Wavefront sensing with optical differentiation powered by deep learning

Biswa R. Swain,[1,2] M. Akif Qadeer,[1,2] Christophe Dorrer,[3] Renuka Manjula Narayanan,[2,4] Jannick P. Rolland,[2,4] and Jie Qiao[1,2,3,*]

[1]*Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, New York 14623, USA*
[2]*The Center for Freeform Optics, 275 Hutchison Rd., Rochester, New York 14627, USA*
[3]*Aktiwave LLC, 150 Lucius Gordon Dr., West Henrietta, New York 14586, USA*
[4]*The Institute of Optics, University of Rochester, 275 Hutchison Rd., Rochester, New York 14627, USA*
*qiao@cis.rit.edu

We report the experimental demonstration of an optical differentiation wavefront sensor (ODWS) based on binary pixelated linear and nonlinear amplitude filtering in the far-field. We trained and tested a convolutional neural network that reconstructs the spatial phase map from nonlinear-filter-based ODWS data for which an analytic reconstruction algorithm is not available. It shows accurate zonal retrieval over different magnitudes of wavefronts and on randomly shaped wavefronts. This work paves the way for the implementation of simultaneously sensitive, high dynamic range, and high-resolution wavefront sensing.

Phase measurement of an optical wave is essential in the metrology of optical components, adaptive optics, and laser beam quality assessment. The Shack–Hartmann wavefront sensor (SHWS) obtains wavefront slopes by measuring centroid displacements at the focal plane of an array of microlenses. It is limited in spatial resolution and dynamic range, which are dependent on the pitch and focal length of the microlens array [1,2]. Interferometry with an adaptive nulling component such as a deformable mirror or a spatial light modulator (SLM) can offer reconfigurable surface measurements. In both cases, a high-performance wavefront sensor is required to monitor the adaptive component [3,4].

One promising phase measurement method is the optical differentiation wavefront sensor (ODWS), which is based on the amplitude modulation in the far-field of the wave under test [5,6]. The resulting near-field fluence, with and without the filter, leads to the slopes of the input wavefront in the direction of the filter transmission gradient. The full input phase map can be reconstructed from two wavefront slope datasets obtained with two orthogonal orientations of the filter. Compared to the SHWS, the ODWS has the advantages of higher spatial resolution that is only limited by the camera pixel pitch, higher signal-to-noise ratio, and a scalable dynamic range with an increase in the filter size [7].

The accurate realization of the filter's spatially varying amplitude transmission is important for ODWS operation. Amplitude modulations based on a liquid-crystal SLM and a static filter functioning as a spatially varying polarization rotator were previously proposed, but they are polarization- and wavelength-dependent [8,9].

In this work, the amplitude transmission profile has been realized by a pixelated filter composed of a distribution of transparent and opaque pixels [10] implemented using a spatial dithering algorithm [11]. Such a filter can be fabricated by the lithography of a metal layer deposited on a glass plate. It is static, polarization-independent, and achromatic [12]. We have previously demonstrated ODWS-based wavefront sensing via implementing a binary pixelated filter with a linear amplitude transmission profile [10,13,14]. Including this work, we scaled the ODWS dynamic range tenfold by increasing the width of the linear filters from 2 to 20 mm [10,13,14]. It is feasible to further scale the dynamic range by another 10 times or more by deploying filters with a width of 250 mm, a commercially available 10-in. wafer. The potential applications of the ODWS include surface quality assessment for freeform or complex shaped optics during fabrication and wavefront sensing for laser beams, ground telescopes, and vision.

The ODWS has an intrinsic trade-off between the dynamic range and sensitivity. The dynamic range of the wavefront slopes measured by an ODWS is proportional to the filter width and inversely proportional to the focal length of the imaging system. However, the sensitivity is inversely proportional to the width of the linear filter [6,15]. We note that the spatial frequency of the near field beam at the far-field filter location increases from the center to the edge of the filters. A filter with a nonlinear transmission profile that generates steep and shallow transmission slopes for low and high spatial frequencies can overcome the trade-off, providing higher sensitivity and higher accuracy for wavefronts dominated by lower spatial frequency components. However, there is no closed-form analytical solution to relate the fluence images to the wavefront for a nonlinear far-field modulation. Recently, deep learning has attracted interest in image-based wavefront sensing [9,16–20]. This data-driven approach establishes complex nonlinear relations between system inputs and outputs, without details of a physical model
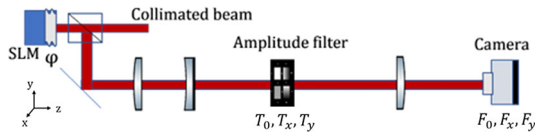
Fig. 1.   Schematic of the experimental ODWS setup.

Fig. 2.   (a) Amplitude transmission and (b) slope profiles for the linear filter (LF), nonlinear filter 1 (NLF1), and nonlinear filter 2 (NLF2).

of the system. Convolutional neural network (CNN) architectures using modal coefficients to restore wavefronts have been designed to predict deformable mirror modes for ODWS that uses liquid-crystal-based nonlinear filters [8,9]. However, such filter implementations are polarization- and wavelength-dependent. Additionally, it is challenging to efficiently restore randomly shaped wavefronts using a modal-coefficient-based CNN [17,18].

In this Letter, we report an ODWS using a binary pixelated filter with nonlinear amplitude transmission and a zonal-reconstruction-based CNN architecture for achromatic, polarization-insensitive, generalized, and robust phase reconstruction. Generalized phase reconstruction refers to retrieving complex and random wavefronts that cannot be efficiently restored by modal-coefficient-based solutions because a large number of modes are required.

We aim to use pixelated nonlinear filters to overcome the dynamic range and sensitivity trade-off, providing higher sensitivity for low wavefront slope within a fixed dynamic range. We also aim to improve the accuracy and robustness for wavefront restoration by devising a zonal-reconstruction-based CNN architecture. While such a CNN architecture has been proposed to obtain spatial phase information for SHWS and deflectometry [17–19], it has not been demonstrated for robust ODWS-based wavefront restoration.

Figure 1 shows the experimental ODWS layout. A high-definition analog SLM (Meadowlark Optics XY Phase Series) is positioned at the object plane to generate wavefronts [4]. The first two lenses form a telephoto system which is confocal with the third lens. The camera is at the image plane of the overall optical system, conjugate to the SLM plane, where the wavefront φ is to be measured. The amplitude transmission filter is placed at the focal plane of the telephoto, on the image side. The setup is built per the design parameters described in [14]. The spatial resolution of the measured wavefront is 50.7 µm. The square aperture that inscribes the circular wavefront has 192×192 sampling points.

For a linear amplitude transmission gradient filter, the relation between the fluence profile at the image plane and the wavefront slope in the x direction (horizontal) is described by Eq. (1) as follows:

$$\frac{\partial \varphi \left(-\frac{x}{m}, -\frac{y}{m}\right)}{\partial \left(-\frac{x}{m}\right)} = \frac{\pi W}{\lambda f} \left(2\sqrt{\frac{F_x(\mathrm{x}, \mathrm{y})}{F_0(\mathrm{x}, \mathrm{y})}} - 1\right), \tag{1}$$

where F_0 is the fluence with 100% uniform filter transmission ($T_0 = 1$), W is the filter width, F_x is the fluence map obtained with the filter transmission gradient along the x direction, f is the effective focal length (EFL) of the telephoto system, and m is the magnification. Similarly, wavefront slopes along the y direction (vertical) can be obtained from the fluence map F_y measured at the image plane with a 90° filter rotation. The wavefront map is reconstructed from the obtained orthogonal slopes using the Southwell's zonal reconstruction procedure [21].
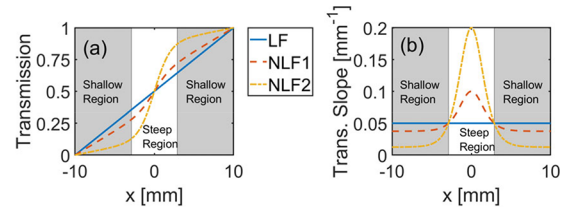
Wavefronts with steep slopes can be measured with an ODWS by employing an amplitude filter of large width. From Eq. (1), the ratio in fluence values, with and without the filter, decreases with increasing filter width for a fixed wavefront slope. Thus, when detection noise is present, the measurement of small phase variations is corrupted by a large filter width, leading to the decrease in sensitivity. A nonlinear transmission profile with a relatively high amplitude transmission gradient near the center region of a filter having an overall large width is thus both sensitive and high dynamic range.

The transmission profiles of the binary pixelated nonlinear filters investigated in this work are described as follows [8]:

$$T_x(x, y) = \left(\frac{1 - \beta}{2}\right)\left(1 + \frac{x}{W/2}\right) + \frac{\beta}{1 + e^{-\frac{2x}{\sigma W}}}. \tag{2}$$

The transmission profile in Eq. (2) is the combination of a linear filter and a step-like function where β and σ determine the height and width of the step, respectively.

In this work, we designed and implemented three filters with the same width of 20 mm but having significantly different slopes in the center and the outer portions of the filters, corresponding to the steep and shallow slope regions, respectively (Fig. 2). Figures 2(a) and 2(b) respectively show the transmission and slope profiles of these three filters: one linear filter (LF) and two nonlinear filters (NLF1 and NLF2). The corresponding transmission slopes at the center regions are 0.05, 0.10, and $0.20\,\mathrm{mm^{-1}}$, respectively. In comparison to the LF, the transmission slopes of NLF1 and NLF2 are increasingly steeper at the center region while becoming progressively shallower in the outer regions, respectively. Such a design consideration allows us to compare the impact of the spatially variant transmission on sensitivity measurement.

Figure 3(a) shows the CNN architecture implemented to retrieve the wavefront from the ODWS fluence images generated with these filters. The CNN is inspired by the U-Net architecture [17–19,22] and implemented in the PyTorch framework [23].

The input and output of the CNN are the fluence image ratios (F_x/F_0, F_y/F_0) and the corresponding wavefronts, respectively. The two-channel input of fluence image ratios is first processed by a series of encoder layers, shown as the blue blocks. As illustrated by the architecture of encoder layers in Fig. 3(b), the fluence ratios are sequentially processed to produce feature maps through batch normalization, convolution operations, average pooling, and activation with a leaky rectified linear unit (ReLU). With each consecutive encoder layer, the spatial size of the feature maps is halved, and the number of feature maps is doubled. After the encoder layers, the feature maps are processed through the decoder layers, shown in green blocks. As per the architecture of decoder layers in Fig. 3(c), the feature maps are
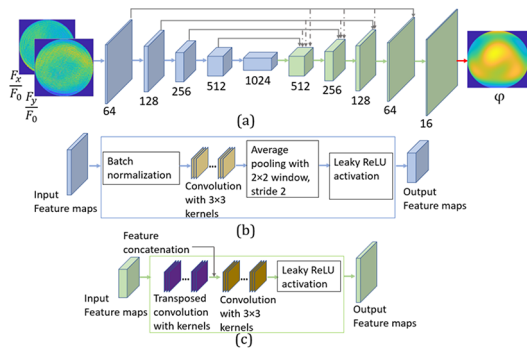
Fig. 3. (a) Overall CNN architecture with encoder (blue blocks) and decoder (green blocks) layers for wavefront reconstruction from ODWS fluence ratio maps; the number under each block is the depth of the feature maps or the number of convolution kernels in that layer. (b) Architecture of an encoder layer. (c) Architecture of a decoder layer.
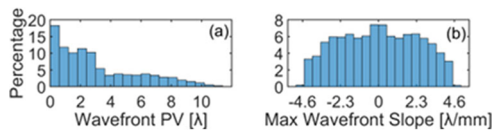

Fig. 4. Histograms of (a) the peak-to-valley (PV) and (b) the maximum slopes of the 10,000-wavefront dataset.

sequentially processed through up-sampling-transposed convolution, followed by convolution with kernels and activation with leaky ReLU. In this layer, the up-sampled feature maps are concatenated with features from the encoder layers which facilitate convergence of the CNN. In each consecutive decoder layer, the number of feature maps is halved, and the spatial dimension is doubled. The output of the last layer is resized to a depth of one by applying 1×1 convolution (red arrow). The output represents the reconstructed wavefront map, where the spatial array dimensions are equal to those of the input fluence ratio maps.

The network is trained by minimizing a loss function. The loss function is defined as the root mean square of the residual between the CNN-predicted wavefront and the ground-truth wavefront. A total of 500 training cycles are used with each cycle trained in steps with mini batches of 64 wavefronts.

The strategy for generating the input training dataset is to provide sufficient wavefront variability while ensuring that the wavefront slopes are within the dynamic range of the ODWS. Figure 4 shows the statistics of the 10,000-wavefront set generated for training and testing purposes. The peak-to-valley spread of the input wavefronts is between 0 and 10 λ ($\lambda = 632.8$ nm), and the maximum slope for each wavefront is between -4.6 and 4.6λ/mm (in each orthogonal direction), primarily limited by the proof-of-concept filter width of 20 mm. Considering the first 37 terms of the Fringe Zernike polynomials, the 10,000-wavefront data consist of three sets: Set 1 is created with randomized individual Fringe Zernike polynomials, Set 2 consists of combined Fringe Zernike polynomials with coefficients chosen from unique normal distributions, and Set 3 is the combination of different Fringe Zernike wavefronts and random patterns. The random patterns are generated by assigning a random value within a set range at each pixel of the wavefront matrix, which is then convolved with a Gaussian filter to ensure that the maximum slope is within the dynamic range [24]. We first trained and
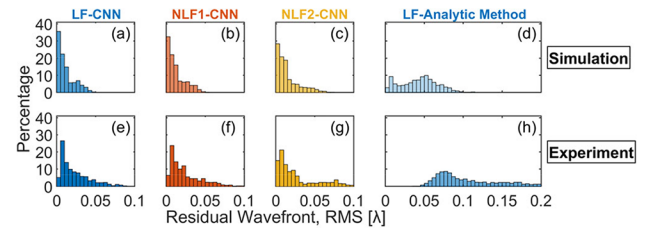

Fig. 5. Histograms of the residual RMS of the 1000 test wavefronts retrieved by the CNN using the fluence data obtained respectively from simulation (top row) and experiment (bottom row). (a) and (e) Linear filter LF; (b) and (f) nonlinear filter NLF1; (c) and (g) nonlinear filter NLF2; and (d) and (h) retrievals using the analytic equation.

tested the CNN for each of the three ODWS filter configurations (Fig. 2) through simulation, then repeated the entire process with experiments. The same 10,000 input ground-truth wavefronts were paired with their corresponding ODWS fluence ratio maps to train and test each filter configuration. Nine thousand ground-truth wavefronts were used for training, and 1000 were used for testing. This separate training process for the three filters, both in simulation and experiments, resulted in a total of six sets of CNN testing results for the same 1000 testing wavefronts. For simulation, the fluence maps were generated theoretically using an ODWS model based on the Rayleigh–Sommerfeld diffraction. For experiments, the input wavefronts were created by the SLM in the setup shown in Fig. 1, and the corresponding fluences were experimentally captured at the image plane of the ODWS. The SLM was custom-calibrated to precisely account for the wavelength used in this experiment, and its 4π phase modulation range [4].

It takes approximately 3 h for each CNN training, using an NVIDIA A100 GPU. A test wavefront is retrieved from the fluence in 110 ms using the trained CNN. This speed can be significantly improved by further optimizing the CNN architecture and reducing its number of parameters (layers, channels, etc.), which is beyond the scope of this work.

Figure 5 shows the histograms of the residual wavefronts (RMS) from simulations [Figs. 5(a)–5(d)] and experiments [Figs. 5(e)–5(h)]. For the linear-filter configuration, the test wavefronts were also reconstructed using the analytic equation, using the simulated and experimentally measured fluence ratios [Figs. 5(d) and 5(h)]. The residual wavefront is defined as the difference between the ground truth and the reconstructed wavefront.

The CNN-based wavefront retrievals for all three filters statistically have lower residual wavefront RMS than the corresponding analytic reconstruction results for the linear filter, validated by simulation and experimental results. The statistical mean values of the residual RMS for the CNN retrieval are 0.015λ for the simulation and 0.027λ for experiments, significantly less than the corresponding statistics from the analytic method, 0.040λ and 0.100λ, respectively.

This result demonstrates that the CNN is statistically more accurate in comparison to the analytic equation. The inadequacy of the analytic equation compared to the simulation-based CNN prediction may be attributed to the slowly varying amplitude assumption made in its derivation and the robustness of the wavefront slope integration [25,26]. Comparing the respective experimental and simulation results for each filter by the CNN and the analytic equation, the experimental accuracy deteriorates
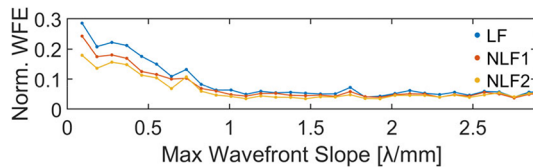
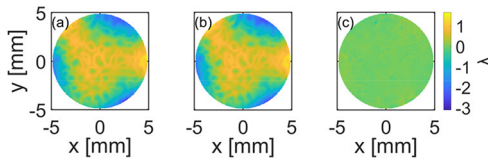**Fig. 6.** Normalized wavefront error (WFE) for the LF, NLF1, and NLF2.

Fig. 7. CNN performance on a random pattern wavefront: (a) ground truth (PV: $4.777\,\lambda$, RMS: $0.780\,\lambda$); (b) CNN prediction (PV: $4.513\,\lambda$, RMS: $0.799\,\lambda$); (c) residual (RMS: $0.056\,\lambda$).

for both the CNN and the analytic methods. However, the deterioration is much higher for the analytic method. The additional deterioration is due to system imperfections, such as detector noise, residual misalignments, and inaccuracy in amplitude transmission realized with the spatially dithered binary pixelated filters. The analytic equation does not account for these imperfections, while the CNN mitigates these undesired effects during the learning process. These results statistically prove that the zonal-reconstruction-based CNN is more robust to these system imperfections as compared to the analytic equation.

We further characterized the effectiveness of using nonlinear filters to achieve higher sensitivity for a given dynamic range limited by the prototype filter width of 20 mm. Figure 6 shows the residual RMS wavefront errors (WFEs), normalized with respect to the input wavefront RMS of the 1000 experimentally measured wavefronts as a function of the maximum wavefront slope. We observe a progressive increase in measurement accuracy for the wavefronts with low wavefront slopes, as filters with a larger center slope are used, specifically 0.05, 0.10, and 0.20 mm^{-1} for the LF, NLF1, and NLF2, respectively. This demonstrates the improved measurement sensitivity without sacrificing the dynamic range by using the nonlinear filter profiles. See Supplement 1 for the fluence and restored wavefront maps for an example of ground-truth wavefront where the effect of the sensitivity gain is clearly observed.

We finally demonstrate the effectiveness of a zonal-reconstruction-based CNN architecture in predicting random phase profiles, which cannot be efficiently restored by a modal-coefficient-based CNN architecture, such as using deformable mirror modes or Zernike coefficients. Figure 7(a) shows a hybrid wavefront that is a combination of Zernike aberrations and a random-pattern phase profile. The corresponding CNN prediction (using NLF2) from the experimentally measured ODWS fluence maps and its difference from the ground truth (residual RMS: $0.056\,\lambda$) is plotted in Figs. 7(b) and 7(c), respectively. This result indicates that the CNN can retrieve complex wavefront shapes that can be present in astronomical imaging and biological samples.

In conclusion, we have demonstrated a binary pixelated nonlinear-filter-based ODWS using a CNN as a robust wavefront reconstructor. We have improved sensitivity without sacrificing dynamic range, overcoming the trade-off between dynamic range and sensitivity. Through simulations and experiments, we show that the zonal-reconstruction-based CNN architecture provides higher accuracy and robustness than the analytic reconstruction, with respective statistical mean residual RMS of $0.027\,\lambda$ and $0.100\,\lambda$ for experiments. Additionally, we have experimentally demonstrated that this CNN architecture is capable of retrieving complex wavefront shapes that modal coefficients cannot efficiently describe.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. R. V. Shack and B. C. Platt, J. Opt. Soc. Am. **61**, 656 (1971).
2. M. Rocktäschel and H. J. Tiziani, Opt. Laser Technol. **34**, 631 (2002).
3. K. Fuerschbach, K. P. Thompson, and J. P. Rolland, Opt. Lett. **39**, 18 (2014).
4. R. Chaudhuri, A. Wansha, R. Porras-Aguilar, *et al.*, Opt. Express **30**, 43938 (2022).
5. J. C. Bortz, J. Opt. Soc. Am. A **1**, 35 (1984).
6. H. Furuhashi, K. Matsuda, and C. P. Grover, Appl. Opt. **42**, 218 (2003).
7. J. Oti, V. Canales, and M. Cagigal, Opt. Express **11**, 2783 (2003).
8. S. Y. Haffert, Opt. Express **24**, 18986 (2016).
9. R. Landman and S. Y. Haffert, Opt. Express **28**, 16644 (2020).
10. J. Qiao, Z. Mulhollan, and C. Dorrer, Opt. Express **24**, 9266 (2016).
11. C. Dorrer and J. D. Zuegel, J. Opt. Soc. Am. B **24**, 1268 (2007).
12. C. Dorrer, Appl. Opt. **52**, 3368 (2013).
13. B. R. Swain, C. Dorrer, and J. Qiao, Opt. Express **27**, 36297 (2019).
14. B. R. Swain, C. Dorrer, and J. Qiao, Opt. Express **29**, 38395 (2021).
15. A. Sagan, T. J. Antosiewicz, and T. Szoplik, Appl. Opt. **48**, 1143 (2009).
16. S. W. Paine and J. R. Fienup, Opt. Lett. **43**, 1235 (2018).
17. R. Swanson, K. Kutulakos, S. Sivanandam, *et al.*, Proc. SPIE **10703**, 52 (2018).
18. L. Hu, S. Hu, W. Gong, *et al.*, Opt. Lett. **45**, 3741 (2020).
19. J. Dou, D. Wang, Q. Yu, *et al.*, Opt. Lett. **47**, 78 (2022).
20. F. Guzmán, J. Tapia, C. Weinberger, *et al.*, Photonics Res. **12**, 301 (2024).
21. W. H. Southwell, J. Opt. Soc. Am. **70**, 998 (1980).
22. O. Ronneberger, P. Fischer, and T. Brox, in *Lecture Notes in Computer Science*, Vol. 9351 (Springer, 2015), pp. 234–241.
23. A. Paszke, S. Gross, F. Massa, *et al.*, Adv. Neural Inf. Process. Syst. **32**, 8026 (2019).
24. S. Schott, J. Bertolotti, J.-F. Léger, *et al.*, Opt. Express **23**, 13505 (2015).
25. J. Qiao, A. Travinsky, G. Ding, *et al.*, Proc. SPIE **9356**, 935608 (2015).
26. B. Pathak and B. R. Boruah, J. Opt. **16**, 055403 (2014).