

# Optics Letters

## Piston alignment of segmented optical mirrors via convolutional neural networks

DAILOS GUERRA-RAMOS,<sup>1,\*</sup> LARA DÍAZ-GARCÍA,<sup>2</sup> JUAN TRUJILLO-SEVILLA,<sup>2</sup> AND JOSE MANUEL RODRÍGUEZ-RAMOS<sup>2,3</sup>

<sup>1</sup>Universidad de La Laguna, C/Padre Herrera s/n 38200, La Laguna, Tenerife, Canary Islands, Spain

<sup>2</sup>Wooptix, S.L. Avda. Trinidad 61. 38204, La Laguna, Tenerife, Canary Islands, Spain

<sup>3</sup>Cibican. Campus Ciencias de La Salud s/n, E-38071 La Laguna, Spain

\*Corresponding author: dailos.guerra@gtc.iac.es

Received 4 June 2018; revised 4 August 2018; accepted 6 August 2018; posted 7 August 2018 (Doc. ID 334328); published 29 August 2018

Most of the methods used today for the alignment of segmented mirrors are based on Shack-Hartman wavefront sensors. Other proposed methods are based on curvature sensors. These can be used to cross-check the measurements given by the primary method. We investigate a different approach which employs convolutional neural networks. This technique allows the piston step values between segments to be measured with high accuracy, as well as a large capture range at visible wavelengths. The technique does not require special hardware, and is fast to be used at any time during the observation. © 2018 Optical Society of America

**OCIS codes:** (100.4996) Pattern recognition, neural networks; (100.5070) Phase retrieval; (120.5050) Phase measurement; (120.6085) Space instrumentation.

<https://doi.org/10.1364/OL.43.004264>

In spite of its multiple advantages, segmentation also has some inconveniences caused by the discontinuities in the reflective surface that need to be addressed. In order to prevent the degradation of the quality of the image, phase differences between segments have to be reduced to a small fraction of the wavelength used.

Some methods proposed for cophasing segmented mirrors use microlenses or prisms that must be aligned with high accuracy over each of the edges of the segments [1]. This strategy becomes less practical for a growing number of segments; hence, a simpler and more robust method is preferred. Piston detection based on curvature sensors was first proposed in Ref. [2]. These methods that rely on one or two defocused images of the pupil correspond to a much simpler optical design. However, they either lack ample capture range [3] or depend strongly on atmospheric seeing [4]. Lastly, other efforts have been done to recover the tip-tilt value of segmented mirrors using a geometric sensor [5]. We propose the former method to work jointly with the piston detector that is described in this Letter.

The method presented here makes use of intensity measurements at visible wavelengths. To this end, a detector is placed at a single defocused plane, so no dedicated hardware is required. The method also contributes robustness and a broad capture range, while being insensitive to changes in the Fried parameter of the atmosphere. The defocus distance of choice is such that the atmospheric coherence diameter and the diffraction signal at the detector are in good agreement to get piston step information from the recorded data. Using monochromatic light, the differences in optical path lengths that are a number of waves away are indistinguishable by any interferometric method. The proposed algorithm is able to resolve this ambiguity through the use of different wavelengths and some machine learning strategies.

The versatility of machine learning makes it possible to solve problems that rely on finding patterns in structured complex data such as images. Convolutional neural networks (CNNs), in particular, are a powerful paradigm to represent solutions that map images into output labels. Hereafter, we will show that CNNs can be utilized for cophasing segmented mirrors. Neural networks (NNs) have been used in the past for wavefront sensing [6] and interferometry, although they have not been used for detecting local pistons between segments for cophasing segmented mirrors, i.e., phase discontinuity sensing and wrapping effect. CNNs do not discard spatial information, and the algorithm determines the appropriate features to extract from the image in order to get the best result.

This Letter is organized as follows: the mathematical and physical background of the problem is explained. Then the optical configuration and the way in which the data are simulated are described. After that, an explanation of the network used in this Letter and the procedure followed to train the network are detailed. Finally, we presented conclusions, future work considerations, and final remarks.

The interference produced by a single piston step can be thought of as a plane wave travelling along the positive  $z$  axis that is intercepted by a semi-infinite transparent screen at  $z = 0$ . The screen introduces a phase shift,  $\Delta\phi$ , in the wavefront at  $x < 0$ , where the screen exists. The normalized intensity distribution at a distance  $z$  is analytically derived in [7]

from the classical Fresnel diffraction treatment given in the following expression:

$$\begin{aligned}\frac{I(x)}{I_0} &= 1 + A[\sin(\Delta\phi)] + B[1 - \cos(\Delta\phi)], \\ A &= C\left(\frac{x}{\rho}\right) - S\left(\frac{x}{\rho}\right), \\ B &= C^2\left(\frac{x}{\rho}\right) + S^2\left(\frac{x}{\rho}\right) - \frac{1}{2}, \\ \rho &= \sqrt{\lambda z/2},\end{aligned}\quad (1)$$

where  $C$  and  $S$  are the usual Fresnel cosine and sine integral functions, respectively, and  $\rho$  is proportional to the width of the first peak in the diffraction pattern.

Equation (1) is valid for collimated light only. In the case of segmented mirror telescopes, the light beam is a spherical wave converging at the focus, so we cannot define  $z$  as the distance from the pupil to the defocus plane, but rather from the pupil to the defocus conjugate plane as

$$z = \frac{(f - l)f}{l}, \quad (2)$$

where  $f$  is the focal distance, and  $l$  is the distance from the defocus plane (where the detector is placed) to the focal plane. The variable  $\rho$  has to be redefined accordingly to  $\rho_{\text{defoc}} = \rho l/f$ . The normalized intensity distribution  $I(x, y)/I_0$  at a distance  $z$  from the screen is plotted in Fig. 1 for different values of the phase shift. Note that it does not depend on coordinate  $y$ . The signals for values from  $-\pi$  to  $0$  are the same as those from  $0$  to  $\pi$ , but flipped around the point at  $x = 0$ .

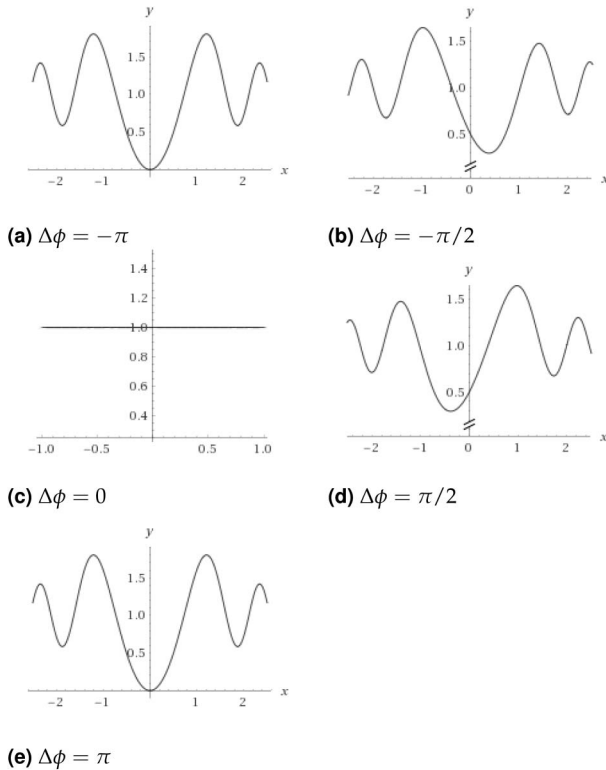


Fig. 1. Intensity distribution perpendicular to the segment edge.

However, the segmentation in the mirror is far more elaborate than the ideal assumption of a piston step with infinite length in Eq. (1). Other mathematical approaches have to be considered. We use the numerical implementation of Fresnel–Kirchoff diffraction formula through convolution to obtain the complex amplitude field at the defocus plane:

$$U(x, y) \propto \mathcal{F}^{-1}\{U(\xi, \eta)e^{-i\pi\lambda z(f_x^2 + f_y^2)}\}, \quad (3)$$

where  $U(\xi, \eta)$  and  $U(x, y)$  are the complex amplitude fields at the pupil and at a distance  $z$  from the pupil respectively. The wavelength of the monochromatic light used is  $\lambda$ , and coordinates in the frequency domain are represented by  $f_x, f_y$ .

Finally, the intensity image at the defocus plane can be found through

$$I(x, y) = |U(x, y)|^2, \quad (4)$$

which is the squared modulus of the complex field.

Here, we describe the procedure followed to generate the simulation used to test the efficiency of the algorithm. We create intensity images applying Eqs. (3) and (4) to the complex amplitude field at the pupil. We only consider piston errors only this Letter. We assume the tip-tilt has been correctly restored previously. A method such as that in Ref. [5] suits our needs, since it requires defocused images as inputs. The wavefront at the pupil is then constant piece-wise. To apply the propagation, we need to set some parameters to the values listed in Table 1 for a telescope with a 36-segment primary mirror.

The propagation distance chosen,  $z$ , is such that the width of the diffraction signal peak at the defocus image plane is at least twice as large as the image blur due to the turbulence for the maximum seeing considered. In addition, the propagation distance is expected to allow the diffraction pattern in the image to be well sampled by the detector.

To surpass the fundamental limit of  $2\pi$  in the measurement range, a sequence of acquisitions with different wavelengths might be taken [8]. In our example we use  $\lambda_0 = 700$  nm as the largest wavelength and three additional shorter wavelengths to disambiguate:  $\lambda_1 = 0.930\lambda_0$ ,  $\lambda_2 = 0.860\lambda_0$ , and  $\lambda_3 = 0.790\lambda_0$ .

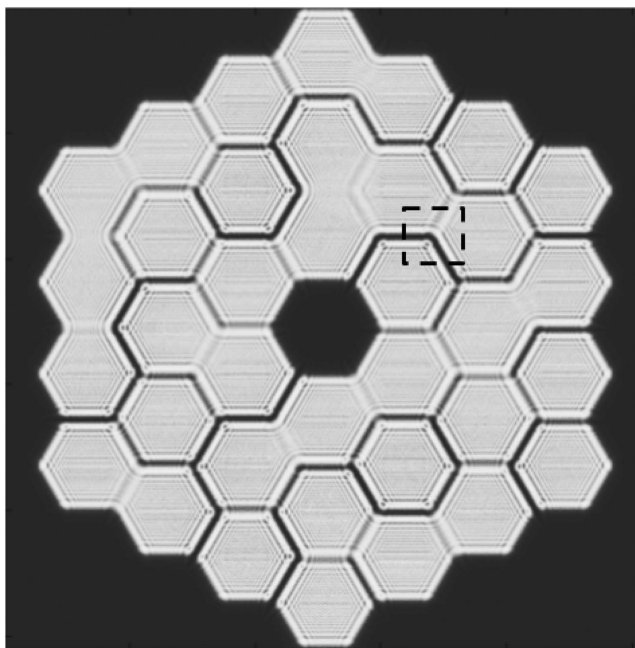
Piston values for each segment were independently extracted from a uniform probability distribution such that the difference between any two adjacent segments is limited to  $\pm 11\lambda_0$ . All piston step values throughout this Letter are measured at the wavefront.

An atmospheric effect is added to simulations by filtering the intensity image at the detector. The filter chosen is the well-known long exposure optical transfer function (OTF) of the atmosphere developed by Fried [9]. The Fried parameter used for this OTF is also randomly picked from a uniform distribution between 0.1 and 0.2 m. Consequently, each image has a different atmospheric seeing.

Figure 2 shows an example of an intensity image obtained with the procedure described above.

Table 1. Simulation Parameters

Parameter	Value
Focal length of telescope	$f = 250$ m
Defocus distance	$l = 4$ m
Diameter of telescope	$D = 10$ m
Detector array size	$1024 \times 1024$

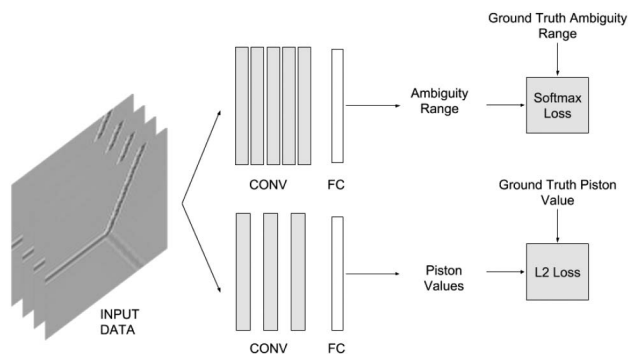


**Fig. 2.** Simulated intensity image. The junction of the three mirror segments used as the input data point is marked with a dashed line.

The network topology that was used is described. The architecture is inspired by that proposed by Alex Krizhevsky [10]. Although much simpler than AlexNet, it remains the same in spirit. We split the network into two branches with different loss functions for each. A schematic representation is shown in Fig. 3.

One branch has three convolutional layers with ReLU activations and a fully connected layer at the end. This is used to predict the piston step values in the range  $[0, \pi]$ . This branch is in charge of performing a regression task, i.e., to predict a continuous value. The loss in this branch is the mean squared error between the output scores of the last fully connected layer and the ground truth labels or, equivalently, the L2 norm of the difference between predicted scores and labels. This prediction can be carried out for the information contained in one single wavelength. Images in the range  $[-\pi, 0]$  are labeled in reverse order to take advantage of the symmetry depicted in Fig. 1.

The second branch has a similar configuration to the first, but contains five convolutional layers and finishes with a



**Fig. 3.** Schematic representation of the network architecture.

softmax loss function. It is used to guess the ambiguity range for every piston step. It performs a classification task. The span of each ambiguity range is equal to  $\pi$ , i.e., piston step values limited to  $\pm 11\lambda_0$  can belong to 44 different ambiguity ranges.

New simulations are generated at every training step of the neural network. First, piston values are selected randomly for every segment, and an intensity image like the one in Fig. 2 is created for each wavelength considered. The effect of the atmosphere is then applied to the images with random seeing. The network should learn how to recover the piston step values, despite the Fried parameter being unknown.

After that, images are cropped into pieces centered on the junctions of every three neighboring segments. An example of these cutouts is shown inside Fig. 2, marked with a dashed line. This input data point is a three-dimensional array. The height and width are those of the cropped window centered at the junction, and the depth is the number of wavelengths taken into account for the analysis. The CNN is trained in a fully supervised manner, so input data and labels must be supplied in the process. We take the image at the junction as data, together with the three piston step values that originate that part of the diffraction image as labels.

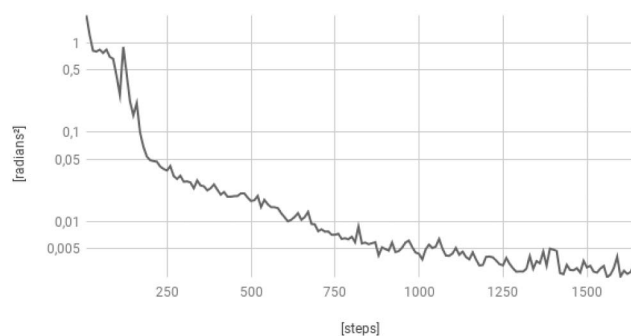
In order to alleviate the computational burden of simulations, we make use of some data augmentation. Given the highly symmetrical properties of the intensity images generated, flipping the images both vertically and horizontally, and reordering labels accordingly, results in a new piece of data useful for training.

The initial learning rate is a hyperparameter that needs to be tuned. We do so by running a random search as suggested in [11].

Network parameters are updated with mini-batch gradient descent algorithm and the Adam update rule [12].

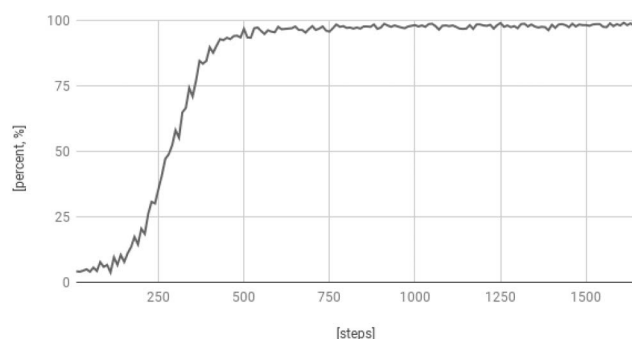
We can see the evolution of the regression loss over the course of the training in Fig. 4. The vertical axis represents the mean squared error between predicted piston steps and ground truth labels, both measured in radians. Values predicted by this part of the network are somewhere within the interval  $[0, \pi]$ , which is why loss starts at 1.6 when guesses are completely random. It is important to point out that the network sees new data at every iteration, so the training error and test error are equivalent. The learning curve plateaus at  $0.003 \text{ rad}^2$ . It reaches the minimum after 1600 training steps which means that 409,600 examples have passed through the network.

**Regression loss over training**



**Fig. 4.** Regression loss over training in a logarithmic scale.

Classification accuracy over training

**Fig. 5.** Percentage of correctly classified ambiguity ranges.**(a)**  $RMSE = 20.63$ **(b)**  $RMSE = 0.056$ **Fig. 6.** Sample outcome of a wavefront with misaligned segments (left) versus a wavefront after segments have been phased (right).

In Fig. 5, the accuracy of the classification task over training is plotted. The highest accuracy attained is around 98%. This means that 2 in every 100 ambiguity ranges are misclassified. Therefore, one must be extra careful when handling them. Fortunately, we are interested in piston values for each segment, rather than piston steps between adjacent segments. The sum of the three piston step values of each junction should add up to zero. Otherwise, an ambiguity range has been misclassified. It can be located through a relationship with adjacent junctions and then discarded. Still, the number of relations, i.e., piston step values, exceeds the number of unknowns, i.e., global piston values. This set of overdetermined linear equations can be solved with singular value decomposition.

Figure 6 shows a sample wavefront before and after the segments have been properly phased. In this case, classification errors were detected and were not considered. Piston steps are between  $\pm 69.115$  rad initially, and the RMSE is 20.63. Once all of the piston global values are computed with the method just described, they are moved back using any segment position as reference. This leads to the wavefront illustrated in Fig. 6 right with  $RMSE = 0.056$ .

It has been proven in this Letter that the method trained with simulated data is efficient. The method is fast, since only a few exposures are needed, and there is no iterative process involved after the network is trained. An imaging camera for visible light suffices for the method to operate, so it requires little or no additional equipment to be installed. It is possible to reach broad capture range through the usage of combined wavelengths. In this example, we employ four wavelengths to obtain a capture range of  $\pm 11\lambda_0$ , in units of the largest wavelength. This is satisfactory to follow up the drift of edge sensor reference periodically; however, it is not suitable for segment integration [13]. An accuracy of around  $\pm 0.0087\lambda_0$  was achieved, which means an error of  $\pm 6.1$  nm for the piston step measured at the wavefront for the given wavelength. Classification errors must be checked for consistency first. The presence of many of them could be problematic, and it is something we have to address in future developments. In addition, the method has been tested to be robust to any atmospheric turbulences with Fried parameters between 0.1 and 0.2 m. Tip-tilt and other sources of errors are planned to be added in following work.

**Funding.** Ministerio de Economía, Industria y Competitividad, Gobierno de España; European Regional Development Fund (ERDF); Spanish Ministry of Economy (DPI2015-66458-C2-2-R).

## REFERENCES

1. G. Chanan, C. Ohara, and M. Troy, *Appl. Opt.* **39**, 4706 (2000).
2. J. M. Rodríguez-Ramos and J. J. Fuensalida, *Optical Telescopes of Today and Tomorrow* (International Society for Optics and Photonics, 1997), Vol. **2871**, pp. 613–617.
3. G. Chanan, M. Troy, and E. Sirko, *Appl. Opt.* **38**, 704 (1999).
4. A. Schumacher and N. Devaney, *Mon. Not. R. Astron. Soc.* **366**, 537 (2006).
5. J. J. Fernández-Valdivia, A. L. Sedano, S. Chueca, J. S. Gil, and J. M. Rodríguez-Ramos, *Opt. Eng.* **52**, 056601 (2013).
6. P. C. McGuire, D. G. Sandler, M. Lloyd-Hart, and T. A. Rhoadarmer, in *Scientific Applications of Neural Nets* (Springer, 1999), pp. 97–138.
7. J. M. Rodríguez-Gonzalez and J. J. Fuensalida, in *Large Ground-based Telescopes* (International Society for Optics and Photonics, 2003), Vol. **4837**, pp. 726–737.
8. M. G. Lofdahl and H. Eriksson, *Opt. Eng.* **40**, 984 (2001).
9. D. L. Fried, *J. Opt. Soc. Am.* **56**, 1372 (1966).
10. A. Krizhevsky, I. Sutskever, and G. E. Hinton, *Advances in Neural Information Processing Systems* (2012), pp. 1097–1105.
11. J. Bergstra and Y. Bengio, *J. Mach. Learn. Res.* **13**, 281 (2012).
12. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv:1412.6980* (2014).
13. N. Yaitskova, F. Gonte, F. Deriche, L. Noethe, I. Surdej, R. Karban, K. Dohlen, M. Langlois, S. Esposito, E. Pinna, and M. Reyes, *Ground-Based and Airborne Telescopes* (International Society for Optics and Photonics, 2006), Vol. **6267**, p. 62672Z.