

Data Preprocessing on Sequential Data for Improved Astronomical Imaging

S.J. Weddell, and R.Y. Webb

University of Canterbury, Dept. Electrical & Computer Engineering.

Email: steve.weddell@canterbury.ac.nz

Abstract

Obtaining diffraction limited astronomical images using ground-based telescopes and sensors operating at optical wavelengths is currently not possible due to the effects of atmospheric turbulence. Adaptive optics is a mature technology that attempts to minimise the effects of atmospheric turbulence through wavefront detection and optical path correction in real-time. Our work involves the acquisition and preprocessing of sequential astronomical data and the application of machine learning algorithms to obtain consistently, high-resolution images called *lucky* frames. This paper will outline the design of a framework used to capture and preprocess sequential data in real-time, and explore machine learning extensions to accurately predict *lucky* frames.

Keywords: digital image coding, machine learning, wavefront detection, adaptive optics

1 Introduction

A framework of software tools is required for research in the prediction of image distortion in real-time. Images of astronomical objects experience distortion due the movement of air in the atmosphere [11]. While adaptive optics (AO) aims to correct image distortion in real-time by controlling the optical path of an instrument, our approach is to predict and separate high- and low-distortion images for concurrent processing.

Frame separation, based on image distortion, serves two purposes. Firstly, a series of high-resolution real-time images can be produced by capturing image frames during periods of low turbulence. Secondly, images that have been distorted by atmospheric turbulence can be used independently to detect and correct for wavefront aberrations.

In turbulent atmosphere, highly resolved image frames can be captured if exposure times are reduced to a minimum. This is essentially due to freezing the image during brief periods when the image wavefront is planar with the image sensor. The term we use to describe such a highly fortunate event is a *lucky* frame.

Image frames, distorted by atmospheric turbulence, can be used to determine wavefront aberrations. We term such images, *AO* frames. An adaptive optics system will typically process every frame, including *lucky* frames. Since *lucky* frames contain minimum information on wavefront aberrations, efficiencies, in terms of

data processing, could be gained if *lucky* frames were removed from the data stream, prior to processing.

A background on atmospheric turbulence and adaptive optics is given in the following section. The architecture of *lucky* frames is discussed in Section 3. Section 4 provides a system overview. The problem space is outlined in Section 5 and this is followed by an brief discussion on our requirements for machine learning in Section 6. To conclude, a summary of our work and future considerations is given in Section 7.

2 Background

Atmospheric turbulence severely limits the performance of ground-based telescopes. Adaptive optics [16] is used on ground-based telescopes to ultimately achieve diffraction limited images. Employment of AO technology provides a significant improvement in image resolution, however, this is currently limited to infrared wavelengths due to atmospheric absorption [2]. The Hubble Space Telescope (HST) is capable of providing near diffraction limited images over a wide electromagnetic spectrum, including optical and ultraviolet wavelengths.

In spite of current limitations within the visual spectrum, the refinement of AO systems has continued to improve [11]. In specialised fields, such as spectroscopy, infrared and visual performance of ground based telescopes using *extreme* adaptive optics has exceeded the Hubble

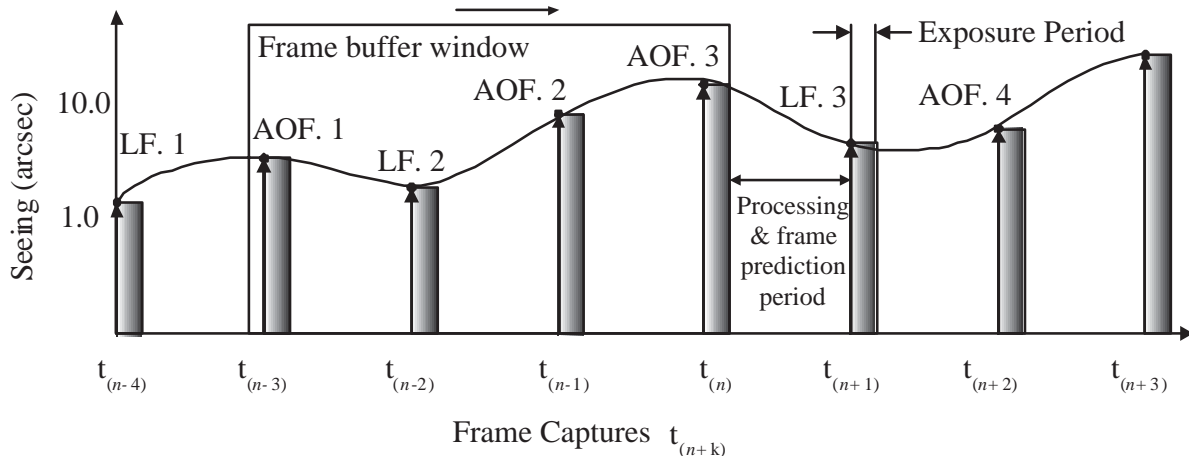


Figure 1: Prediction of *Lucky* Frames

Space Telescope (HST) [2]. Advances, in terms of sensor technology [13], improved control algorithms [1], and more recently, wavefront prediction using machine learning [11], are demonstrations of further enhancements to AO systems. New fields in adaptive optics, such as multi-conjugate adaptive optics (MCAO) [4] and wide-field adaptive optics (WFAO) [3], that utilises the tomography of atmospheric turbulence [14], will further enhance the performance of AO over the next decade.

3 Lucky Frames Architecture

A *lucky* frame is a single, high-resolution, minimally distorted (blurred) image, the occurrence of which is predicted by the analysis of preprocessed data, comprising of temporal and spatial information, representing astronomical science objects.

Given a fixed frame buffer representing preprocessed image data, our aim is to accurately predict the period between frame captures, when atmospheric turbulence is at a minimum.

During periods of low turbulence, image resolution will be optimal. High resolution images will be captured during such periods and form a series of *lucky* frames. It is proposed that by separating high- and low-resolution images, using wavefront prediction [11], a reduction in processing time may be achieved.

Three stages are required to accurately predict *lucky* frames. Firstly, a series of images is captured, preprocessed and stored within a frame buffer. Secondly, a machine learning algorithm is used to classify the preprocessed data and will determine if the next image will be either a *lucky*

or *AO* frame. The process required to accurately predict the next n -frame types is more complex than determining local minima and maxima. The effects of turbulence, based on three or more astronomical objects within a wide field-of-view, and employing *a priori* classification data, will be used as an improved method of prediction.

Image data from *lucky* frames can be used to provide high resolution imaging. Image data from *AO* frames can be used to extract information for wavefront reconstruction in real-time. Results from frame classifications are used to calibrate camera settings for optimum capture of the next frame. By employing two cameras, imaged through a 50/50 beam-splitter, individual camera setting could be optimised for recording high-resolution i.e., *lucky*, and distorted i.e., *AO* frames, using dedicated cameras. Camera settings that can be optimised in real-time would include frame rate, exposure time, and image gain.

Figure 1 shows our approach in predicting *lucky* frames. At time $t_{(n)}$, prediction of the next frame $t_{(n+1)}$ is based on a classification of the previously detected *lucky* frame, $t_{(n-2)}$ and *AO* frames, $t_{(n)}$, $t_{(n-1)}$, and, $t_{(n-3)}$. Frame preprocessing of the most recently acquired frame, and classification using preprocessed data within the frame buffer, will allow $t_{(n+1)}$ to be accurately predicted as a *lucky* frame.

To demonstrate the application of *lucky* frames, two hundred continuous images of Alpha Centauri were captured at the prime focus of a 20 cm, Cassegrain telescope. Due to a combination of low viewing angle (60 Degrees from Zenith), and a turbulent atmosphere, the seeing conditions over the

science object varied between an estimated 2, to 20 arcseconds.

From the images captured, two images were manually selected, four frames, or 133 *ms* apart. The first of these images, Figure 2, shows Alpha Centauri as a binary star with a separation of 13 arcseconds; each pixel within the image represents 1.5 arcseconds. This image was recorded with minimum distortion due to low atmospheric turbulence between the science object and the imaging device, throughout the duration of the exposure period. An image such as this would be classified as a *lucky* frame and is labelled in Figure 1 as *LF.1*.

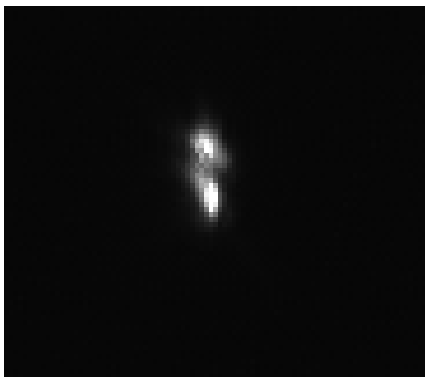


Figure 2: A Lucky Frame (*LF.1*) of α Cen

A second frame, taken 133 *ms* after the first, is shown in Figure 3. Distortion, primarily due to atmospheric turbulence during image capture, is clearly evident. Using our terminology, this image would be classified as an *AO* frame. A corresponding *AO* frame is shown in Figure 1 and labelled *AOF.2*.

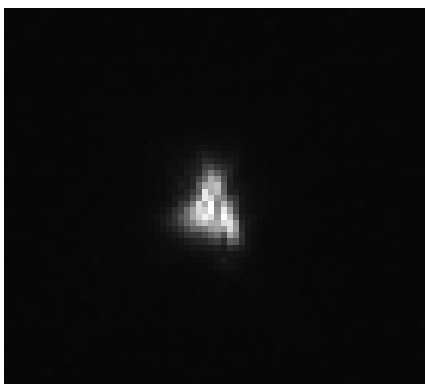


Figure 3: A *AO* Frame (*AOF.2*) of α Cen

4 System Overview

A frame buffer was required to capture and preprocess a series of images in real-time. The function of

the preprocessor is to decompose each image into a series of primitive data objects. A framework, comprised of functions and data objects, will be required. To provide accurate prediction of high-resolution images from sequential data, a machine learning module will be employed. A control module is required to synchronise camera functions and coordinate the selection and storage of image data. An overview of our proposed system is shown in Figure 4.

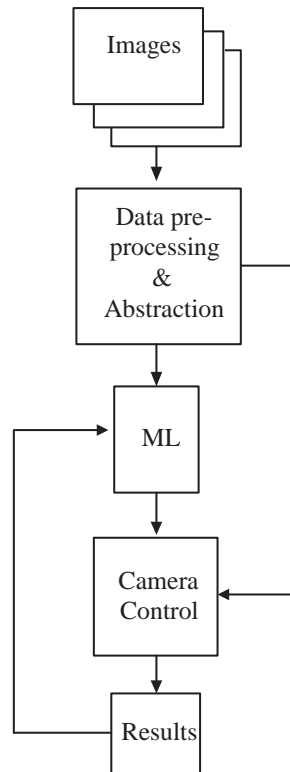


Figure 4: System Overview

Source data from the capture of a astronomical image will be processed at an initial frame rate of 30 frames-per-second (FPS). The data objects requiring analysis, comprise the science object and three or more guide stars. A tomographic approach [14], over a wide field-of-view (FOV) will be employed to predict the occurrence of high-resolution images.

Hickson [6] used a single-pupil curvature sensor, on slightly defocused images, to retrieve wavefront maps. Ragazzoni used tomography to increase the corrected FOV over a science object and four guide stars using images from a single-pupil curvature sensor. By employing Fourier transforms, the Poisson propagation equation can be solved for wavefront reconstruction [14]. Such methods are computationally intensive and as a result have yet to be implemented in real-time on a closed loop system.

To adopt a tomographic approach over a wide FOV, computational demands need to be reduced [5]. Our method will employ machine learning to predict images at the focal plain that have undergone minimum distortion and separate them into two, concurrent processing streams. Signal processing techniques, such as double or *ping-pong* buffering, will be used to maximise processing efficiency.

A description of the program, and core routines required to implement frame prediction is provided in Listing 1. System initialisation fills the frame buffer with *MaxFrames* image data while preprocessing each frame into a series of data objects. These core routines are also used in the main loop to maintain the image buffer. The function *PredictLuckyFrame*, will call control routines that will subsequently be used to synchronise camera primitives. The frame buffer will load and preprocess a new frame for each iteration of the loop, discarding data from the oldest preprocessed image frame. This process is commonly referred to as *circular buffering* [7], and is used extensively in digital signal processing.

```

\\ initialisation
for i ← 1 to MaxFrames do
    CaptureFrame(); \\ get frame data
    PreprocessFrame(); \\ preprocess frame
    StoreDataBuffer(); \\ & store data
end

\\ process frame buffer
do
    ProcessBufferData(); \\ Buffer process
    PredictLuckyFrame(); \\ ML routines
    CaptureFrame() \\ get next frame
    PreprocessFrame(); \\ preprocess &
    StoreDataBuffer(); \\ store frame data
while

```

Listing 1: Proposed System Structure

An API framework is used to control two Micropix Firewire cameras. Two cameras will allow higher frame rates to be achieved and provide independent streaming of data, based on their frame type. Independent triggering of both cameras will be controlled by a ML module. Optimum exposure times and frame capture periods between cameras will be determined by the ML module. A similar equipment configuration is currently in use for SCIDAR research at Mt John observatory [10], [12].

5 The Problem Space

Creating an effective problem space establishes a platform to refine our model, given the availability of reusable data objects. Mapping the problem space to a solutions space allows developers to evaluate results based on the application of alternative design methods. Abstraction from implementation detail, such as storage capacity, data collection, parsing, indexing and encoding of data objects for example, will establish the environment for application of a predictive algorithm.

5.1 The Data Preprocessor

The data preprocessor module, outlined in Sections 3 and 4, must be capable of acquiring up to n image frames from M cameras and process data within period P . Period P is bound by the minimum frame rate FP , supported by the slowest camera. The minimum acquisition and processing period, P , is given by,

$$P = \frac{1}{M.FP} 10^{-3} \text{ ms} \quad (1)$$

The maximum number of cameras, M , that can be used in this system is related to the exposure period, E . For example, if we assume $M = 2$ and $FP = 33.3 \text{ ms}$, then the minimum processing period is, $P = 15 \text{ ms}$.

Data reduction is performed by firstly determining the centroids of all data objects within the field of view, and secondly clipping the image of all astronomical data objects, such as guide stars, and the science object. In order to meet real-time processing requirements, a VLSI preprocessor is proposed. CCD image sensor preprocessors have been developed specifically for adaptive optics [13], however by combining both centroid and clipping functions, a significant reduction in preprocessing would be achieved. The data preprocessor module, outlined in Figure 4, is shown in Figure 5.

5.2 Specialised Functions

In addition to the system requirements outlined in Section 4, the implementation of several specialised functions will be required. These functions include:

1. Encoding and decoding routines to support data compression.
2. Track and report methods for dropped frames.
3. Random access to centroid & scintillation data.

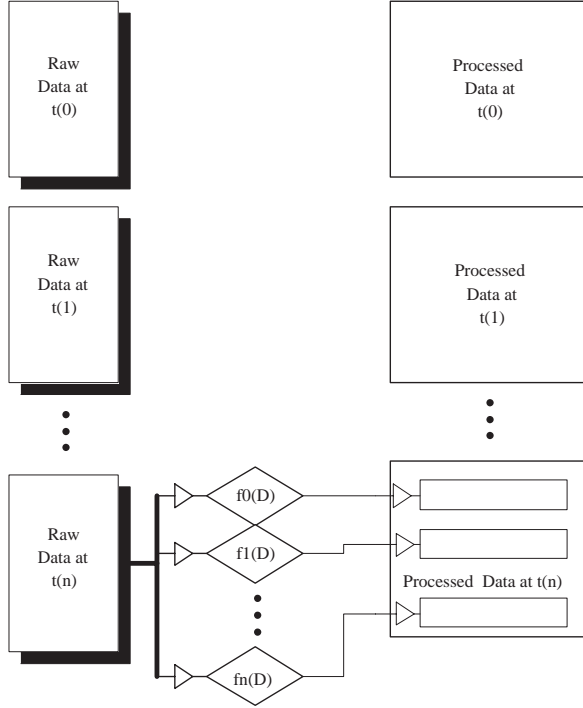


Figure 5: The Data Preprocessor

4. Image enhancement and interpolation routines.

Currently, Items 2 & 3 have been implemented. The employment of specialised, VLSI encoding algorithms, discussed in Section 5.1, is under development. Several approaches, including the application of the K-Maximum Sum algorithm [15] are being considered.

6 ML Requirements

Machine learning has been used to predict atmospheric turbulence [11], however, few classification models based on discriminative learning have been reported.

Bayesian maximum *a posterior* (MAP) estimation has been used in the blind restoration of adaptive optics telescope images [3], and provides a formal approach for manipulating non-deterministic, complex models by describing or extending a probability density function, over the variables in question [8]. Within this probability density function, partial knowledge can be specified *a priori* and a coarse model can be refined using empirical observations and data. This can be described using Bayesian methods as,

$$Posterior = \frac{Likelihood}{Evidence} \times Prior \quad (2)$$

Given the intuitive definition of Bayes' rule in

Equation 2, a major problem with generative based systems is that aspects of this prior knowledge may be inconsistent or incorrect [9].

In contrast to generative learning that use prior models as the basis for probabilistic learning, discriminative learning, used for example in artificial neural networks (ANNs) and support vector machines (SVMs) [8], adopts a mechanism that maps all input variables directly to system outputs. In the case of SVMs, data is mapped to a higher dimensional space by using the kernel method, and computing the largest possible margin separating the hyperplane [8]. Such direct mapping methods, however, are computationally intensive.

A predictive algorithm that can accurately select high resolution images from a continuous series of distorted images will need to be highly efficient in performance, and simple in terms of implementation. Current research in the classification of astronomical images, using ANNs and SVMs for reinforcement learning, is underway. A ML framework has been constructed [17] and will be used to further this requirement.

7 Conclusion and Future Work

This research was motivated by the design and realisation of an image processing framework that could support the application of a machine learning algorithm to predict the quality of images, and selectively process images in real-time. In this paper we have discussed the first stage of this project, data preprocessing on sequential data.

The second stage of this project will require the selection and implementation of a machine learning algorithm for the determination of *lucky* frames. In addition, preprocessing of image data in hardware will be required to reduce data processing demands on a real-time system.

Lastly, further research is required to reduce the computational complexity of reconstructing wavefront maps in real-time using tomography. It is our intention to extend this work in frame prediction, to the prediction of wavefronts using machine learning for application in wide-field adaptive optics.

References

- [1] N.J. Doelman, K.J.G. Hinnen, F.J.G. Stofelen & M.H. Verhaegen, "Optimal Control Strategy to Reduce the Temporal Wavefront Error in AO Systems", *Proceedings of SPIE*

- Annual Conference on Adaptive Optical System Technologies*, Vol. 5490, pp. 121-130, 2004.
- [2] H. Furguson, "Space-based vs. Ground-based telescopes with Adaptive Optics (AO)", emphA summary prepared by the staff of the Space Telescope Science Institute, The Association of Universities for Research in Astronomy (AURA), Inc., NASA contract NAS5-26555, 2003.
- [3] T. Fusco, J.-M. Conan, L.V. Michau & G. Rousset, "Efficient phase estimation for large-field-of-view adaptive optics", *Optics Letters*, Vol. 24, No. 21, pp. 1472-1474, 1999.
- [4] T. Fusco, J.-M. Conan, L.V. Michau, G. Rousset & F. Assemat, "Multi-conjugate adaptive optics: Comparison of phase reconstructions for large field of view", *Proceedings of SPIE - The International Society for Optical Engineering*, Vol. 4167, pp. 168-179, 2001.
- [5] D. Gavel, "Tomography for Wide-Field Adaptive Optics", Center for Adaptive Optics, CfAO Summer School for Adaptive Optics, <http://cfao.ucolick.org/aosummer/aosummer2005/pdf/Gavel-tomography.pdf>, visited on 15/8/2005.
- [6] P. Hickson, "Single-image wavefront curvature sensing", *Proceedings from SPIE Annual Conference on Adaptive Optical System Technologies*, Vol. 2201, pp. 549-554, 1994.
- [7] E.C. Ifeachor & B.W. Jervis, *Digital Signal Processing - A Practical Approach*, 2nd Ed., Pearson Education Ltd., Ch. 12, pp. 766-768, 2002.
- [8] T. Jabara, *Machine Learning Discriminative and Generative*, Kluwer Academic Publishers, 2004.
- [9] A.K. Johan, T.V. Gestel, J. De Brabanter, B. De Moor & J. Vandewalle, *Least Squares Support Vector Machines*, World Scientific Publishing Co. Pte. Ltd., 2002.
- [10] R.A. Johnston, J.L. Mohr, P.L. Cottrell & R.G. Lane, "A Bread-Board SCIDAR System at Mount John", *Proceedings from the Imaging and Vision Conference (IVCNZ), New Zealand 2004*, pp. 209-214, 2004.
- [11] P.C. McGuire, D.G. Sandler, M. Lloyd-Hart & T.A. Rhoadarmer, "Adaptive Optics: Neural Network Wavefront Sensing, Reconstruction, and Prediction", *Scientific Applications of Neural Nets, Proceedings of the 194th W. E. Heraeus Seminar, 1998*. Edited by J.W. Clark, T. Lindenau, & M.L. Ristig, Springer-Verlag Publishers, pp. 97, 1999.
- [12] J.L. Mohr, J.A. Johnston, R.G. Lane & P.L. Cottrell, "Adaptive Optics at Mt John Observatory", *Proceedings from the Image & Vision Computing Conference (IVCNZ), New Zealand 2004*, pp. 233-236, 2004.
- [13] B.H. Pui, B. Hayes-Gill, M. Clark, M. Somekh, C. See, J.F. Pieri, S. Morgan & A. Ng, "The Design and Characterisation of an Optical VLSI Processor for Real Time Centroid Detection", *Proceedings on the design, test, integration, and packaging of MEMS/MOEMS, Proceedings SPIE*, Vol. v4408, pp 73-80, 2001.
- [14] R. Ragazzoni, E. Marchetti & G. Valente, "Adaptive Optics Corrections Available for the Whole Sky", *Nature Letters*, Vol. 403, Issue 6765, pp. 54-56, 2000.
- [15] T. Takaoka, "A faster algorithm for the all-pairs shortest path problem and it's application", *Computing and Combinatorics, Proceedings Lecture Notes in Computer Science*, Springer-Verlag, Vol. 3106, pp. 278-289, 2004.
- [16] Tyson R. K. *Introduction to Adaptive Optics*, Optical Engineering Vol. TT41, SPIE Press, 2000.
- [17] R. Webb, *DURA: A Machine Learning toolbox*, Electrical & Computer Engineering, University of Canterbury, 2005.