S. J. Weddell, R. Y. Webb, 'A Neural Network Architecture for the Reconstruction of Turbulence Degraded Point Spread Functions', *Proceedings of Image and Vision Computing New Zealand 2007*, pp. 103–108, Hamilton, New Zealand, December 2007.

A Neural Network Architecture for the Reconstruction of Turbulence Degraded Point Spread Functions

S. J. Weddell, R. Y. Webb

Dept. Electrical & Electronic Engineering, University of Canterbury, Christchurch, New Zealand.
Email: steve.weddell@canterbury.ac.nz

Abstract

Astronomical images, degraded by the effects of atmospheric turbulence, can be corrected in real-time using adaptive optics (AO). A reconstructed wavefront, measured using a reference object with a small angular separation from a target object, is used to compensate aberrations effecting a target object. The conjugate of the recovered wavefront is used to alter the optical path of a telescope for correction. To ensure high performance, a constraint is imposed on the angular separation between a reference object, such as a bright guide star, and target object which is typically fainter. This constraint is referred to as the isoplanatic angle, θ_0. Given the sparsity of natural guide stars, combined with limitations on the use of artificial guide stars, severe restrictions are imposed on the field of view (FOV) and turbulence compensation over the anisoplanatic region. Our aim is to use aberration data from two or more reference objects to extend the isoplantatic angle. This paper outlines a system architecture and proposes the use of artificial neural networks (ANN) to provide extended FOV coverage for real-time astronomical image restoration.

Keywords: image restoration, adaptive optics, echo state networks, wavefront prediction

1 Introduction

A planar wavefront entering the Earth's atmosphere is subject to distortion as light propagates through turbulence. The effects of such distortion, in terms of spatial resolution and sensitivity, can be corrected by up to two orders of magnitude using adaptive optics (AO) [1].

AO systems require a wavefront sensor (WFS) to measure aberrations induced by atmospheric turbulence. Since target objects are typically too faint to measure directly, natural or artificial source objects (guide stars) are imaged to achieve adequate signal to noise ratio (SNR). However, to maximise correlation of source and target aberrations the angular separation between objects must be limited to the isoplanatic angle, θ_0. Turbulence induced aberrations of target and source objects can be considered identical if $\theta_0 \leq 40$ arcseconds [2].

Target image reconstruction using a point spread function (PSF) in anisoplanatic regions requires more than one source. Our approach is to use a parallel architecture, based on a set of artificial neural network (ANN) modules, for real-time aberration prediction and correction. This paper describes a simulation platform for evaluation of an ANN module for the prediction of the space-varying PSF required for image reconstruction.

The problem statement is defined in the following section. Phase perturbations defined by Zernike polynomials are outlined in Section 3, and this is followed by an introduction to artificial neural networks in Section 4. The system architecture is described in Section 5 and an overview of the simulation platform is given in Section 6. Our interim results are presented in Section 7 and this is followed by our conclusion and outline of future work in Section 8.

2 Problem Statement

The simultaneous capture of two slightly defocused images, comprised of multiple point source objects, i.e. natural guide stars (NGSs), propagating through separate regions and multiple layers of atmospheric turbulence, are used to determine modal wavefront aberrations.

Wavefront aberrations can be characterised in terms of Zernike polynomials [3]. The utilisation of these polynomials and employment of machine learning to determine wavefront aberrations in regions extending the isoplanatic region is the basis for this study.

Figure 1 shows the geometry of three objects and the angular separations that define two regions of variable wavefront correlation. Considering the sep-

aration angle, $\theta_{\gamma\beta}$, the aberrated PSFs that result from imaging a target object, γ, and source object, α, can be considered equivalent, since $\theta_{\gamma\beta} \leq \theta_0$. However, considering the angular separation of source object, α, and target object, γ, the aberrated PSFs that result from imaging these objects will be significantly different, since $\theta_{\alpha\gamma} > \theta_0$.

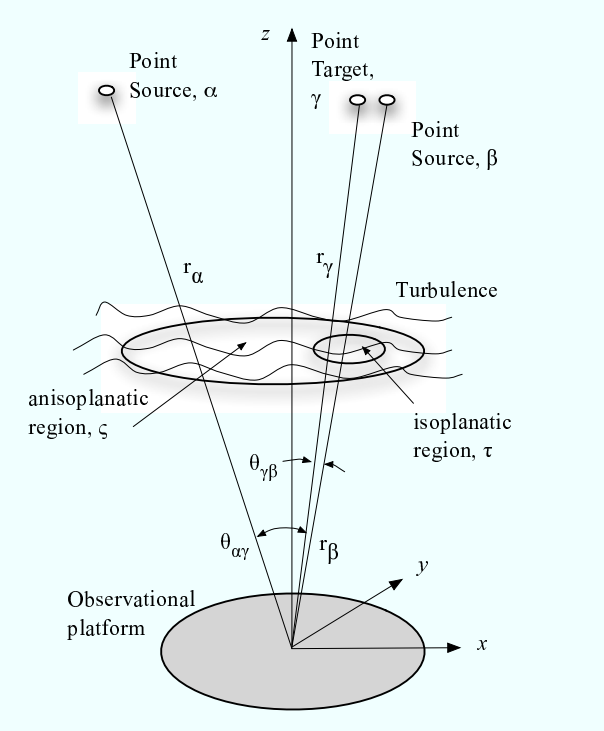

Figure 1: Geometry of Anisoplanatic Effects.

Determining the aberrations of a target PSF, such as γ in Figure 1, from aberrations measured using one or more source objects, such as α (shown in the same figure as exceeding θ_0), is ill-posed if any of the following conditions are satisfied [4]:

1. A nonexistent input vector results in an output vector i.e., the existence condition is violated.
2. Input data is insufficient to uniquely reconstruct the non-linear dynamics in the solution space.
3. The presence of noise or errors caused by imprecise measurements introduces discontinuity in the reconstruction.

Solutions to such inverse problems have been proposed, however architectural considerations for hardware implementation are rare. For example, Tikhonov regularization and the expectation-maximisation (EM) algorithm have been shown to improve predictability of the geometrical properties of the space-varying PSF with a mean squared error (MSE) of less than 2.7% [5]. Time-delay neural network (TDNN) architectures have been used for prediction of low-order wavefront aberrations [6]. Modal tomography has been adapted from medical imaging applications to estimate NGS wavefront distortions over a wide FOV [7].

Our approach to make such dynamic reconstruction problems well-posed is to use *a priori* knowledge of aberrations from multiple sources, such as point sources α and β in Figure 1, and through optimisation of the dynamic structure of an ANN. The use of spatially variable parameters, such as $\theta_{\alpha\gamma}$, will provide the basis for training and system evaluation.

3 Zernike Polynomials

Zernike polynomials are 2D orthonormal basis functions that represent an ordered series of aberrations on a unit circle. This set of polynomials is defined as

$$\left.\begin{aligned} Z_{even_i}(r,\theta) &= \sqrt{n+1}\, R_n^m(r)\sqrt{2}\cos(m\theta) \\ Z_{odd_i}(r,\theta) &= \sqrt{n+1}\, R_n^m(r)\sqrt{2}\sin(m\theta) \end{aligned}\right\} \quad m \neq 0,$$

$$Z_0(r) = R_n^0(r)\sqrt{2},\ m = 0. \tag{1}$$

The radial functions $R_n^m(r)$ are defined as

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^n\,(n-s)!}{s!\left[\frac{(n+m)}{2}-s\right]!\left[\frac{(n-m)}{2}-s\right]!} r^{(n-2s)}. \tag{2}$$

Each Zernike mode can be represented by a 2D image, commonly referred to as a *phase map*. For example, an aberrated PSF will be represented by N phase maps on the pupil plane. The linear combination of N aberrations over a unit circle of radius R results in the phase perturbation

$$\Psi(Rr,\theta) = \sum_{i=2}^{N} a_i Z_i(r,\theta), \tag{3}$$

where the piston term, Z_1, has been removed, as is common for single aperture instruments.

The Zernike coefficients are defined as

$$a_i = \int_0^1 \int_0^{2\pi} \mathcal{W}(r)\Psi(Rr,\theta)Z_i(r,\theta)d\theta dr, \tag{4}$$

where $\mathcal{W}(r)$ is an area weighting function,

$$\mathcal{W}(r,\theta) = \begin{cases} \frac{1}{\pi} & r \leq 1 \\ 0 & r > 1 \end{cases}. \tag{5}$$

Given the linear collection of aberrations, represented by Zernike polynomials in Equation 3, where $\Psi(Rr,\theta)$ can be represented by a vector location, $\vec{x}$, a wavefront can be defined as,

$$u_t(\vec{x}) = exp[j\Psi(\vec{x})]. \quad (6)$$

Simulating the propagation of several source objects through multiple turbulent layers, imaging each source using Fresnel diffraction, and recovering the Zernike coefficients a_i given in Equation 4, where $i = 2 \dots 20$, is the basis for our work to predict the aberrated PSF over an anisoplanatic region.

4 Artificial Neural Networks

Artificial neural networks are used for blackbox modelling when an analytical model is either not able or is too complex to be practically realisable. In this regard, ANNs do not necessarily model the internal physical mechanisms but are trained to support the input/output behaviour of a target system [8].

McGuire et al., provides extensive background on ANNs and details how such networks have been applied in the field of adaptive optics [9].

4.1 Applications

ANNs have demonstrated their practicality in the following applications:

1. Noise reduction: ANNs are very effective in recovering signals from noisy data, particularly for image processing applications [10].
2. Estimation/Classification: Zernike polynomial coefficients are estimated either directly [11], or indirectly using dimensionally reduction [12], from image data.
3. Prediction: Single- and multi-step prediction is used as an alternative to Kalman filters [13] to compensate for computational delays in wavefront sensing and servo lag due to temporal decorrelation of the atmosphere.

4.2 Architecture

For optimal performance, in terms of MSE, ANN architectures are application dependent and highly sensitive to parameters, such as the number of inputs and hidden node configurations that define their architecture. For example, given a discrete time series, $u(n)$, where $n = 0, 1, \dots N$, predicting the value of u, one sample period into the future requires the ability to store and access time history of input data. This is expressed as

$$u(n+1) = f(u(n), u(n-1), \dots, u(n-N)). \quad (7)$$

The simple delay-line architecture described in Equation 7 is referred to as a time delay neural network (TDNN). Recurrent neural networks (RNNs) also satisfy this requirement with the use of feedback, however training algorithms have been difficult to implement.

Recently, echo state networks (ESNs) have been proposed that simplify training of RNNs [14] and can be optimised using a signal processing interpretation and metric to describe the richness of their dynamics [15].

4.3 Echo State Networks

Echo state networks are recurrent, sparsely connected, artificial neural networks. The distinguishing features of ESNs as RNNs includes their random, fixed recurrent weights, damped dynamics, and simple training algorithm.

Given a sampled data series $\mathbf{u}(n)$, where $n = 1 \dots N$, an ESN can be defined in terms of the previously updated state vector, $\mathbf{x}(n\text{-}1)$, current state vector $\mathbf{x}(n)$, and predicted output, $\mathbf{y}(n)$

$$\begin{aligned} \mathbf{x}(n) &= \tanh(\mathbf{W}_{\text{in}}\,\mathbf{u}(n)^{\text{T}} + \mathbf{W}_{\text{DR}}\,\mathbf{x}(n\text{-}1)^{\text{T}}), \\ \mathbf{y}(n) &= \mathbf{W}_{\text{out}}\,\mathbf{x}(n)^{\text{T}}, \end{aligned} \quad (8)$$

where $\mathbf{W}_{\text{in}}$ is the input weight matrix, $\mathbf{W}_{\text{DR}}$ is the dynamic reservoir matrix, and $\mathbf{W}_{\text{out}}$ is the output weight matrix.

The predicted output of the input series, $\mathbf{u}(n)$, one time-step into the future, is given by

$$\hat{\mathbf{u}}(n{+}1) = \mathbf{y}(n). \quad (9)$$

ESNs were considered for the prediction of the space-varying PDF for the following reasons:

1. Training of an ESN is simplified; a simple linear regression is required for each output node.
2. Due to their simplified architecture, efficient hardware implementations should be possible.
3. Dynamic optimisation of the network may be more easily facilitated by the simplicity of the architecture.

Various studies on the design [15] and architectural enhancements [16] of ESNs have been conducted.

4.4 Training

ANNs require training to minimise weight vectors to ensure convergence and provide generalised performance. Batch training algorithms pass all the training data through the network prior to use in an application; on-line algorithms require one

training pattern, such as the previous data sample, and update the weight vectors before processing the next input.

In practice, the gradient decent algorithm is commonly used for on-line training; batch mode training typically employes linear regression or the Moore-Penrose pseudoinverse. Advantages and disadvantages of both training methods can be summarised as a trade-off between training time and MSE. The principle requirement of any regression algorithm is to ensure that ill-posed conditions, such as outlined in Section 2, are minimised or circumvented.

ESNs use a simple least-squares algorithm to train only the output nodes; the reservoir is not trained and relies on the spectral radius (the largest eigenvalues of the reservoir matrix) as the basis for selection of design parameters [15].

5 Architectural Overview

A concurrent system architecture for the prediction of wavefront aberrations, employing M point source functions (PSF) over a wide FOV, is shown in Figure 2. This model provides a basis for real-time AO operation and may be considered for wide-field adaptive optics and modal tomography [7].

To extend θ_0 beyond 40 arcseconds [9] a system architecture is described as follows. Two high-resolution, high-speed CCD cameras are used to continuously capture image frames of multiple NGSs (each considered an aberrated PSF), and a target object, over a wide spatial field. Phase retrieval [17] or an ANN classifier [12] is used to estimate wavefront aberrations from multiple source objects using defocused (intra- & extra-focal) images. To support the demands of a real-time system, pipelined image acquisition and preprocessing modules are used. Two sets of these modules (one per image) are shown on the right of image planes I_1 and I_2 in Figure 2.

A region of interest (ROI) surrounding each source and target is determined, *a priori*. Each source represents a point-spread function (PSF) and the image aberrations resulting from the adverse effects of the atmosphere are estimated, as described in the previous paragraph. To improve performance, a set of parallel modules (estimators) are used (one module per source) to define the set of Zenike polynomials, $Z_{2...20}$, that represent aberrations effecting each PSF. The set of estimators, E_1 to E_M, is shown to the center right of Figure 2.

Lastly, a modal predictor, P, employs an echo state network (ESN) to predict turbulence surrounding a target object, T_{Est}. The structure of a suitable predictor module is shown on the right of Figure 2. The set of Zenike coefficients $Z_{2...20}$ for each source $S_{1...N}$ is used as time-series input to determine aberrations surrounding a target region.

Additional parameters, such as angular separations of source and target regions (not shown in Figure 2), are used to predict modal expansions $Z_{2...20,\,T_{Est.}}$, over a constrained, anisoplanatic region.

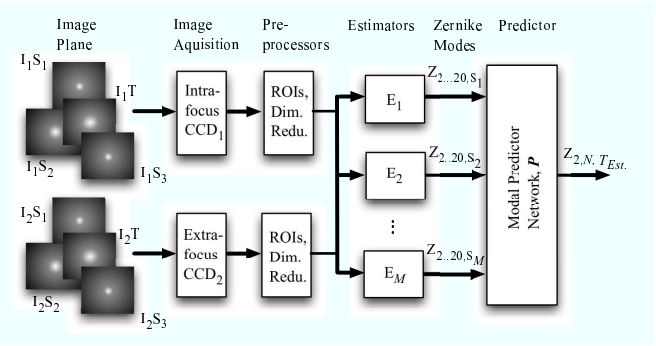

Figure 2: Proposed ANN System Architecture.

6 Simulations

Diffraction limited imaging systems typically employ Fourier optics to characterise an optical system [18]. Such systems are considered linear and the point spread function (PSF) is used extensively to define inverse problems using either incoherent and coherent light [19].

Thus, given multiple source objects with planer wavefronts i.e., NGSs in the absence of turbulence, an optical system will image PSFs for each source object within the FOV. After considering the effects of Fresnel diffraction and the required wavefront sensor (WFS), two slightly defocused images were produced. Each image showed multiple, diffraction-limited sources in the absence of turbulence.

Simulations were developed in Matlab and two 1024 $\times$ 1024 focal plane images were produced, using a simulated shutter speed of T_s milliseconds, and a frame rate of f_r FPS, where $T_s < \frac{1}{f_c}$. The imaging model is based on a Cassegrain optical telescope of 1 m diameter, with focal length of 7.7 m, supporting two, 4.7×4.7 mm CCDs. This provided a FOV of approximately 2.1 arc-minutes.

Independent and movable phase-screens, employing Komogorov statistics, representing one or more atmospheric layers at a height of $0 \ldots Z$ kilometers, were generated using the fractal method [20]. Planar wavefronts from M NGSs were convolved with each layer resulting in highly aberrated sub-images. Examples of non-aberrated sub-images are shown on the far left of Figure 2.

Predefined regions of interest, representing the simulated image of each NGS were taken and pro-

cessed to extract aberrations, in terms of Zernike coefficients of order $Z_{2...20}$. At each time-step the phase-screen is moved over each sub-image, at a velocity of V_p meters per second - Taylor's frozen turbulence hypothesis was assumed [21]. A curvature WFS [17], shown in Figure 2 as $E_{1...M}$, was used to estimate M sets of Zernike coefficients, $Z_{2...20}$, from sub-images $I_{1,2}$, and sources $S_{1...M}$.

The capture of several thousand frames, and subsequent estimation of Zernike coefficients from aberrated point-spread functions, $S_{1...M}$, resulted in time-series data, $Z_{2...20,\,S_{1...M}}$. To demonstrate in principle the effectiveness of these methods, only low-order *Tilt* terms, Z_2 and Z_3, were used in our analysis. An example of the resulting Zernike polynomial time-series ensembles for the *Tilt* term, Z_3, is shown in Figure 3. These data series, $S_1 \ldots S_3$, were used to train an ESN using batch mode training for prediction of a corresponding aberration, Z_3, adversely affecting a target PSF, also shown in Figure 3. The ESN module used to predict the target aberration is shown as the modal predictor network, P, in Figure 2.

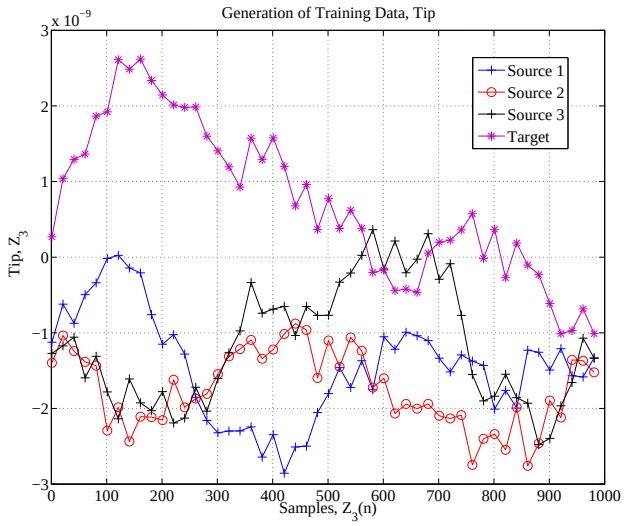

Figure 3: Simulated Low-order Aberrations.

7 Results

A set of results from a low-order PSF target aberration is shown in Figure 4. These results represent an extract of 26 samples from a time-series generated using three source objects and orientated in the configuration of an equilateral triangle.

Test data were collected using 30 random angular separations within the anisoplanatic region, with a maximum of 50 and minimum of 20 arcseconds separation between target, T and source objects, $S_{1...3}$. Batch mode training employed the Moore-Penrose pseudoinverse.

Each sample run used 1024 iterations from a moving phase screen. The results from over 30 simulations were used to provide a statistical average of $Z_{3,\,T_{Est.}}$.

The predicted results of the tilt aberration $Z_{T_{Est.}}$, of a space-varying target PSF, given input aberrations Z_3, and angular separations $\theta_{1...3}$ between each source $S_{1...3}$ and target object T, are compared to actual tilt aberrations $Z_{3,\,T_{Act.}}$ in Figure 4(a). A corresponding error plot for a portion of the original time-series is shown in Figure 4(b). The MSE over this portion was 0.0497.

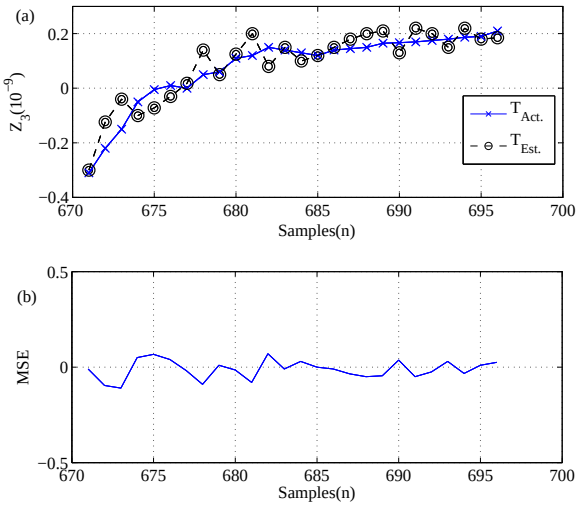

Figure 4: (a) Z_3: Actual ($T_{Act.}$) & Predicted ($T_{Est.}$) (b) Error Plot.

Currently, we are optimising for MSE using various orientations of source objects over aniosoplanatic regions and multiple turbulence layers. Additive photon and CCD read noise profiles have also been considered.

8 Conclusion & Future Work

In this paper we have described a system architecture suitable for demonstrating single time-step prediction of low-order aberrations from anisoplanatic target objects. A more thorough analysis, however, in terms of ESN optimisation and orientation of source objects is required to complete this study.

A hardware implementation within an FPGA of each of the modules discussed in Sections 4.2 and 4.3, is being investigated for real-time application.

9 Acknowledgements

The author's wish to thank Theam Yong Chew for the use of his curvature wavefront sensor routines and Dave van Leeuwen for his assistance with the Electrical & Computer Engineering *Beowolf Cluster* to complete our simulations.

References

[1] R. Tyson, *Principles of Adaptive Optics*. Academic Press, 1991.

[2] P. McGuire, D. Sandler, M. Lloyd-Hart, T. Rhoadarmer, A. of, N. Lecture, and i Physics, "Adaptive optics: Neural network wavefront sensing, reconstruction, and prediction," 1999.

[3] R. Noll, "Zernike polynomials and atmospheric turbulence," *Journal of the Optical Society of America*, vol. 66, no. 3, 1976.

[4] S. Haykin and J. Principe, "Using neural networks to dynamically model chaotic events such as sea clutter," in *IEEE Signal Processing Magazine*, I. Press, Ed., May 1998, pp. 66–81.

[5] M. Aubailly, M. Roggerman, and T. Schulz, "Prediction of space-varying point spread function for reconstruction of anisoplanatic adaptive optics images," in *Proceedings, Astronomical Adaptive Optics Systems and Applications II*, T. I. S. for Optical Engineering SPIE, Ed., vol. 5903, Sep 2005, pp. 59–64.

[6] S. Weddell and R. Webb, "Dynamic artificial neural networks for centroid prediction in astronomy," in *Sixth International Conference, Hybrid Intelligent Systems (HIS'06), Auckland, New Zealand*, I. Press, Ed. IEEE Press, Dec 2006, pp. 68–72.

[7] R. Ragazzoni and E. Marchetti and G. Valente, "Adaptive Optics Corrections Available for the Whole Sky," *Nature Letters*, vol. 403, no. 6765, pp. 54–56, 1999.

[8] H. H.A. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, pp. 78–80, 2004.

[9] P. McGuire, D. Sandler, M. Lloyd-Hart, and T. Rhoadarmer, "Adaptive optics: Neural network wavefront sensing, reconstruction, and prediction," *Scientific Applications of Neural Nets, Proceedings of the 194th W. E. Heracus Seminar, 1998. Edited by J.W. Clark, T Lindenau, and M.L. Ristig, Springer-Verlag Publishers*, p. 97, 1999.

[10] D. Mandic and J. Chambers, *Recurrent Neural Networks for Prediction - Learning Algorithms, Architectures and Stability*. John Wiley & Sons Ltd., 2002.

[11] G. Vdovin, "Model of an adaptive optical system controlled by a neural network," *Optical Engineering*, vol. 34, no. 11, pp. 3249–3253, 1995.

[12] G.S. Chundi and M. Lloyd-Hart and M.K. Sundareshan, "Training Multilayer Perceptron and Radial Basis Function Neural Networks for Wavefront Sensing and Restoration of Turbulence-Degraded Imagery," *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 3, pp. 2117–2122, 2004.

[13] S. Goh and D. Mandic, "An augmented extended kalman filter algorithm for complex-valued recurrent neural networks," *Neural Computation*, vol. 19, no. 4, pp. 1039–1055, 2007.

[14] H. Jaeger, "Adaptive nonlinear system identification with echo state networks," 2003. [Online]. Available: citeseer.ist.psu.edu/jaeger03adaptive.html

[15] M. Ozturk, D. Xu, and J. Principe, "Analysis and design of echo state networks," *Neural Computation*, vol. 19, pp. 111–138, 2007.

[16] R. Webb, "A cascade architecture for online predictive echo state networks," *Proceedings of the 2007 International Conference on Machine Learning: Models, Technologies & Applications (MLMTA'07)*, 2007.

[17] T. Chew and R. Lane, "Estimating phase aberrations from intensity data," *Image and Vision Computing NZ 2003 (IVCNZ'03)*, 2003.

[18] J. Goodman, *Introduction to Fourier Optics*. McGraw-Hill, 1968.

[19] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. Institute of Physics (IOP), 1998.

[20] R. Johnston, "Inverse problems in astronomical imaging," Ph.D. dissertation, Dept. of Electrical & Computer Engineering, University of Canterbury, 2001.

[21] R. Tyson, "Introduction to adaptive optics," *Optical Engineering*, vol. TT41, 2000.