Improving spatial domain based image formation through compressed sensing.

Gene Stoltz[a, b] and André Leon Nel[b]

[a]Council for Scientific and Industrial Research of South Africa
[b]University of Johannesburg, South Africa

ABSTRACT

In this paper, we improve image reconstruction in a single-pixel scanning system by selecting an detector optimal field of view. Image reconstruction is based on compressed sensing and image quality is compared to interpolated staring arrays. The image quality comparisons use a "dead leaves" data set, Bayesian estimation and the Peak-Signal-to-Noise Ratio (PSNR) measure.

Compressed sensing is explored as an interpolation algorithm and shows with high probability an improved performance compared to Lanczos interpolation. Furthermore, multi-level sampling in a single-pixel scanning system is simulated by dynamically altering the detector field of view. It was shown that multi-level sampling improves the distribution of the Peak-Signal-to-Noise Ratio.

We further explore the expected sampling level distributions and PSNR distributions for multi-level sampling. The PSNR distribution indicates that there is a small set of levels which will improve image quality over interpolated staring arrays. We further conclude that multi-level sampling will outperform single-level uniform random sampling on average.

Keywords: Image formation, Compressed sensing, Single pixel scanning system, Bayesian estimation.

1. INTRODUCTION

Since the development of compressed sensing, the image formation process within the optical domain can be divided into two categories; spatial domain sampling and global domain sampling. Global domain sampling can be defined as the sampling of the total output image field of view by transforming the wavefront into a basis such as single pixel imaging using a Digital Mirror Device[1] or multiple scatter imaging using scattering media.[2] Spatial domain sampling, the traditional method, is achieved by using the irradiance measurement of an optical detector with a specific field of view. Multiple of these measurements can be used to reconstruct an output image much larger than the detector field of view with the measurements linearly correlated to the output image. The two most notable methods are the use of a staring array and the scanning of a single pixel such as in a Rosette Imager[3] using a Rosette scanner or a scanning optical microscopy using an optical probe.[4]

In Rosette imaging, compressed sensing has been used to transform non-grid based spatial domain samples into grid-based samples, simulating a staring array.[5] Using the Rosette Imager, it has been shown that non-grid based sampling can match the performance of a traditional staring array (grid-based sampling).[3] These results are not unexpected as compressed sensing require signals to be sparse in the sampling domain and natural scenes are not sparse in the spatial domain.

In recent years, compressed sensing principles have been expanded from incoherence, sparsity, and random sampling to more general principles; asymptotic sparsity, asymptotic incoherence, and multilevel sampling,[6] with multi-level uniform sampling further extended to structured sampling.[7] With these expanded principles we investigate compressed sensing firstly with regards to structured sampling and secondly with regards to multi-level sampling. Structured sampling is explored through grid-based sampling in the spatial domain and can be implemented by equivalent hardware such as Single Fiber Endoscopy(SFE).[8] Secondly, multi-level sampling in

the spatial domain is explored by assuming the change in detector field of view simulates different sampling levels which can be implemented in systems such as a Rosette Imager.

In Section 2, the compressed sensing setup for spatial domain sampling is discussed. It is shown how compressed sensing can be used for interpolation, how an optimal probe size can be selected for single pixel scanning systems and a multi-level sampling approach is developed. In Section 3, the tools to evaluate image quality and compare different algorithms are discussed. In Section 4 the comparative results for interpolation and multi-level sampling are presented. Finally, in Section 5, the results are discussed and possible future work is presented.

2. SINGLE DETECTOR SPATIAL DOMAIN IMAGE FORMATION

The image formation process in the spatial domain uses the irradiance measurements of single detectors and related directional positions. A staring array can be simulated by using a square detector and moving the detector in a grid-like pattern until all positions have been covered correlating each detector position to a pixel in the output image. The single detector scanning system does not always have a one-to-one relation to the output image, such as probe overlap when using circular probes, and therefore regularisation methods are required for image formation.

Compressed sensing provides a regularisation method to reconstruct an image from overlapping probes, different probe shapes, and unstructured sampling positions; and is elaborated on in Section 2.1. The capability of a single detector to simulate a staring array creates the opportunity for compressed sensing to function as an interpolation algorithm. Compressed sensing as an interpolation algorithm as well as the use of an optimal square or circular probe size in grid-based sampling is explored in Section 2.2.

Compressed sensing theory suggests multi-level sampling to improve signal reconstruction. Levels within the spatial domain is ill-defined compared to levels in the Wavelet Transform. In Section 2.3 we define levels equivalent to different probe sizes and develop an approach to multi-level uniform random sampling.

2.1 Compressed Sensing

The compressed sensing framework requires a measurement matrix and a regularisation algorithm within a sparsity basis. In this paper the measurement matrix is formed through spatial domain sampling.[5] A sample within the measurement matrix is constructed by discretizing a probe onto the output image at a predetermined position. The probe simulates the irradiance measurement of a detector taking the sum of the pixels covered in the output image.

The regularisation algorithm is chosen to support real-time reconstruction from structured sampling and thus the Fourier Domain Regularisation Inversion (FDRI) algorithm is an adequate choice.[3] The FDRI algorithm only approximates an $l1$-optimisation algorithm but does match the performance of Nesterov's algorithm (NESTA) and Basis Pursuit Denoising (BPDN).[9]

It is noted that in the the paper an output image size of 64x64 will be used with 19.1% sampling, equating to 784 samples. The 784 samples allow for a 28x28 sampling grid, equating to a 28x28 staring array with appropriate square probe, as explained in Section 2.2.

2.2 Grid-based Sampling

The grid-based sampling method simulate detector positions similar to elements in a staring array where various probe forms and sizes can be used to sample at these positions. Mapping a sampling grid of 28x28 onto an output image of 64x64 pixels using square probes of 2.28x2.28 pixels placed on the sampling grid to simulate a 28x28 staring array. The measurement matrix can be created by using each square probe as an entry. This measurement matrix can then be used to generate the inversion matrix using the FDRI algorithm. The compressed sensing framework effectively becomes an interpolation algorithm as developed in the this section. The interpolation algorithm is called the *CS-FPA*.

The probe size of a scanning detector is not limited by the grid size as in the case of a staring array. One can select an optimal size for a square and circular probe by following a greedy optimisation approach. The approach scans from small to large probe size using the same sampling grid and compressed sensing to measure

the average PSNR of the data set for each probe size (see Section 3 for more information.). The results are shown in Figure 1 where the optimal probe size for the circular probe, a diameter of 3.71 pixels, and a side of 3.34 pixels for the square probe. The sampling approach, relative probe size, and regularisation algorithm for the circular probe is called the *CS-Circle*, and for the square probe the *CS-Square*.

(a) Square probe optimisation curve   (b) Circular probe optimisation curve
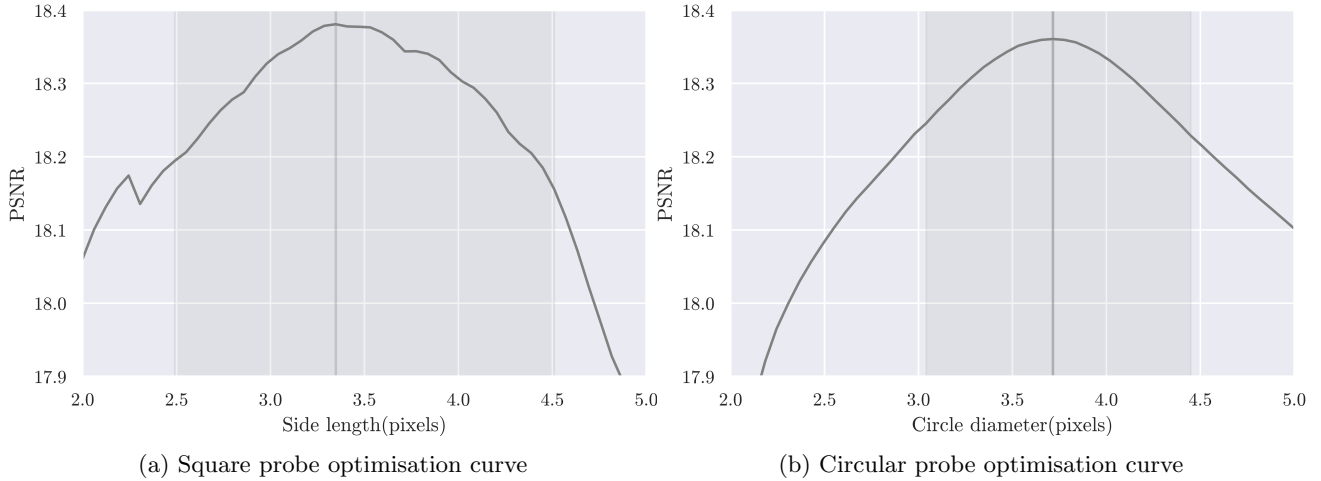
Figure 1: The PSNR results of different probe sizes with a dark gray region indicating the 89% HDI using the maximum average PSNR value as the reference. Figure 1a gives the optimisation curve for a square probe with the maximum average PSNR value of 18.38dB at 3.34 pixel side length.Figure 1b gives the optimisation curve for a circular probe with the maximum average PSNR value of 18.36dB at 3.71 pixel diameter.

2.3 Multi-Level Sampling Setup

The extended theory of compressed sensing introduces multi-level sampling. In certain transformations, such as the Wavelet domain, levels can be interpreted as single coefficients within the transform or as a range of coefficients. Each level is then sampled from a uniform random distribution with a pre-determined number of samples. There is currently no guidelines in choosing the level ranges but it is suggested that samples are to be taken per level should it be distributed by a power law.[6]

In the spatial domain, a single coefficient directly translates to a pixel value. Working with probes covering multiple pixels, the probe size can be interpreted as a level with the probe distributed uniformly across the output image. From this perspective, a range of levels will consist of a range of probe sizes. This interpretation also makes it possible to implement in hardware by changing the field of view of a detector.

In Section 2.2, a greedy approach was followed to find the single level providing optimal results. Using a greedy approach to find the optimal set of levels and adequate power law distribution in a large search space is not viable. We devise a model to be used to explore the parameter space and expected PSNR distribution of the multi-level sampling scheme.

The model uses five parameters. Four indicate a radius size; and one manages the coefficient of the power law which determines the number of samples per radius size. Let $r_i \in [0..31] \in \mathbb{R}$ be an arbitrary radius ranging from 0 to 31 pixels, using four different radiusses $i \in [0..3] \equiv R \in \mathbb{N}$ and requiring a total of $m \in \mathbb{N}$ samples. The number of samples $m_i \in \mathbb{R}$ per radius r_i is given by

$$m_i = \left\lfloor c \cdot e^{-\alpha\left(\frac{i}{\|R\|}\right)} \cdot m \right\rfloor \tag{1}$$

where c is a constant such that $\sum_i e^{-\alpha\left(\frac{i}{\|R\|}\right)} = 1$ and $m \leq \sum_i m_i$. The model parameter distributions are determined using the Markov Chain Monte Carlo (MCMC) algorithm called the "t-walk"[10] with the PSNR value as the likelihood measure.

3. ASSESSMENT METHODS AND PROTOCOLS

The assessment of the various algorithms is performed by means of a data set, a quality measure and hypothesis testing to compare algorithms. A "dead leaves" data set is generated with a power law distribution parameter set to 3 for the simulation and a small size cut-off on the circle radius at 0.7 pixels. This allows for a scale invariant approximation of natural scenes.[11] The data set contains 200 64x64 images. The quality measure for assessing image reconstruction accuracy is based on the PSNR measure. (See Appendix A.1). Images are converted to 16 bit, before measuring the PSNR, placing the maximum possible PSNR value at 96.33 dB.

The algorithm comparison is performed by using two group comparisons using Bayesian estimation as presented by Kruschke.[12] A Student's t-distribution for model approximation and uniform priors are used (See Appendix A.2). The Highest Density Interval (HDI) of 89% measured on the effect size is used to determine statistical significance.[13] The relative performance of two algorithms are evaluated by approximating the posterior distribution of the effect size via a MCMC algorithm. Typically, negative effect size values support the first algorithm and positive effect size value supports the second algorithm. However, from a Bayesian perspective the algorithm is assumed to perform better if the 89% HDI of the effect size are strictly positive or negative.

3.1 Grid-based Sampling

Two evaluations are required. The first is to evaluate whether compressed sensing can be used as an interpolation technique and the second is to determine whether improved image formation performance can be achieved by optimising the probe size while sampling within a grid.

All three algorithms will be compared against each other as well as some well known interpolation techniques; namely nearest neighbour, linear, cubic and Lanczos interpolation. The interpolation techniques will upscale a 28x28 sized image which was decimated through an averaging process simulating staring array sampling. The seven algorithms will be compared using Bayesian estimation and the HDI of the effect size to determine whether any method is statistically better than another.

3.2 Multi-level Sampling

To investigate multi-level sampling, the PSNR distribution of the proposed model is compared to the PSNR distribution of single-level sampling. The model for single-level sampling contains a single parameter, indicating the probe radius and also uses the PSNR measure as a likelihood indicator. The two PSNR distributions will be compared to determine whether multi-level sampling does improve over single-level sampling in the spatial domain.

4. RESULTS

4.1 Grid-based Sampling

An example image for each interpolation method and compressed sensing method is shown in Figure 2; with the final results shown in Table 1.

(a) Original

(b) Nearest
16.52 [16.51]

(c) Linear
17.40 [17.38]

(d) Cubic
17.95 [17.94]

(e) Lanczos
18.01 [17.99]

(f) CS-FPA
18.18 [18.19]

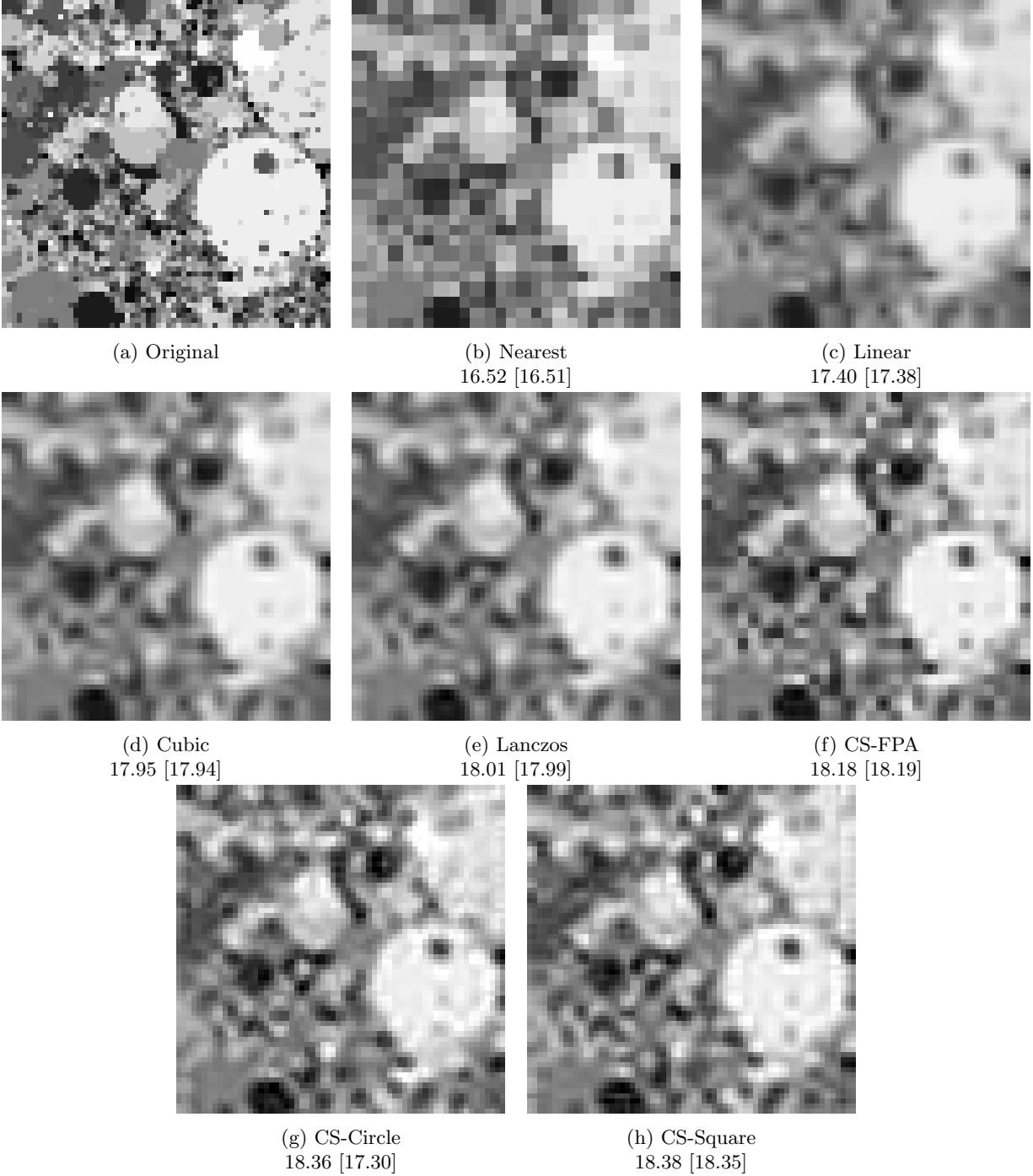(g) CS-Circle
18.36 [17.30]

(h) CS-Square
18.38 [18.35]

Figure 2: Example images for each algorithm with the average PSNR value indicated below and the specific image PSNR shown in square brackets.

The numerical experimentation has shown that compressed sensing as an interpolation method (CS-FPA) outperforms Lanczos interpolation as well as cubic interpolation. It is interesting to note that Lanczos interpolation does not outperform cubic interpolation in this case. Furthermore, the compressed sensing algorithms with optimised probes (CS-Circle and CS-Square) outperform all image formation algorithms but do not differ significantly from one another.

Table 1: The table provides the results for grid based sampling and interpolation techniques. The bold values give the average PSNR for each algorithm while all other values are the minimum HDI value calculated as row algorithm versus column algorithm. A negative HDI value indicates that the row algorithm is not statistically better than the column algorithm.

.	Nearest	Linear	Cubic	Lanczos	CS-FPA	CS-Circle	CS-Square
CS-Square	2.71	1.34	0.50	0.38	0.13	-0.11	**18.38(1.06)**
CS-Circle	2.64	1.30	0.41	0.30	0.05	**18.32(1.07)**	
CS-FPA	2.39	1.03	0.16	0.06	**18.18(1.08)**		
Lanczos	2.14	0.77	-0.08	**18.01(1.08)**			
Cubic	2.04	0.66	**17.94(1.08)**				
Linear	1.22	**17.40(1.06)**					
Nearest	**16.52(1.04)**						

4.2 Multi-level Sampling

The PSNR distributions generated through the MCMC for single-level and multi-level sampling is shown in Figure 3 while the parameter distributions are shown in Figure 4.
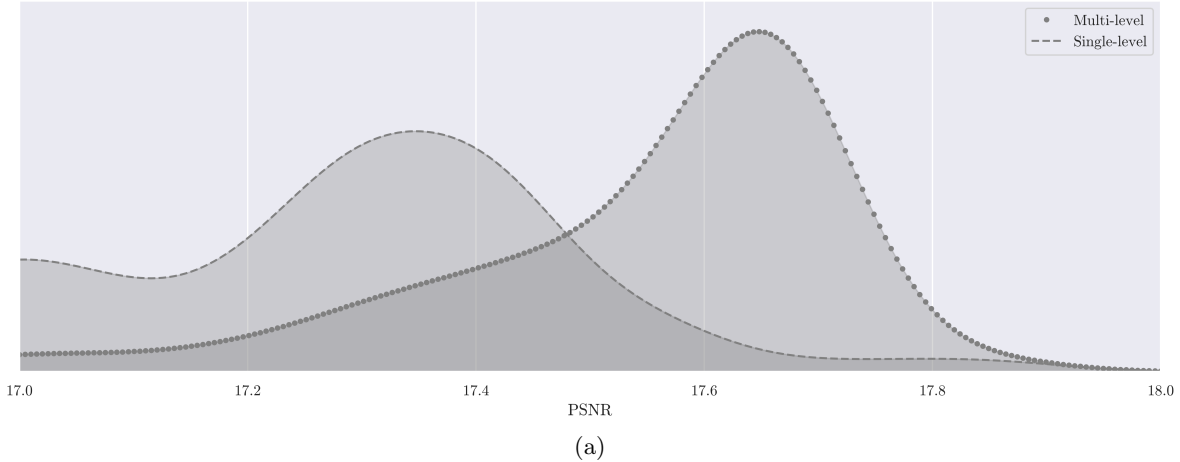

(a)

Figure 3: The PSNR distributions for multi-level and single-level sampling. The average PSNR value for multi-level sampling is 17.54dB and 17.27dB for single-level sampling.

The parameter distributions shown in Figure 4b show that only a small range of radii produce high PSNR values and, on average, all other parameters perform much worse. The parameter distributions in Figure 4b for r_1 is tending to the same parameter value found in the single-level sampling while the parameter α tends to be skewed to a lower value. Lower values of α create a flattened power law distribution indicating that values are more uniformly distributed over the four radius parameters. There is no correlation identified which will produce high PSNR values as in the single-level case. This might be due to the MCMC not finding adequate steps and might require custom adjustments.
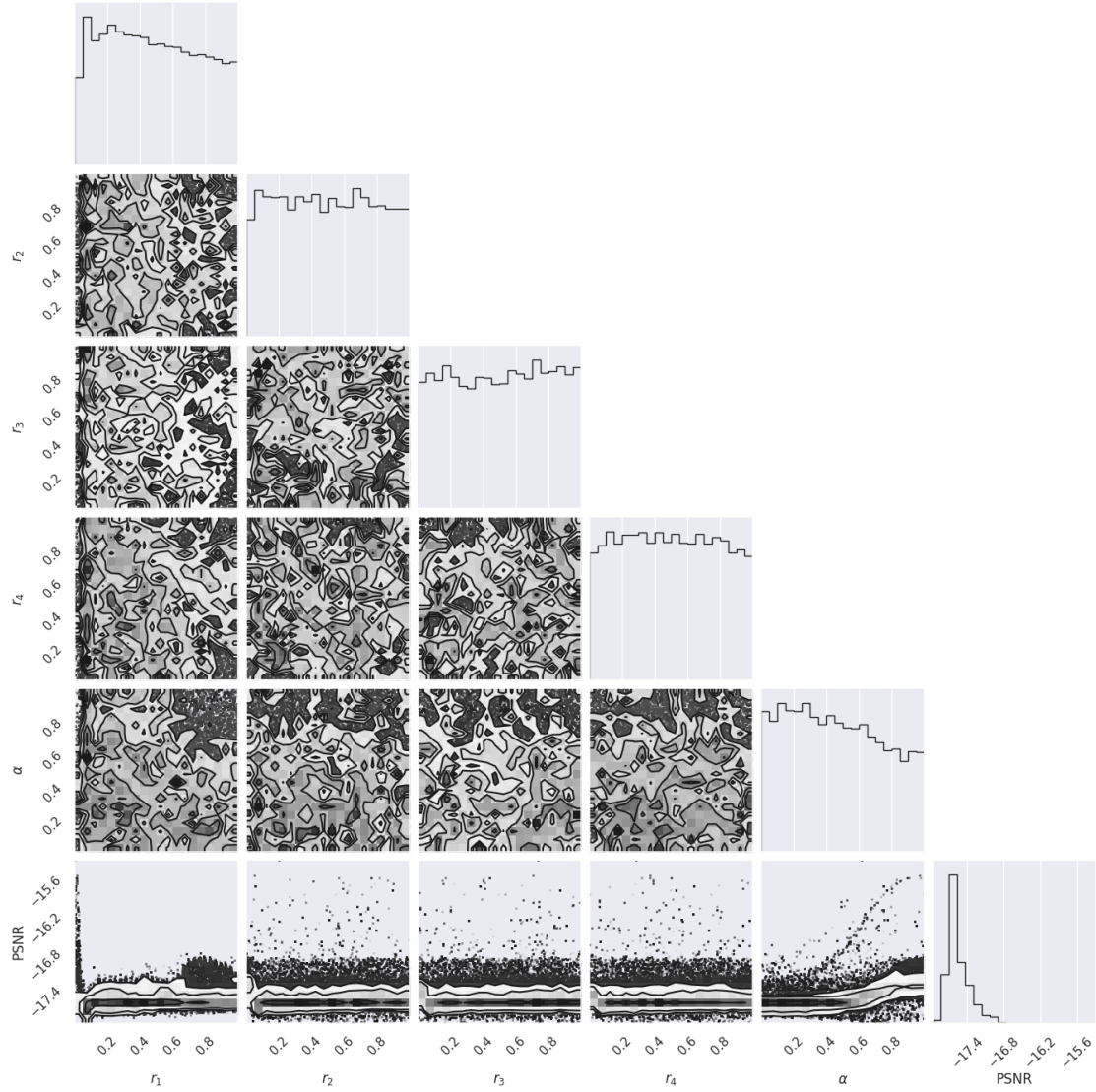
(a) The parameter distributions for the multi-level model.
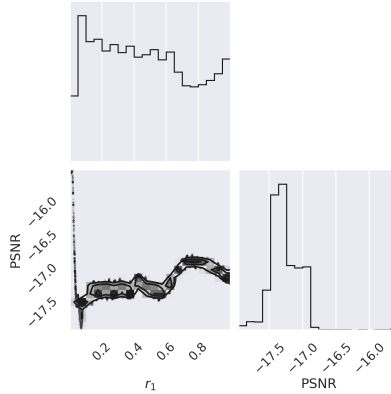
(b) The parameter distributions for the single-level model.

Figure 4: The parameter co-variance matrices for the MCMC chains produced by the "t-walk".

5. DISCUSSION AND FUTURE WORK

It has been shown that single pixel scanning systems have the potential to outperform staring array based imaging systems in regards to image quality by adequately optimising the detectors and sampling patterns. It was first shown that compressed sensing can be used as an interpolation algorithm that outperforms traditional interpolation algorithms such as Lanczos and cubic interpolation. This was then extended to use a sample grid equivalent to a staring array but optimising the probe size used in the sampling. This further improved the image quality beyond what a staring array could produce for the same amount of samples. Finally it was shown that multi-level sampling improve its PSNR distribution compared to single-level sampling.

A 28x28 staring array was used and up-scaled to a 64x64 image by using an interpolation algorithm. The compressed sensing based interpolation algorithm, CS-FPA, was developed by modelling each detector in the staring array as a separate sample, with each sample corresponding to a square with sides equal to 2.28 pixels related to the output image. The chosen image reconstruction algorithm, FDRI, allows for generating an interpolation matrix usable on any image to upscale from 28x28 to 64x64. An interpolation matrix must be calculated for every set of up-scaling. The CS-FPA outperformed Lanczos interpolation by producing an average PSNR value of 18.18 dB compared to 18.01 dB. It was then fathomed that if a single detector scanning system would sample at the same positions as the staring array, taking the centre of each element of the staring array as a position, it is possible to adjust the size of the sampling probe or detector field of view too possibly improve image quality. Not only can the probe size be changed but the probe shape can be selected as well. In the paper the probe shape was limited to a square and circular probe with algorithms called CS-Square and CS-Circle respectively. The optimal probe size was determined by selecting the probe size producing the highest average PSNR value. CS-Square with a probe size of 3.34 pixels produced a PSNR value of 18.38 dB while CS-Circle with a diameter of 3.71 pixels produced a PSNR value of 18.36dB; both outperformed CS-FPA and Lanczos.

It must be noted that CS-Circle and CS-Square is not interpolation algorithms as they are dependant on the physical sampling processes used. To summarise, when given a choice to take a set amount of samples with either a staring array or a single detector scanning system, and the detector field of view in the scanning system can be adjusted, the results presented here indicated that the scanning system will outperform the staring array.

Compressed sensing theory indicates that multi-level sampling can further improve signal reconstruction quality. Too investigate this, a multi-level sampling approach was developed and the PSNR distribution was compared to a single-level sampling approach. Only the PSNR distributions was compared due to the high parameter space present in multi-level sampling; there is currently no guarantee that optimal parameters will be found. The average PSNR distribution for multi-level sampling was improved compared to single-level sampling, 17.54 dB versus 17.17 dB respectively. It can be concluded that using a multi-level sampling model will improve the PSNR distribution compared to a single-level sampling model.

Spatial domain sampling can further be developed by investigating possible different probe forms and improved real-time regularisation algorithms. Multi-level sampling has shown potential and research in optimisation approaches for finding optimal sampling structures or identifying optimal probe ranges and corresponding number of samples can be investigated.

APPENDIX A. EQUATIONS

A.1 Peak-Signal to Noise Ratio Equations

Given an image I of height H and width W such that $I(m,n) \mapsto i_{m,n} \in 0, \dots, 2^{16} - 1$ with $|m| \leq H - 1$, $|n| \leq W - 1$ and $m, n \in \mathbb{Z}$. The Mean Square Error between a reference image I_R and a test image I_T is given be

$$MSE(I_T, I_R) = \frac{1}{HW} \sum_{m=0}^{H-1} \sum_{n=0}^{W-1} \left[I_T(m,n) - I_R(m,n) \right]^2$$

The Peak Signal-to-Noise Ratio (PSNR) in decibel (dB) is defined as

$$PSNR(I_T, I_R) = \begin{cases} 20 \cdot \log_{10} \left(\frac{2^B - 1}{\sqrt{MSE(I_T, I_R)}} \right) & MSE(I_T, I_R) \geq 1 \\ 20 \cdot \log_{10} \left(2^B - 1 \right) & Otherwise \end{cases} \tag{2}$$

A.2 Student's t-distribution

The Student's t-distribution is given by

$$\text{StudentT}(y|\nu,\mu,\sigma) = \frac{\Gamma\left((\nu+1)/2\right)}{\Gamma(\nu/2)} \; \frac{1}{\sqrt{\nu\pi}\,\sigma} \; \left(1 + \frac{1}{\nu}\left(\frac{y-\mu}{\sigma}\right)^2\right)^{-(\nu+1)/2}$$

with ν as the normality parameter, μ as the mean, σ as standard deviation and y as the data point.

The effect size is calculated by:

$$\text{Effect Size} = \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{1}{2}\left(\sigma_1^2 + \sigma_2^2\right)}}$$

REFERENCES

[1] Duarte, M. F., Davenport, M. A., Takhar, D., Laska, J. N., Sun, T., Kelly, K. F., and Baraniuk, R. G., "Single-Pixel Imaging via Compressive Sampling," *IEEE Signal Processing Magazine* **25**(2), 83–91 (2008).

[2] Liutkus, A., Martina, D., Popoff, S., Chardon, G., Katz, O., Lerosey, G., Gigan, S., Daudet, L., and Carron, I., "Imaging with nature: Compressive imaging using a multiply scattering medium," *Scientific Reports* **4**, 1–7 (2014).

[3] Stoltz, G. and Stoltz, M., "Performance estimation of a real-time rosette imager," (2020).

[4] Seibel, E. J. and Pollack, G. H., "Imaging 'intact' myofibrils with a near-field scanning optical microscope," *Journal of Microscopy* **186**(3), 221–231 (1997).

[5] Uzeler, H., Cakir, S., and Aytaç, T., "Image reconstruction for single detector rosette scanning systems based on compressive sensing theory," *Optical Engineering* **55**(2), 023108 (2016).

[6] Adcock, B., Hansen, A. C., Poon, C., and Roman, B., "Breaking the coherence barrier: A new theory for compressed sensing," in [*Forum of Mathematics, Sigma*], **5**, Cambridge University Press (2017).

[7] Li, C. and Adcock, B., "Compressed sensing with local structure: Uniform recovery guarantees for the sparsity in levels class," (2016).

[8] Seibel, E. J., Brown, C. M., Dominitz, J. A., and Kimmey, M. B., "Scanning Single Fiber Endoscopy: A New Platform Technology for Integrated Laser Imaging, Diagnosis, and Future Therapies," *Gastrointestinal Endoscopy Clinics of North America* **18**(3), 467–478 (2008).

[9] Czajkowski, K. M., Pastuszczak, A., and Kotyński, R., "Real-time single-pixel video imaging with fourier domain regularization," *Opt. Express* **26**, 20009–20022 (Aug 2018).

[10] Christen, J. A. and Foxy, C., "A general purpose sampling algorithm for continuous distributions (the t-walk)," *Bayesian Analysis* **5**(2), 263–282 (2010).

[11] Gousseau, Y. and Roueff, F., "The dead leaves model: general results and limits at small scales," *arXiv preprint math/0312035* **0**, 1–38 (2003).

[12] Kruschke, J. K., "Bayesian estimation supersedes the T test," *Journal of Experimental Psychology: General* **142**(2), 573–588 (2013).

[13] McElreath, R., [*Statistical rethinking: A Bayesian course with examples in R and Stan*], CRC press (2020).