Comparison of Phase Retrieval Methods

Oliver Johnson

Abstract—Phase retrieval is a common problem that presents in many different fields. In this project several several phase retrieval methods based on the alternating projections were compared and contrasted. These methods were Gerchberg–Saxton, Hybrid Input-Output and Oversampling Smoothness. These methods were tested on the MNIST handwritten digit dataset and examples from the BSDS300 dataset. It was found that Hybrid Input-Output and Oversampling Smoothness had higher performance, but their performance could not reliably differentiated over the small test suite conducted in this report. Overall the results show the successful application of these phase retrieval methods to test data.

Index Terms—Phase Retrieval

1 Introduction

PHASE retrieval is an important problem in many fields such as optics, single particle imaging, X-ray crystallography, signal processing and diffraction imaging. In diffraction imaging, an object is placed in front of laser, and based on how the light interacts with object, a diffraction pattern emerges, and usually only the magnitude of this light can be measured with a detector [1]. In order to fully reconstruct the image we need magnitude and phase information and trying to estimate the phase in this kind of situation is the problem that phase detection seeks to solve. Mathematically the problem involves recovering the entire signal x when only the Fourier transform magnitude information $|\mathcal{F}(x)|$ is known. This is a common situation for many other measuring devices where particular systems may only record magnitude information, but not phase information.

The difficulty is that the problem is fundamental ill-posed. If a is a complex number indicating a point in Fourier space, then any of $ae^{i\theta}$ will have the same magntiude $|a|$. Especially if the measurements are noisy, it can become very difficult to reconstruct the original image [1]. However many methods have been able to achieve good phase estimation with various methods, which are described in related work. As shown in Figure 1 it is often the case that the phase information in the FFT of the image is often more important for reconstruction than the magnitude of the FTT of the image.

In this project we seek to study and compare various algorithms to solve the phase retrieval problem. We will particularly focus on alternating projection methods tested image samples on a variety of datasets.

2 Related Work

One of the most fundamental methods in phase retrieval is the Gerchberg–Saxton algorithm [2], which employs an iterative method to retrieve the phase. In many cases multiple image measurements will be combined to give more data to determine phase. Building on this, Fienup [3] introduced several other methods, including the Hybrid input-ouput algoirthm, which changed the object domain step to have a negative feedback term. This is one of the most widely used and sucessful methods in phase detection. A further extension of this is oversampling smoothness, which helps address the problem of noisy Fourier magnitude data. Recent methods also include PhaseLift [4], which uses a method based on complex programming and matrix completion to estimate the phase. Signal sparsity has also been used to give additional prior information [5]. Some more recent work has used end-to-end deep learning methods for phase retrieval, where a neural network is trained to calculate the inverse mapping. Manekar et al. [6] employed this approach and used the U-Net architecture and trained on the MNIST handwritten dataset. They obtained good qualitative performance, but did not undertake a thorough quantitative comparison with other methods. They also experimented with the effect of symmetries of the image on learning. Deep learning methods might be especially useful when there is specific imaging problem where the training data can be used to make a deep learning model to solve the inverse problem for this situation.

• *Oliver Johnson is with the Department of Electrical Engineering, Stanford University*

3 Method

3.1 Problem Formulation

For the general case, the phase retrieval problem can be expressed as:

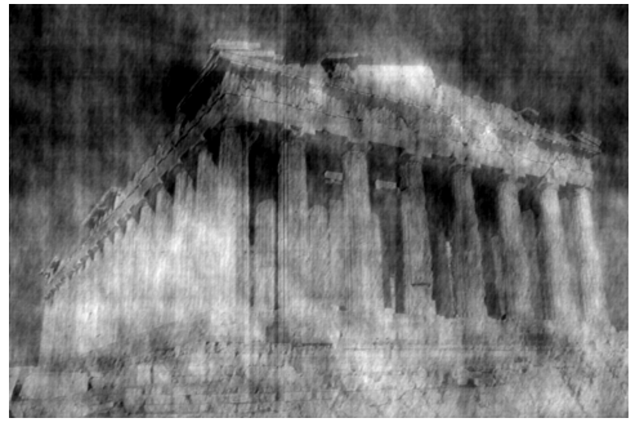
$$\boldsymbol{y}_k = |\langle \boldsymbol{a}_k, \boldsymbol{x} \rangle|^2, \quad k = 1, ..., M \tag{1}$$

where $\boldsymbol{x}$ is the signal vector and $\boldsymbol{a_k}$ are the measurement vectors. For the case of Fourier phase retrieval this reduces to the case where $\boldsymbol{a}_k[n] = e^{-i2\pi \frac{kn}{M}}$ [7]. This means that equation 1 can be expressed as:

$$\boldsymbol{y} = |\mathrm{DFT}[\boldsymbol{x}]|^2 \tag{2}$$

In general the recovery of the signal does not lead to a unique solution. When trying to recover the original signal, there are several trivial ambiguities that need to be considered. If the image undergoes any of the following three transformation, then the Fourier transform magnitude will be remain unchanged. The first is global phase shift, which is where the entire image is shifted by a particular phase. However for a real input image this is not as relevant, since

(a) Image 1



(b) Magnitude from image 1, phase from image 2

(c) Image 2

(d) Magnitude from image 2, phase from image 1

Fig. 1: The FFT magnitude is switched in these two images, this shows that the phase contains comparatively more information than the magnitude.

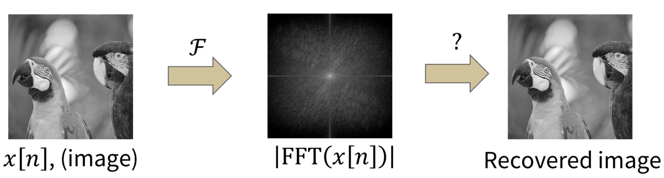


Fig. 2: Problem Definition

we always know that our signal is positive and real. The second is conjugate inversion, which corresponds to a flip of the (real) input image. The third is a spatial shift of the image [7]. In the results section of this project there will be examples of some of these phenomena.

3.2 Techniques applied

We propose to implement several phase retrieval algorithms from the literature and apply them on the the MNIST dataset [8] and BSDS300 dataset [9], and compare the relative performance. The methods we propose to use are listed below:

- Gerchberg-Saxton algorithm (GS) [2]
- Hybrid Input–Output (HIO) [3]
- Oversampling Smoothness (OSS) [10]

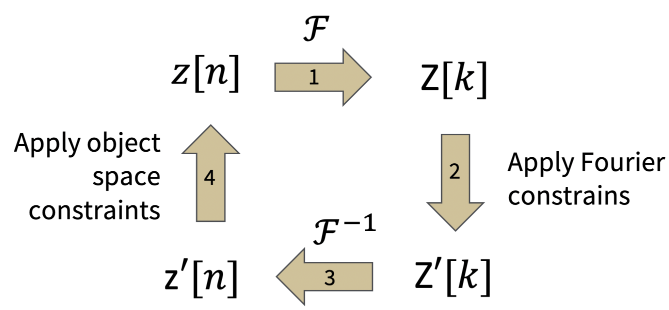


Fig. 3: Alternating projection methods

The alternating projection methods switch between the object and Fourier domains and apply various constraints as shown in Figure 3.

The first three steps of the alternating projection methods are described below:

1) Take Fourier transform of $z[n]$ to obtain $Z[k]$, i.e.

$$Z'[k] = \text{DFT}[z[n]] \tag{3}$$

2) Apply Fourier magnitude constraints:

$$Z_i[k] = |X[k]| \cdot Z[k]/|Z[k]| \tag{4}$$

3) Take inverse Fourier Transform of $Z[k]$ to obtain $z'[n]$, i.e.

$$z'[n] = \text{DFT}^{-1}[Z'[k]] \tag{5}$$

The fourth step differs depending on the method used. For GS, this is a object-space magnitude constraint as

$$z[n] = |x[n]| \cdot z'[n]/|z'[n]| \tag{6}$$

For HIO, the fourth step is negative feedback term with parameter $\beta \in (0, 1]$ and γ is the set of indices that contradict the constraints in object space [7].

$$z[n] = \begin{cases} z'[n] & n \notin \gamma \\ z[n] - \beta z'[n] & n \in \gamma \end{cases} \tag{7}$$

And for OSS, the fourth step switches to the Fourier domain and applies a Gaussian filter in the off-support region [7].

$$z''[n] = \begin{cases} z'[n] & n \notin \gamma \\ z[n] - \beta z'[n] & n \in \gamma \end{cases} \tag{8}$$

$$z[n] = \begin{cases} z''[n] & n \notin \gamma \\ \text{DFT}[Z''[k]W[k]] & n \in \gamma \end{cases} \tag{9}$$

where $W[k]$ is a Gaussian function that decreases its variance as the iterations progress [10].

3.3 Oversampling of Fourier Transform

A standard approach to help solve these phase retrieval problems is to oversample the Fourier transform. This makes the problem slightly less overdetermined so it is easier to solve. We know that taking the Fourier transform of a zero-padded image gives the oversampled Fourier transform of the image, so we zeropadded our image before taking the Fourier Transform. This padding can then be used as an additional constraint in the object domain.

Another approach is to use masks (that could be random) that modify the image before the fourier transform is applied. Since the masks are known, this essentially increases the number number of equations available to solve this underdetermined problem, so it is better posed. A disadvantage of this approach is that multiple captures of the same scene are required. Depending on the application this may or may not be a problem. We did not apply masks in this report but could be an avenue of future work.

4 Results

Figure 4 shows the the recovery using the three methods for the MNIST dataset with either no over sampling no no oversampling applied.

Table 1 shows some PSNR values for various iteration numbers for the large cat image (over-sampling applied), and Figure 5 shows the images for 10000 iterations.

Figure 6 shows the results for a smaller section of the cat image, but with $\mathcal{N}(0, 1)$ noise added to Fourier magnitudes.

TABLE 1: PSNR results for cat image

Number Iterations	GE	HIO	OSS
10000 iterations	16.47	26.0	38.9
20000 iterations	16.3	72.0	58.2

5 Discussion

The results show that the various phase detection methods worked successfully on the test data. It is noted that without the oversampling, the reconstruction of the MNIST digits is fairly poor, although some remnants of the numbers can be made out. Comparing similar results in the literature, the end-to-end deep learning approach to phase retrieval [6] which was trained on the MNIST dataset gave superior results to the ones we have presented here (clear representations of the numbers). In this work the authors did not oversample the Fourier transform. However, it is possible that their approach is highly specialized for the MNIST dataset only.

Note how for Figure 4 (a) sometimes the numbers wrap around the edge, but for (b) they do not. This is because the masking operation in (b) has effectively constrained the image so it must appear in one connected piece inside the boundaries. Since around each character, the pixel values are zero, the digit is free to shift in the object domain up until the edge of the digit without changing the Fourier magnitude. In the oversampled signal the mask effectively constrains the digit to appear all in one piece.

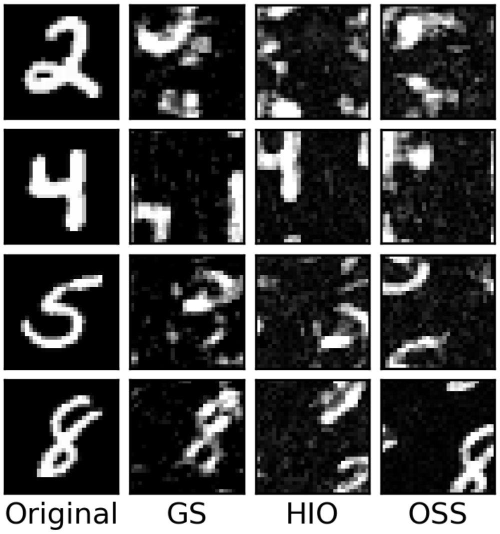
Comparing the various methods for the oversampled case, it is noted that HIO and OSS seem to almost perfectly reconstruct the digits (up to a rotation/shift). given by [6], The PSNR was not calculated for these examples as due to the shifting behaviour of the output due the result would not have much meaning. We instead calculated the PSNR where the edges of the image were not zero valued so that it would easier to align the reconstructed image to the output.

Considering Figure 5, it can be seen that HIO and OSS reconstruct the original image quite well. While OSS did to better on this iteration, it was found that this was quite variable, and sometimes HIO would perform better.

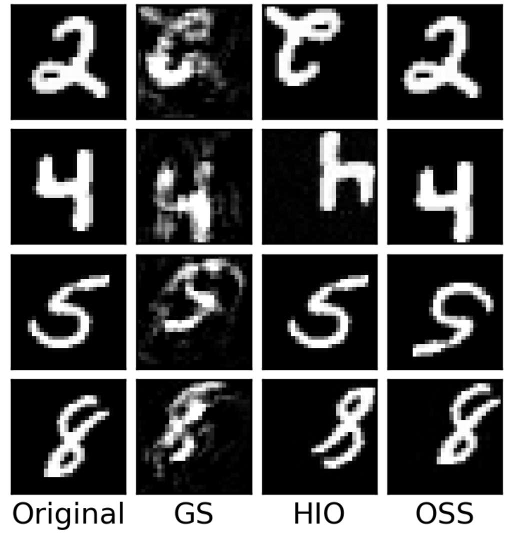
Table 1 shows that the reconstruction ability performance of GE was worse than HIO and OSS. However, for the tests conducted, HIO and OSS could not be reliably differentiated. It was found that the performance was highly variable depending on the run, so the two methods could not reliably be differentiated statistically in the time given. In the future it would be better to run these algorithms with more initial conditions on a wider set of images so they can be more reliably compared. It would also be good to further characterize the effect of adding noise to the Fourier magnitude coefficients to simulate a real measurement situation.

6 Limitations and Future Work

One limitation of this report is that it only considered simulated data, not raw measurements from an actual experiment, which would have provided a better indication of how these techniques worked in the real world. The testing could also have been completed on a wider dataset, as the outputs of the alternating-projection methods are



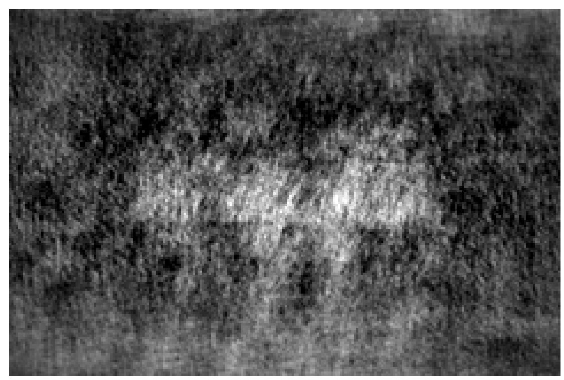
(a) No oversampling applied



(b) Oversampling applied

Fig. 4: Phase Retrievals from the MNIST dataset

(a) Original Image



(b) GE

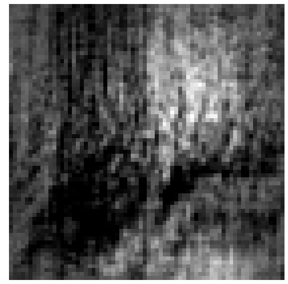
(c) HIO

(d) OSS

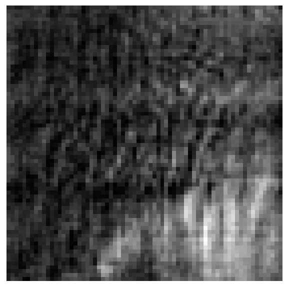
Fig. 5: Methods applied to image from BSDS300 dataset. This corresponds to the PSNR values in column 1 in table 1 (10000 iterations).

fairly sensitive and sometimes inconsistent, so the results presented here might not be representative of the wider context. It would also be desirable to test some more modern approaches such as PhaseLift, and deep learning based methods, however there was not time to implement these methods for this report. It would also be good to further

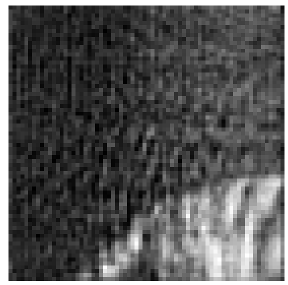
(a) Original Image



(b) GE. PSNR = 15.42



(c) HIO. PSNR = 19.55



(d) OSS. PSNR = 20.33

Fig. 6: Methods applied to image from BSDS300 dataset. In this case noise (zero mean, variance of one Gaussian) was added to the Fourier magnitudes. This test was repeated 10 times to give an estimate of the benefit in noise performance for OSS (Only one image is shown, but PSNR values are averaged). Qualitatively it looks as if OSS has slightly less wide ripples in the reconstruction.

characterize the noise added the Fourier magnitudes, as the OSS method did not see a large improvement. It is also suggested to use a different distribution of noise, such as Possion (photon noise), which may be more realistic for the typical measurement scenario when capturing data.

7 Conclusion

This report has compared and contrasted several different methods to reconstruct an image given only the magnitude of the FFT. It was shown that HIO and OSS provided a large improvement compared to GS when oversampling is applied. Without oversampling applied while HIO and OSS performed better, they still did not extract the finer details of the image. The relationship between the noise and the method used was unclear, so work needs to be done to characterize the methods over a wider range of data.

8 Code

The source code for this report is available at https://github.com/oliver-johnson1/phase-retrieval A part of the code was adapted from https://github.com/hhu-machine-learning/phase-retrieval-cgan, and the OSS MATLAB implementation from the OSS article's authors [10] was used to guide the python implementation of the OSS method.

References

[1] G. zhen Yang, B. zhen Dong, B. yuan Gu, J. yao Zhuang, and O. K. Ersoy, "Gerchberg–saxton and yang–gu algorithms for phase retrieval in a nonunitary transform system: a comparison," *Appl. Opt.*, vol. 33, no. 2, pp. 209–218, Jan 1994. [Online]. Available: http://opg.optica.org/ao/abstract.cfm?URI=ao-33-2-209

[2] R. W. Gerchberg and W. O. Saxton, "Practical algorithm for determination of phase from image and diffraction plane pictures," *OPTIK*, vol. 35, no. 2, pp. 237–&, 1972.

[3] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, Aug 1982. [Online]. Available: http://opg.optica.org/ao/abstract.cfm?URI=ao-21-15-2758

[4] E. J. CandÃšs, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Review*, vol. 57, no. 2, pp. 225–251, 2015. [Online]. Available: https://doi.org/10.1137/151005099

[5] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Convex algorithms and limitations," 2013.

[6] R. Manekar, K. Tayal, V. Kumar, and J. Sun, "End to end learning for phase retrieval," *ICML workshop on ML Interpretability for Scientific Discovery*, 2020. [Online]. Available: https://par.nsf.gov/biblio/10198729

[7] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[8] L. Deng, "The mnist database of handwritten digit images for machine learning research," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[9] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int'l Conf. Computer Vision*, vol. 2, July 2001, pp. 416–423.

[10] J. A. Rodriguez, R. Xu, C.-C. Chen, Y. Zou, and J. Miao, "Oversampling smoothness: an effective algorithm for phase retrieval of noisy diffraction intensities," *Journal of Applied Crystallography*, vol. 46, no. 2, p. 312–318, Feb 2013. [Online]. Available: http://dx.doi.org/10.1107/S0021889813002471