P7 Compressed Sensing Image Reconstruction for CASA

Jonas Schwammberger

December 17, 2018

Abstract

abstract plcaceholder

Contents

1 Large Scale Image Reconstruction for the MeerKAT Telescope **1**
1.1 The Major Cycle Architecture . . . 1
1.2 Compressed Sensing Reconstructions . . . 2

2 Challenges for imaging MeerKAT data **3**
2.1 Wide Field of View Imaging . . . 3
2.2 State of the Art: Distributing the W-Term with WSCLEAN . . . 4

3 Eliminating the Major Cycle **5**
3.1 Non-uniform FFT as Optimization Problem . . . 5
3.2 Spherical Harmonics . . . 5
3.3 UV-Smooth . . . 5
3.4 Coordinate Descent . . . 5

4 Coordinate Descent for MeerKAT Data **7**
4.1 The Starlet Regularization . . . 7
4.2 Active set heuristic with Starlets . . . 7
4.3 Results on Simulated Data . . . 7
4.3.1 Point Sources . . . 7
4.3.2 Gaussian and Point Sources . . . 7
4.4 Scalability estimates with ideal heuristics . . . 7

5 Conclusion **9**

6 Ehrlichkeitserklärung **12**

1 Large Scale Image Reconstruction for the MeerKAT Telescope

Image reconstruction on the large scale

Large scale image reconstruction applied to data from the MeerKAT Telescope.

$$V(u,v) = \int\int x(x,y)e^{2\pi i(ux+vy)}\,dx\,dy \tag{1.1}$$

sampling pattern is defined by the instrument, in MeerKAT's case the sampling Visibilities are:

1. Not Uniform: Pockets of dense samples.
2. Incomplete: Information for the Image Reconstruction was not measured by the instrument

1.1 The Major Cycle Architecture

Algorithms in Radio Astronomy employ the major cycle architecture. The figure shows the major cycle framework. In a major cycle consists of two parts: The non-uniform FFT and an optimization algorithm.

The non-uniform FFT is responsible for approximating a regularly spaced image from the measurements, and for approximating the measurements corresponding to an image. Non-uniform FFT's are fast approximation algorithms, but by being approximation algorithms, they introduce errors.

A optimization algorithm uses the image and removes the effect of incomplete samples. In the past, CLEAN algorithms were used to remove the effects. For the future, algorithms based on the theory of compressed sensing show promises in image quality.

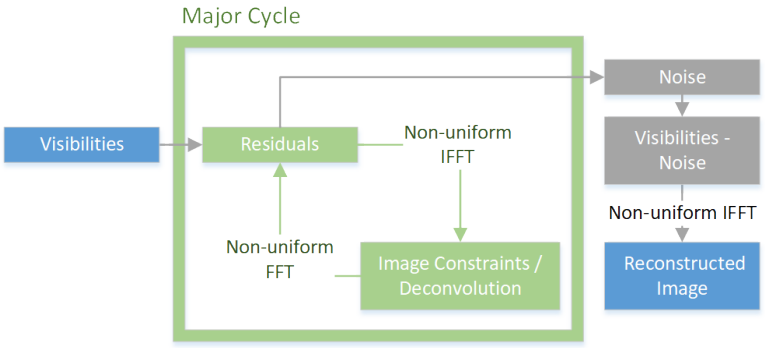

Figure 1: The Major Cycle Framework

A full major cycle consists of the following operations: First, it approximates the regularly spaced the regularly spaced image from the measurements. Then the optimization algorithm removes the effects of Incomplete measurements and returns the corresponding image. The major cycle then approximates the measurements corresponding to the image with the non-uniform FFT. The residual measurements are used in the next Major Cycle. In each cycle, two errors get simultaneously reduced:

1. The Error introduced by the non-uniform FFT.
2. The Error introduced by the incomplete measurements.

After several major cycles the algorithm converges on a regularly spaced image which has a small error from non-uniform samples, and a small error from incomplete measurements.

For the optimization algorithm, the CLEAN class of algorithms get used. But algorithms based on the theory of compressed sensing have been shown to produce superior images.

1.2 Compressed Sensing Reconstructions

However, compressed Sensing algorithms come with the drawback of requiring more major cycles.

Current Compressed Sensing reconstructions reduce the number of major cycles. However, the question is if Compressed Sensing can use a different architecture, and scale better to problems of the size of MeerKAT.

There are ways to get rid of the major cycle, but overall the complexity could not be reduced.

2 Challenges for imaging MeerKAT data

There are several challenger for imaging meerkat data. One problem is the new amount of data.

terabytes of measurements. Large image size 32k squared are the obvious problems to solve. Distributing the problem is not part of this work.

In this work, it is focused on Wide field of view issue.

Calibration gets not explicitly called, but

Further issues that do not get handled here

- (Beam Pattern, A Projection)
- Full polarization
- Wide band imaging

2.1 Wide Field of View Imaging

So far the small Field of View inverse problem has been introduced where each antenna pair measures a Visibility of the sky brightness distribution. This leads to the small Field of View measurement equation (2.1). It is identical to the two dimensional Fourier Transform. In practice the Fast Fourier Transform (FFT) is used, since it scales with $n\,log(n)$ instead of n^2 pixels.

$$V(u,v) = \int\int x(l,m)e^{2\pi i(ux+vy)}dldm \tag{2.1}$$

For wide Field of View imaging, two effects break the two dimensional Fourier Transform relationship: Non-coplanar Baselines and the celestial sphere which lead to the measurement equation (**??**). Note that for small Field of View $1-x^2-y^2 \ll 1$, and (**??**) reduces to the 2d measurement equation (2.1).

$$V(u,v,w) = \int\int \frac{X(x,y)}{\sqrt{1-x^2-y^2}} e^{2\pi i(ux+vy+w\sqrt{1-x^2-y^2})}dxdy \tag{2.2}$$

Non-coplanar Baselines lead to a third component w for each Visibility. Figure 2 shows the the u v and w coordinate system. w is essentially the pointing direction of the instrument. The UV-Plane is the projection of the antennas on a plane perpendicular to the pointing direction. Which point in the UV-Plane get sampled and what w component it has depends on the pointing direction. If the instrument points straight up, the UV-Plane is a tangent to earth's surface, and the w term compensates for earth's surface curvature. If however the instrument points at the horizon, the projected UV-Plane gets squashed and w compensates for antennas which lie far behind the UV-Plane. In essence, w is a phase delay that corrects antenna positions in three dimensions. The wide Field of View measurement equation (**??**)

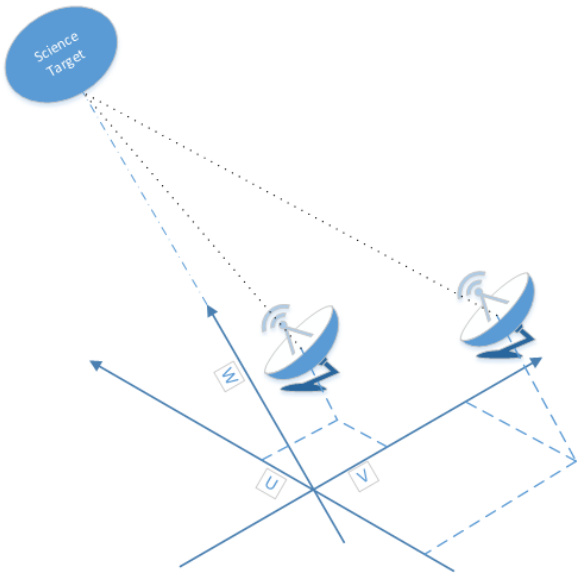

Figure 2: U V and W coordinate space

would account for the w phase delay, but it breaks the the two dimensional Fourier relationship and the FFT cannot be used. The W-Projection [**?**] algorithm approximates the effect of the w term restores the two dimensional Fourier relationship.

2.2 State of the Art: Distributing the W-Term with WSCLEAN

Tackling the problem State of the Art:Distributing the W-Term with WSCLEAN

Spherical Harmonics

FFT

uv smooth

Coordinate Descent

A-Projection [**?**]

3 Eliminating the Major Cycle

There are some ways of potentially eliminating the major cycle.

3.1 Non-uniform FFT as Optimization Problem

The major cycle minimizes two errors simultaneously: The error introduced by the non-uniform FFT and the error introduced by incomplete measurements. The idea is to separate the two errors, and solve for each separately. First, we minimize the objective (3.1), which searches for the optimal image x given the non-uniformly sampled measurements and then remove the effects of incomplete measurements.

$$\underset{x}{minimize} \ \|V - Ax\|_2^2 \tag{3.1}$$

Two possible upsides. A specialized optimization algorithm can be used for each sub-problem, hopefully converging faster than the major cycle. and because the image x is magnitudes smaller than the measurements, it is easier to handle downstream, throwing away the measurements.

However, the problem is that the imaging algorithm should also be able to calibrate. With this approach, there is a strict one way flow from measurements to image. For calibration, a backwards flow from image to measurements is necessary.

For calibrated data, this approach may be potentially faster. but since MeerKAT will be difficult to calibrate correctly, this is unlikely.

3.2 Spherical Harmonics

The signal naturally lives on the celestial sphere. The wide field of View measurement equation (2.2) can be factorized into spherical harmonics. This means spherical harmonics are equivalent, calculating an image from the measurements with the three dimensional fourier transform is the same as using the spherical harmonics transform.

There exist fast Spherical harmonics transforms, but replacing the nuFFT with W-Correction does not gain a speed advantage out of the box. It uses a non-uniform FFT as part of the algorithm.

Spherical harmonics were researched in the context of Compressed Sensing image reconstructions (cite wiaux). They improved the quality of the image reconstruction, but no speed advantage.

The push on solving on the sphere directly. Using spherical haar wavelets, (cite) was able to speed up the simulation problem. Simulation tries to solve to problem of finding the measurements corresponding to a given image x. However, this has so far not been used for imaging.

3.3 UV-Smooth

3.4 Coordinate Descent

Coordinate descent is an optimization algorithm, which can be interesting on LASSO synthesis objectives (3.2). The image $x = D\alpha$ and F^{-1} stands for the three dimensional inverse Fourier transform. It is interesting

because it does not need a major cycle, and can handle the w term without any extra operation.

$$\underset{\alpha}{minimize} \left\| V - F^{-1} D\alpha \right\|_2^2 + \lambda \left\| \alpha \right\|_1 \tag{3.2}$$

Coordinate descent minimizes the objective by descending one coordinate at a time. At a given point, all α are fixed except for one, which gets minimized. If the problem has the LASSO form, it forms a parabola which can be minimized analytically.

The Matrix Product $F^{-1} * D$ has the dimensionality $M * S$, where M is the number of measurements and S is the number of components in the dictionary. Since compressed sensing gets used with over-complete dictionaries, S is larger than the number of pixels N. For MeerKAT scale data, the matrix product is too big to calculate explicitly. The good news is that only small number of α are non-zero, so if there is a heuristic to find candidate α's we only need to calculate a subset of $F^{-1} * D$. It turns out, with the Starlet Basis one such heuristic exists.

The question is, can the Coordinate Descent with exact transformations out-perform the major cycle approaches, which uses approximations.

4 Coordinate Descent for MeerKAT Data

The basic algorithm

Full algorithm with starlets

bunch of heuristics,

4.1 The Starlet Regularization

Starlet is a multi-scale wavelet representation which were specifically developed for astronomy.

Over-complete representation. More starlets than there are pixels. Sparse representation, the number of non-zero starlets is smaller than the number of pixels

Starlets as a series of convolutions.

Forward transform, from image to starlets

From Starlets to image

4.2 Active set heuristic with Starlets

Even though the starlets are an over-complete dictionary, they have an approximate transform from image to starlet space. For Coordinate Descent, this can be used as a active set heuristic: We try to find the coefficients which are likely to be non-zero. This helps us so we do not need to calculate the whole matrix product $F^{-1}D$. We only use columns that are likely to be not zero.

(Image of the starlet level zero) The higher the number, the more likely this component is to be non-zero. It is essentially a probability distribution for which starlet components are non-zero.

Stupid approach with line search. Could be done more efficiently by using the histogram of the starlet level.

4.3 Results on Simulated Data

4.3.1 Point Sources

4.3.2 Gaussian and Point Sources

4.4 Scalability estimates with ideal heuristics

Bunch of heuristics. There are a lot of little ways to optimize this algorithm. The question is, is it worth going further and try to improve the algorithm further. So we want to estimate the lower bound of the algorithm.

Convergence is hard to determine in general. Following simplified assumptions were made: We have a heuristic with oracle performance. It returns only the locations of α in constant time. Furthermore, we assume all axes are independent from each other, only one descent per non-zero axis is necessary.

S

$res * starlet = M$ descent: $genfcol = 3 * M\ a = 3 * M\ b = 4 * M\ residuals, fcol * diff = M$

$total_m it_g encol = 11M$

total CD = $S * 11M + 2M * starlets$

S depends on the image content directly. For example if the image contains 15 point sources and five Gaussian extended emissions, then S equals 20 non-zero components (if we assume the Gaussian sources require only one starlet for representation). Coordinate Descent therefore is independent of the image size N. It solely depends on the size of the measurements M, and the number of non-zero components in the dictionary S.

It does not use any approximation for the fourier transform.

(Cite New W-stacking approach) Nufft: $M + 2Nlog2N$ W-stacking = $M + W * (2Nlog2N + 2N) + NlogN$ Deconvolution = ??

The major cycle algorithm depends on more parameters. Assumptions were made in favor of Coordinate Descent.

the number of w-stacks W. If this is the case

5 Conclusion

References

[1] National Radio Astronomy Observations. tclean overview, 2016.

List of Figures

1 U V and W coordinate space 1
2 U V and W coordinate space 3

List of Tables

6 Ehrlichkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende schriftliche Arbeit selbstständig und nur unter Zuhilfenahme der in den Verzeichnissen oder in den Anmerkungen genannten Quellen angefertigt habe. Ich versichere zudem, diese Arbeit nicht bereits anderweitig als Leistungsnachweis verwendet zu haben. Eine Überprüfung der Arbeit auf Plagiate unter Einsatz entsprechender Software darf vorgenommen werden.
Windisch, December 17, 2018

Jonas Schwammberger