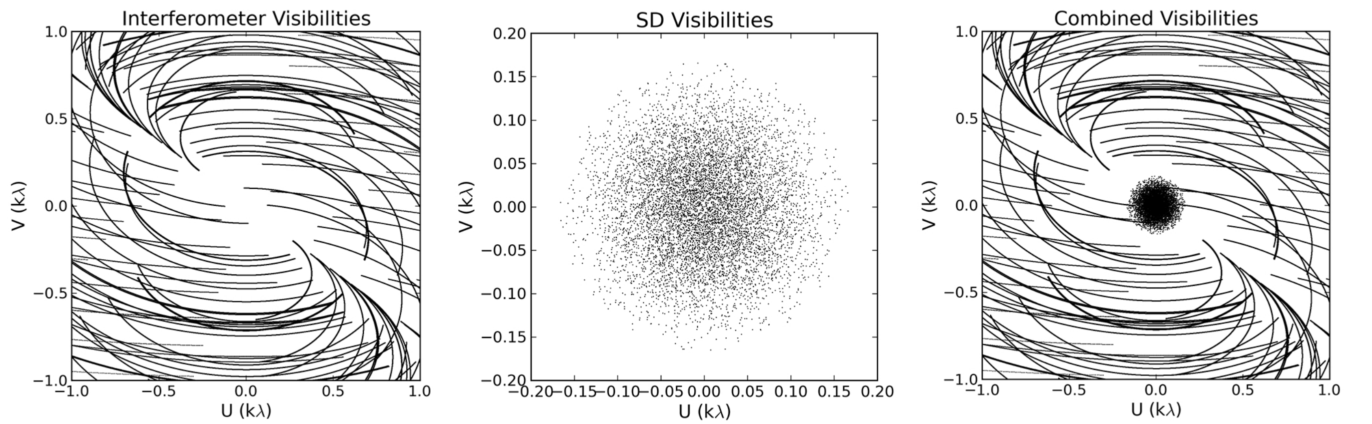

Figure 2. A look at how the visibilities are added through using tp2vis for a model simulated on the VLA A configuration. In the center of the interferometer image in the first plot shows sparsity that is filled in on the third plot. The second plot shows the approximation tp2vis makes of the single dish.

The last method of combination we use is *tp2vis*, which takes the total power image and spoofs a measurement set. The way *tp2vis* accomplishes this is by taking Fourier data from the total power image and sparsely sampling the Fourier data at close spacings (see Koda et al. 2011). This data can be concatenated with the interferometer data and deconvolved in *tclean*. Figure 2 shows how the single-dish is simulated and concatenated with the interferometer data. This should help fill the close-spacing gap created by the interferometer data and subsequently help increase total power information.

2. METHODS

In order to evaluate the different combination methods, a metric needed to be created and tested on a suite of models. The metrics used were CLEAN residuals, fidelity images, and a ratio of the power spectrum densities (PSD). The fidelities are given Equation 1

$$\text{Fidelity}(i,j) = \frac{|\text{Model}(i,j)|}{\max\left(|\text{Difference}(i,j)|, 0.7 \times \text{rms}(\text{Difference})\right)} \tag{1}$$

where model is the reference image and diff is the difference between the test and reference images (Pety et al. 2011, p.19). The exigence for using PSDs is that the zero-spacing power is readily visible for every image. By using these ratios, the closer the ratio is to 1.0 at short spacings (when compared to the true model), the more accurate the combination. In addition, when compared with the total power image, it shows how much effective weight is given to the total power image in the combination. The ratio of the PSDs is given by Equation 2

$$\mathrm{Ratio}(UV) = \frac{\mathrm{Power}_{test}(UV)}{\mathrm{Power}_{ref}(UV)} \cdot \frac{\mathrm{BA}_{ref}}{\mathrm{BA}_{test}} \tag{2}$$

$$\mathrm{err}(UV) = \left(\frac{\mathrm{Power}_{ref}(UV) + \mathrm{Power}_{test}(UV)}{2}\right)^{-1} \tag{3}$$

where Power is the power from the PSD, BA is the beam area, ref refers to the reference image (model or SD), and $test$ refers to whichever image is being tested. The uncertainty given is Equation 3 such that the uncertainty is the reciprocal of the average of the PSDs.

Each combination was compared to both the true model and the single dish, total power image. Many models were tested with various extra parameters. Three models were generated from component lists. One has 4 point sources, one has a single Gaussian source, and one has a mixture of 4 point sources, one very broad Gaussian, and on off-center, stronger Gaussian. There were also various models based off real structure, including M51 (based off an H-α image), Orion, RXJ1347, and a protoplanetary disk (PPD) simulation. Figure 3 shows all of the models. It is important to note that these models have been regridded onto a common coordinate system that is not representative of the true astronomical targets. This was in effort to simplify the simulation process for creating measurement sets.

The code base for testing the effectiveness is hosted publicly for anyone to use[1]. In the source code there are many scripts and methods to facilitate simulating, combining, and comparing. Every model is described in *src/_models.py* with a dictionary defining the simulation and *tclean* parameters. The pipeline for testing these models was as follows:

1. Simulate measurement set (MS) based on VLA configurations

[1] https://gitlab.com/mileslucas/rica

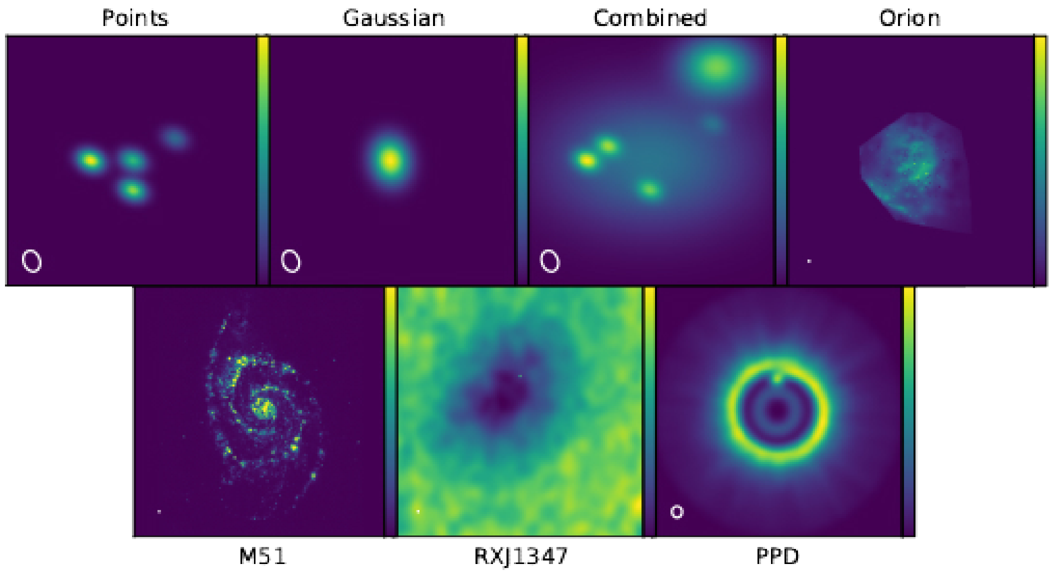

Figure 3. The models used for testing combination methods convolved with some restoring beam to show structure better.

2. Simulate single dish image by convolving model with Gaussian
3. Image the MS
4. Do the combinations (feather, startmodel, joint deconvolution, and tp2vis)
5. For each combination, do a comparison with the true model (convolved with the restoring beam) and the single dish image

and is implemented in *src/pipeline.py*.

2.1. *Simulations and Deconvolution*

All simulations were done using the VLA for the telescope model except for the PPD simulation, which was done using *simalma* and was only simulated once. Each test model could define which of the VLA configurations to use, from A, B, BnA, C, CnB, D, and DnC. The spectral window for the simulations mimic the C-band at 4.5 GHz to 5.5 GHz (central 5 GHz) with 101 channels at 10 MHz intervals. The field observed was J2000 21h28m31.0 45°0′0.0″. The observation date and time was 2018/06/01 at 12:00:00.0 UTC. The integration time was 10 s each with a total of 30 000 s. Finally, the data was corrupted with 1 mJy of simple noise.

The single dish images were created by taking the true model and convolving with a Gaussian equivalent to the single dish beam at the given frequency.

$$FWHM(\nu, D) = \frac{3.66 \times 10^8\,\mathrm{m\,s^{-1}}}{\nu \cdot D} \tag{4}$$

Equation 4 gives the Gaussian kernel full-width half-maximum (FWHM) in radians for a frequency in Hz and dish diameter in meters. For instance, to simulate data from the Green Bank Telescope (GBT) with a dish diameter of 100 m at 5 GHz gives a FWHM of $2'31''$, which is in agreeance with values from the GBT proposer's guide[2]. Note that this is not the most accurate equation for other telescopes, because there is a constant involved in this equation based off the taper-length of the telescope's feedhorns. In addition, these single dish images have 5×10^{-6} uniform noise added.

Each model defines its own parameters for cleaning. Appendix A lists every single parameter for each model. As a default, a model simulated in VLA D configuration has an image size of 64 with a cell size of $7.0''$. The restoring beam for the D configuration is approximately $20.41''$ by $15.60''$ so the cell accounts for half to a third of the beam width. In general, cleaning was done with very high numbers of iterations, relying on a threshold to stop the algorithm for consistency. This differs for the models tested at explicitly different cleaned levels.

2.2. *Comparisons*

Each comparison was ran with the same clean parameters for consistency. Because *feather* does not actively clean, it has its own set of parameters, but all of the models tested used the default parameters. After all the combination permutations were imaged, they were compared against the true model convolved with the restoring beam and the single dish image alongside the uncombined interferometer image. All of the ratio were saved into individual CSV files for further analysis. In all, each model has 10 comparisons (5 against the models and 5 against the single-dish images) and 4 separate cleans to perform.

[2] p.9, https://science.nrao.edu/facilities/gbt/proposing/GBTpg.pdf

In order to run a new model to test comparison methods, there are two general methods. If there exists a true model, simply add an entry to the models dictionary defining the simulation and clean parameters. It is also important to edit *src/simulate.py* to accommodate the new model. Follow the existing code base for guidance. It is also important to create a copy of the model and regrid it to the common coordinate system (use an existing model to snag the coordinate system). The model can then be ran through the pipeline along with any other models using *src/pipeline.py*.

If there is no true model, it is still useful to test and compare against the total power image, alone. In this case, using a cleaned image, a measurement set, and a total power image with the *src/combine.py* script will produce all of the combinations. These can then be compared using *src/compare.py*. It is possible to edit *src/pipeline.py* to accommodate special models, see how the PPD model is handled in the pipeline.

3. RESULTS

4. CONCLUSION

This work was funded by the National Science Foundation in partnership with the National Radio Astronomy Observatory and Associated Universities Incorporated. Thank you to Dr. Kumar Golap and Dr. Takahiro Tsutsumi for their guidance and assistance.

Software: CASA, tp2vis (Koda et al. 2011)

REFERENCES

Koda, J., Sawada, T., Wright, M. C. H., et al. 2011, ApJS, 193, 19, doi: 10.1088/0067-0049/193/1/19

Pety, J., Gueth, F., & Guilloteau, S. 2011, ALMA Memo #398 Impact of ACA on the Wide-Field Imaging Capabilities of ALMA, Tech. rep.

APPENDIX

A. MODEL PARAMETERS

Table 1. The parameters for every model ran through the pipeline

Name	Model	Config	SD size	imsize	cell	niter	cycleniter	threshold	deconvolver	scales	mask
points	points	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	hogbom		No
gauss	gauss	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	hogbom		No
combined	combined	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	hogbom		No
combined-n0	combined	VLA d	100 m	[64, 64]	7.00arcsec	0	100	1e-5Jy	hogbom		No
combined-n100	combined	VLA d	100 m	[64, 64]	7.00arcsec	100	100	1e-5Jy	hogbom		No
combined-n1000	combined	VLA d	100 m	[64, 64]	7.00arcsec	1000	100	1e-5Jy	hogbom		No
combined-ms	combined	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	multiscale	[0, 3, 10, 30]	No
m51	m51	VLA d	100 m	[128, 128]	7.00arcsec	100000	100	1e-12Jy	hogbom		No
m51-ms	m51	VLA d	100 m	[128, 128]	7.00arcsec	100000	100	1e-12Jy	multiscale	[0, 3, 10, 30]	No
m51-b	m51	VLA b	100 m	[1400, 1400]	0.65arcsec	100000	100	1e-12Jy	hogbom		No
m51-b-ms	m51	VLA b	100 m	[1400, 1400]	0.65arcsec	100000	100	1e-12Jy	multiscale	[0, 3, 10, 30]	No
orion-b	orion	VLA b	100 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	hogbom		Yes
orion-b-n0	orion	VLA b	100 m	[700, 700]	0.65arcsec	0	100	1e-9Jy	hogbom		Yes
orion-b-n100	orion	VLA b	100 m	[700, 700]	0.65arcsec	100	100	1e-9Jy	hogbom		Yes

Table 1 continued on next page

Table 1 *(continued)*

Name	Model	Config	SD size	imsize	cell	niter	cycleniter	threshold	deconvolver	scales	mask
orion-b-n1000	orion	VLA b	100 m	[700, 700]	0.65arcsec	1000	100	1e-9Jy	hogbom		Yes
orion-b-ms	orion	VLA b	100 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	multiscale	[0, 3, 10, 50]	Yes
orion-b-25	orion	VLA b	25 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	hogbom		Yes
orion-b-25-n0	orion	VLA b	25 m	[700, 700]	0.65arcsec	0	100	1e-9Jy	hogbom		Yes
orion-b-25-n100	orion	VLA b	25 m	[700, 700]	0.65arcsec	100	100	1e-9Jy	hogbom		Yes
orion-b-25-n1000	orion	VLA b	25 m	[700, 700]	0.65arcsec	1000	100	1e-9Jy	hogbom		Yes
orion-b-25-ms	orion	VLA b	25 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	multiscale	[0, 3, 10, 30]	Yes
orion-c	orion	VLA c	100 m	[216, 216]	2.12arcsec	100000	400	1e-9Jy	hogbom		No
RXJ1347	RXJ1347	VLA a	100 m	[2268, 2268]	0.20arcsec	100000	100	1e-6Jy	hogbom		No
RXJ1347-masked	RXJ1347	VLA a	100 m	[2268, 2268]	0.20arcsec	100000	100	1e-6Jy	hogbom		Yes
ppd	ppd	ALMA	ACA	[192, 192]	0.01arcsec	100000	100	1e-7Jy	hogbom		No

B. MODEL RESULTS

Fig. Set 1. Model Results

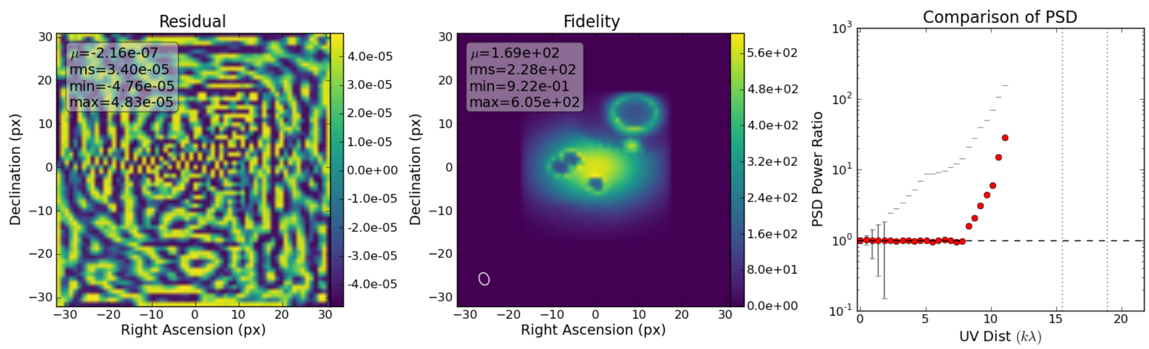

Figure 4. Comparison of feather combination to the true model for combined simulation. The ratio is Equation 2. The black ticks are error bars on a scaled uncertainty based on Equation 3. The vertical lines are the maximum UV distance of the interferometer at the longest and shortest wavelengths, respectively. The complete figure set (125 figures) is available online