Optics Communications 382 (2017) 519–527

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

A Bayesian regularized artificial neural network for adaptive optics forecasting

Zhi Sun [a,c,*], Ying Chen [b], Xinyang Li [b], Xiaolin Qin [a], Huiyong Wang [a,c,d]

[a] Chengdu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu 610041, PR China
[b] Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, PR China
[c] University of Chinese Academy of Sciences, Beijing 100049, PR China
[d] Guangxi Key Laboratory of Cryptography and Information Security, Guilin University of Electronic Technology, Guilin 541004, PR China

ARTICLE INFO

Article history:
Received 29 May 2016
Received in revised form
12 August 2016
Accepted 13 August 2016

Keywords:
Adaptive optics
Predictive controller
Back propagation network
Bayesian regularization
Parallel computing

ABSTRACT

Real-time adaptive optics is a technology for enhancing the resolution of ground-based optical telescopes and overcoming the disturbance of atmospheric turbulence. The performance of the system is limited by delay errors induced by the servo system and photoelectrons noise of wavefront sensor. In order to cut these delay errors, this paper proposes a novel model to forecast the future control voltages of the deformable mirror. The predictive model is constructed by a multi-layered back propagation network with Bayesian regularization (BRBP). For the purpose of parallel computation and less disturbance, we adopt a number of sub-BP neural networks to substitute the whole network. The Bayesian regularized network assigns a probability to the network weights, allowing the network to automatically and optimally penalize excessively complex models. The simulation results show that the BRBP introduces smaller mean absolute percentage error (MAPE) and mean square errors (MSE) than other typical algorithms. Meanwhile, real data analysis results show that the BRBP model has strong generalization capability and parallelism.

1. Introduction

Adaptive optics (AO) [1] is an indispensable technology in large telescopes to improve image quality, degraded due to atmospheric disturbances. The major components involved in a simple AO system are deformable mirrors (DM) inserted in the telescope's optical path and measurements provided by a wavefront sensor (WFS), as well as a control algorithm. WFS measures the error in fitting the DM to the atmospheric distortions in the pupil plane. In the past 20 years, kinds of AO systems have been used to correct distorted wavefront in many domains, especially astronomical observation. An AO system can be regarded as a real-time servo system [2,3]. From detection to correction, there is usually a fixed delay time of 2–3 cycles. The delay time is caused by wavefront sensor, reconstruction calculation, control algorithm [4] and so on. Error caused by delay time is a key cause on how the AO system performs.

Atmospheric turbulence through a telescope or a Hartmann sensor (HS) is a nonstationary Gauss stochastic process [5]. In particular, an advanced wavefront controller based on forecasting technology can help reduce the servo-lag (residual atmosphere) error. Several groups have addressed these problems, and have begun to do experimental work at the telescope. In addition, it has shown that the characteristic of the phase difference between two points is a chaotic function of time, thereby suiting to short-term prediction. Some numerical simulations in which the effect of the delay in the servo loop is reduced by using history data vector from the wavefront sensor to predict the value of the vector sometime in the future, when the phase correction is actually applied to the DM. Jorgenson and Aitken [6] predicted the wavefront slope by using the back propagation network (BP) and Wild [7] did it by using the recursive least square (RLS).

Empirical results suggested that the probabilistic neural network (classification model) outperforms the standard BP neural network (level estimation model) in predicting time-delay systems. However, there are still some drawbacks [8] in the classical model: slow training convergence, easy to fall into local optimum, etc. To address the potential overfitting of neural network weights, some researchers have developed hybrid neural network techniques [9,10]. The results of their study showed significant improvement over other standard models, while Bayesian regularization [11–13] can overcome these shortcomings and improve the

* Corresponding author at: Chengdu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu 610041, PR China.
E-mail address: china.sunzhi@gmail.com (Z. Sun).

http://dx.doi.org/10.1016/j.optcom.2016.08.035
0030-4018/

generalization capability of network.

In this paper, the BP neural network coupled with Bayesian regularization (BRBP) is introduced as a novel hybrid model to forecast DM voltages. Meanwhile, in order to further improve the performance of the BRBP method, we also make it work in parallel mode [14], according to the prediction process. The performance of the BRBP is tested by two different AO systems and compared to other advanced hybrid models, such as the steepest descent algorithm BP neural network (SDBP), the BFGS quasi-Newton method BP neural network (BFGBP) and the Levenberg–Marquardt method BP neural network (LMBP). Simulation results show that the BRBP offers an enhanced level of performance, and the BRBP can promote AO system performance.

2. Principle of DM voltage forecasting

The concept of Taylor frozen flow is a good approximation on short time scales ($t \ll \tau$, where τ is the atmospheric coherence time) for simulation single layer atmospheric turbulence. With the assumption of frozen flow, wind from some preferred directions will blow a static wavefront of aberration across the telescope aperture. Although it is within the decorrelation time, this would introduce significant inaccuracy in the wavefront correction process. A possible partial solution to this problem is to progressively predict wavefronts arriving after a delay time equivalent to the servo time lag. Prediction technique in an AO system consists of two main types: gradient-based and voltage-based.

According to Taylor's hypothesis [15] of atmospheric turbulence, we know that the relative spatial structure of atmospheric turbulence remains unchanged in a short time. At this moment, the change of the turbulence is caused by the transverse wind. And this information can be directly reflected in the gradients of a certain subaperture and its neighboring subapertures. In such a case, it makes much sense to predict the gradient based on the earlier gradients over the subaperture of interest and the neighboring subapertures. In an AO system which adopts the gradient-direct reconstruction algorithm, there is a direct linear relation between gradients and voltages:

$$V = R_{xy}^{+}G \tag{1}$$

where R_{xy}^{+} is the transfer matrix and G represents the gradient. Considering the linear relation, the control voltage of DM can also be used in prediction.

An adaptive optics system with a prediction controller is shown in Fig. 1. The system includes a Shack–Hartmann sensor (HS) that measures the gradient signal, a wavefront controller (CC) based on proportion integration (PI) method that produces control voltages for driving the DM, a prediction controller (P) based on BRBP that generates prediction voltages for improving the performance of the system and the DM used for generating surface shape. The reconstruction voltages are obtained by the gradient-direct method. Wavefront controller and prediction controller are used to transform the reconstruction voltage to the control voltage loaded on the DM. Due to the existing 2–3 cycles delay time, the prediction controller can use history voltages belonging to the actuator and actuators around it to get a next certain voltage. The frames of history voltages that give the best predictability of the near future is one cause that needs optimization in DM voltage prediction.

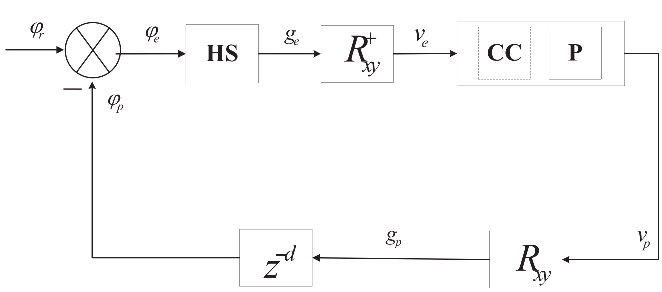

Fig. 1. Block diagram of AO closed-loop predictive control system based on the BRBP.

3. DM voltage forecasting based on BRBP

The linear prediction method is usually used to predict DM voltages in stationary situation. But it is not suitable for a non-stationary state, especially when the speed of transverse wind is very high. When in a non-stationary state, it is difficult to give a definite evaluation and find a concrete function expression. A major issue for these prediction techniques is the potential of overfitting and overtraining. To eliminate the potential for overfitting, a mathematical technique known as Bayesian regularization was developed.

3.1. Linear forecasting method

The RLS method [16] is usually used as a classic linear forecasting method to predict DM voltages. It enjoys a faster convergence speed than the least mean squares (LMS) algorithm. Now we describe the theory of the DM voltage prediction based on RLS linear method briefly. We define $\mathbf{V}_f$, $\mathbf{V}_h$ as future voltages and history voltages, respectively. On the assumption that there is a linear relationship between $\mathbf{V}_f$ and $\mathbf{V}_h$

$$\mathbf{V}_f = [\mathbf{W}_i, \mathbf{W}_{i-1}, \ldots, \mathbf{W}_1]\cdot\mathbf{V}_h + \mathbf{e} \tag{2}$$

let $\mathbf{W} = [\mathbf{W}_i, \mathbf{W}_{i-1}, \ldots, \mathbf{W}_1]$, $\mathbf{V}_f = \mathbf{W}\cdot\mathbf{V}_h + \mathbf{e}$, and solve the weighting matrix $\mathbf{W}$ by RLS. In order to get the minimum error, $S(\mathbf{V}_h)$ is obtained by LMS error criterion:

$$S(\mathbf{W}) = \mathbf{e}^T\mathbf{e} = (\mathbf{V}_f - \mathbf{W}\mathbf{V}_h)^T(\mathbf{V}_f - \mathbf{W}\mathbf{V}_h) \tag{3}$$

Hence $S(\mathbf{W})$ is given by:

$$\mathbf{W} = (\mathbf{V}_h^T\mathbf{V}_h)^{-1}\mathbf{V}_h\mathbf{V}_f \tag{4}$$

We take Eq. (4) into Eq. (2), where the value of $\mathbf{V}_f$ is the desired prediction voltage.

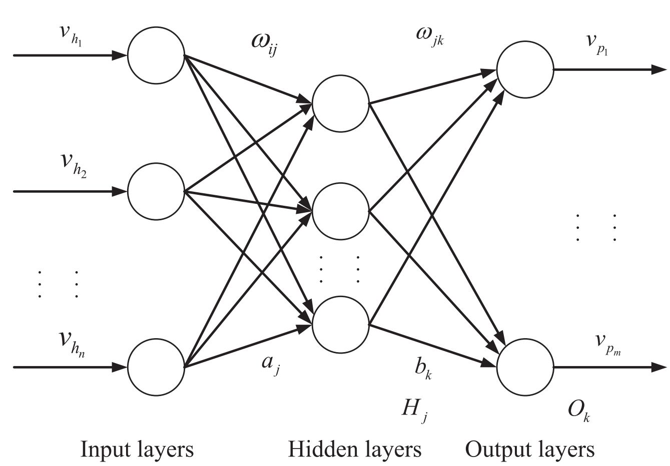

Fig. 2. Three-layer BP neural network for DM voltage prediction.