# Extended scene deep learning wavefront sensing

de Bruijne, Bas; Vdovin, Gleb; Soloviev, Oleg

**DOI**
10.1364/JOSAA.443436

**Publication date**
2022

**Document Version**
Final published version

**Published in**
Journal of the Optical Society of America A: Optics and Image Science, and Vision

Research Article Vol. 39, No. 4 / April 2022 / Journal of the Optical Society of America A 621

Journal of the Optical Society of America A

OPTICS, IMAGE SCIENCE, AND VISION

# Extended scene deep learning wavefront sensing

Bas de Bruijne,[1,*] Gleb Vdovin,[1,2] and Oleg Soloviev[1,2]

[1]Delft Center for Systems and Control, TU Delft, Mekelweg 2, 2628 CD Delft, The Netherlands
[2]Flexible Optical B.V., Polakweg 10-11, 2288 GG Rijswijk, The Netherlands
*Corresponding author: BasdBruijne@gmail.com

Received 5 October 2021; revised 19 February 2022; accepted 21 February 2022; posted 24 February 2022; published 17 March 2022

**We have applied a combination of blind deconvolution and deep learning to the processing of Shack–Hartmann images. By using the intensity information contained in spot positions, and the fine structure of the separate images created by the lenslets, we have increased the sensitivity and resolution of the sensor over the limit defined by standard processing of spot displacements only. We also have demonstrated the applicability of the method to wavefront sensing using extended objects as a reference.** 

https://doi.org/10.1364/JOSAA.443436

## 1. INTRODUCTION

The Shack-Hartmann (SH) sensor is commonly applied in a wide range of imaging applications. The SH sensor consists of two parts: a microlens array (MLA) and imaging sensor, usually based on CCD or CMOS technology. The MLA, usually positioned in the system pupil, divides the light into $M$ images, each formed in a separate subaperture (SA). Conventionally, only the tip–tilt modes are reconstructed for each SA by measuring the $x$ and $y$ shifts of the image from its reference position. For this calculation, the found wavefront slopes can be either fit to a number of Zernike polynomials (modal wavefront reconstruction) or integrated into an array of wavefront values connected to the SAs (zonal wavefront reconstruction) [1].

Using this method, a SH sensor with $M$ SAs can be used to retrieve a wavefront with maximum $2M$ modes, and the wavefront reconstruction is limited to the value of $r_0 > \frac{D}{\sqrt{M}}$.

Different techniques have been proposed to attempt to elevate these limitations, e.g., by using the second moment information of the SAs [2]. The most promising of these techniques in terms of wavefront reconstruction accuracy is the use of machine learning. Earlier research showed that artificial neural networks (ANNs) can be used to accurately reconstruct the wavefront from SH slope measurements, by using either Zernike polynomials [3] or a convolutional neural network (CNN) to reconstruct the wavefront directly [4]. Suárez Gómez*et al.* [5] showed that CNNs can be used to reconstruct wavefront slopes from SH images. Recently, it was shown that CNNs could be used not only for wavefront slope sensing and reconstruction separately, but that one CNN could be used to fulfill both functions [6–8].

Deep learning techniques are superior to conventional methods for wavefront sensing, as they do not solely rely on the shifts of SA images but can also analyze their shapes. The shapes of SA images provide information about the higher order wavefront aberrations that are neglected by conventional wavefront sensing methods.

The above mentioned papers share a major simplification that prevents their methods from being applied in certain imaging applications: they take only point-source observations into account.

While Sánchez-Lasheras *et al.* [9] showed that it is possible to train an ANN to reconstruct the wavefront from an extended scene SH image directly, their methods were applicable only to solar spots, a very specific type of scene. To the authors' best knowledge, no deep learning wavefront sensing (DLWS) technique has been proposed that can be applied to extended scene observations in general.

In this paper, a method is proposed that is aimed to extend DLWS methods to extended scene imaging applications. This method makes use of a semi-blind image deconvolution algorithm that pre-processes the SH image by removing the influence of the extended scene, and returns an estimated point-source SH image. This image is then used in a CNN that is designed to reconstruct the wavefront given uncertainties in the SH image introduced by the pre-processing step. These methods are tested using numerical simulations of a deconvolution from wavefront sensing (DFWS) system as shown in Fig. 6.

First, Section 2 will discuss the mentioned pre-processing step. Second, Section 3 will introduce the modified DLWS methods. Section 4 will discuss how the newly proposed methods compare to traditional extended scene wavefront sensing systems as well as earlier proposed DLWS architectures.

## 2. BLIND SHACK–HARTMANN IMAGE DECONVOLUTION

This section discusses the pre-processing of SH images to remove the dependency of the object on the SH image.

First, consider the mathematical model dictating how the SH image is constructed:

$$i_{\mathrm{sh}} = o * k_{\mathrm{sh}} + n. \tag{1}$$

Here, $*$ refers to the convolution operator, $o$ is the undistorted object, $k_{\mathrm{sh}}$ is the SH point spread function (PSF), $i_{\mathrm{sh}}$ is the SH image, and $n$ represents a noise component. In the case where a point source is observed, $o$ is a delta function, and $i_{\mathrm{sh}} = k_{\mathrm{sh}} + n$.

$k_{\mathrm{sh}}$ is related to the wavefront $\phi$ by

$$k_{\mathrm{sh}} \propto \left|\mathcal{F}\left\{P e^{i(\phi+\phi_{\mathrm{sh}})}\right\}\right|^2. \tag{2}$$

Here, $P$ is the pupil function, and $\phi_{\mathrm{sh}}$ is a piece-wise linear phase component that is added to simulate the influence of the MLA, as described by Soloviev *et al.* [10].

In this paper, blind deconvolution refers to the retrieval of $k_{\mathrm{sh}}$ without the knowledge of $o$. However, more information is assumed to be available than usual in a blind image deconvolution setting.

In many adaptive optics (AO) applications, a second imaging sensor is used to capture high-resolution images with a higher diffraction limit. The construction of this image ($i$) is equivalent to the construction of $i_{\mathrm{sh}}$ as described above, but with a different pupil function and $\phi_{\mathrm{sh}} = \vec{0}$. The down-sampling of $i$ is used as an initial estimate of the object in the blind deconvolution problem. Compared to using one individual SA image or a shifted averaging of all SA images, this initial estimate is distorted in an unbiased manner compared to the individual SA images while requiring minimal processing.

The blind image deconvolution algorithm used in this paper is an adaptation of the tangential iterative projections (TIP) algorithm as introduced by Wilding *et al.* [11]. TIP is representative of a broad class of blind deconvolution techniques that work by alternating between estimating the PSF and the object from the image and the latest estimates of the PSF and object (see [12–14] for the most related work and [15] for a general overview), and differs from other algorithms by using minimal *a priori* assumptions, namely, only the support size of the PSFs and the nonnegativity of the PSFs and the object. In TIP, one iteration looks as follows. Given the latest estimate of the SH PSF ($\hat{K}_{\mathrm{sh}}$) and the SH image ($I_{\mathrm{sh}}$), an estimate of the object ($\hat{O}$) can be calculated as

$$\hat{O} = \mathcal{P}_O \operatorname*{arg\,min}_{O \in \mathbb{C}^{M\times M}} \left\| I_{\mathrm{sh}} - O\hat{K}_{\mathrm{sh}} \right\|^2, \tag{3}$$

where the capitalization of a variable refers to its Fourier transform. Given the updated estimation of the object, the estimate of the SH PSF follows from

$$\hat{K}_{\mathrm{sh}} = \mathcal{P}_{\mathcal{K}} \frac{I_{\mathrm{sh}}}{\hat{O}}. \tag{4}$$

Here, the projection operators $\mathcal{P}_{\mathcal{X}}$ project the result of the least-squares deconvolution onto a set of valid objects or PSFs $X$. In general, this valid set consists of positive and normalized images. Additional constraints can be added to this set to guide the convergence of the algorithm.

For this application, the TIP algorithm is altered. Rather than using the projection operator to enforce a constraint on the results from the deconvolution, the projection operators are used to over-constrain the images to direct the TIP algorithm towards the right solution. These projection operators are relaxed as the iteration continues. For construction of the projection operator, some prior information of the SH image is used. Consider the two possibilities describing the relation between the latest estimate of the object and the SA images (ignoring the odd case where the SA image is equal to the estimated object).

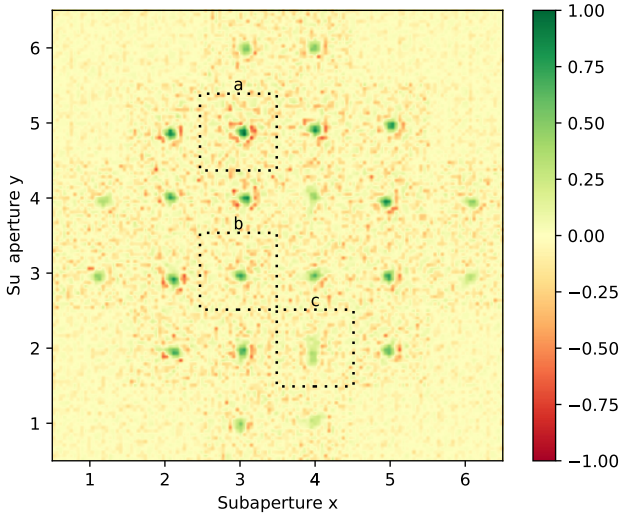


**Fig. 1.** Estimated SH pattern in the first TIP iteration, before the projection step. Three types of PSFs can be identified in this image: (1) the SA image is more blurred than the estimated object, resulting in a blurred PSF, e.g., SA c; (2) the SA image is similar to the estimated object, resulting in a sharp PSF, e.g., SA b; (3) the SA image is sharper than the estimated object, resulting in a PSF with a large positive values as well as negative values, e.g., SA a. Filtering out the SA images corresponding to (1) and discarding the negative values will direct the TIP algorithm to an accurately estimated object.

(1) The SA image is more blurred than the estimated object. This will result in a large PSF (e.g., SA c in Fig. 1).
(2) The SA image is sharper than the estimated object. In this latter case, the PSF does not represent a blurring function, but instead a sharpening function. For the PSF to perform a sharpening action, it must contain negative values as well as large positive values. This effect is highlighted in Fig. 1. SA a is sharper than the estimated object and contains negative values.

To improve the estimate of the object, the algorithm must focus on SA PSFs in the second category. While convolution with these PSFs results in sharper images, deconvolution (which is applied in the next step) with these PSFs will again result in a more blurred estimate of the object. To preserve the sharpness of the images in the deconvolution step, the negative component of the PSFs must be removed. By increasing the contrast of the image, the influence of SA images in the first category is reduced, as these PSFs have lower intensities. This way the TIP algorithm can be directed to a sharper estimate of the object.

The pseudocode shown in Fig. 2 shows a snippet of the exact implementation of the TIP algorithm.

Lines 2–4 perform a least-squares deconvolution and extract the positive real values from the estimated PSF. Note that the

```
1)  for n in range(n_max):
2)    K_sh = I_sh/(Ob+1*(abs(Ob)<1))
3)    k_sh = real(ifft2(K_sh))
4)    k_sh *= (k_sh > 0)
5)
6)    k_sh -= min(k_sh)
7)    k_sh /= max(k_sh)
8)    k_sh **= f(n_max, n)
9)
10)   k_sh -= min(k_sh)
11)   k_sh /= sum(k_sh)
12)   K_sh = fft2(k_sh)
13)
14)   conj = conj(K_sh)
15)   Ob = (conj*I_sh)/(conj*K_sh+1e-9)
16)   ob = abs(ifft2(Ob))
17)   ob *= fov
18)   ob -= min(ob)
19)   ob /= max(ob)
20)   Ob = fft2(ob)
```

**Fig. 2.** Snippet of pseudocode showing the exact implementation of the modified TIP algorithm. Note that for the clarity of the reader, the code shown is not optimized for speed. The object $o$ is referred to as ob.

regularization parameter is added only to the frequencies with a lower absolute value than the regularization parameter in order not to distort valid frequencies. Line 4 removes the negative values in the PSF. Lines 6–7 normalize the estimated PSF between zero and one. Line 8 increases the contrast of the estimated PSF by raising it to the power of $f(n_{\max}, n)$. $f(n_{\max}, n)$ must be selected according to the number of iterations ($n_{\max}$), noise levels, and scenery and is constrained to $f(n_{\max}, n_{\max} - 1) = 1$ and $f > 1$ for $n < n_{\max} - 1$ (note that $0 \leq n < n_{\max}$). In the final implementation of this algorithm, $f(n_{\max}, n) = n_{\max} - n$, and a value of $n_{\max} = 3$ was used as a balance between evaluation speed and accuracy.

Lines 10–12 normalize the PSF such that the sum is equal to one and calculate the Fourier transform of the PSF. Lines 14–20 calculate the estimate of the object and normalizes it. In line 17, fov refers to a binary array that assures that $o$ does not exceed the field of view of each SA (a finite support constraint for the object estimate).

Unlike in the implementation of Wilding *et al.*, the estimated object is the absolute value of the inverse Fourier transform rather than its real component (see line 16). This choice of operator introduces slight discrepancies in the estimated object, which helps avoid a trivial solution where one single SA-PSF reduces to a delta function.

Figure 3 shows how the implementation of the TIP algorithm performs on a simulation of the optical system. It can be seen that the estimated SH pattern is nearly identical to the true SH pattern except for a stronger noise component, which is apparent on the PSFs in log scale.

It must be noted that the pre-processing algorithm does not reliably conserve the tip and tilt modes of the global wavefront. It is observed that the pre-processing algorithm introduces shifts in both the estimated SH pattern and estimated object. As long as the shifts are equal and in the same direction, these images provide a valid solution to the blind deconvolution problem. These shifts, however, corrupt the information of the tip and tilt modes of the wavefront. As a result, the estimated object can be shifted from its true position, but this does not have an inference on its resulting sharpness.

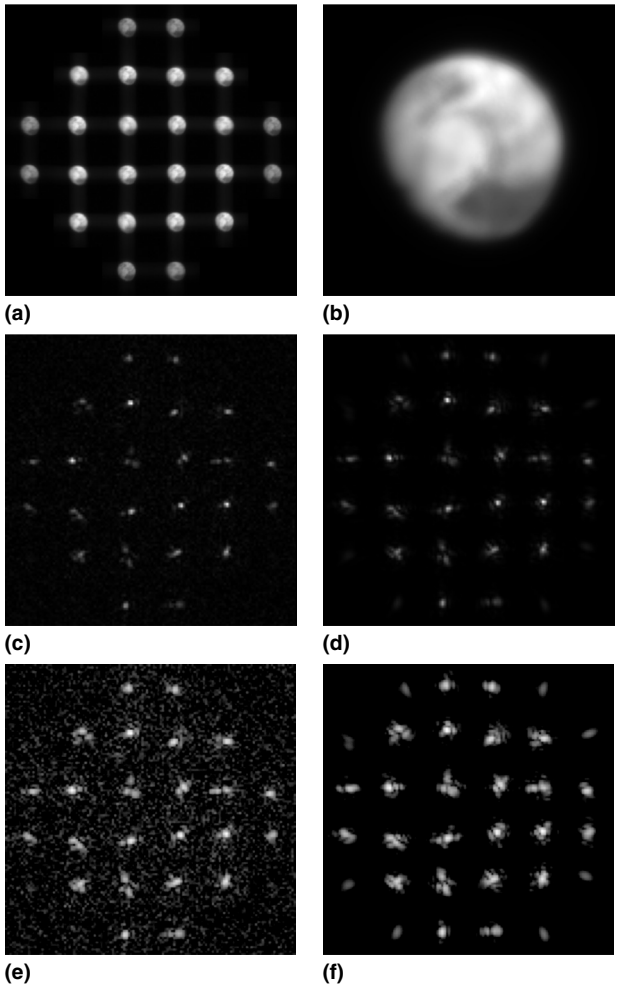


**Fig. 3.** Example of the performance of the TIP algorithm for extraction of a point-source SH pattern from an extended-source SH pattern, shown on a simulation of the optical system with a turbulence strength of $D/r_0 = 20$. In (c), some SA PSFs are bright and cover few pixels (i.e., are less aberrated), while others are dim and cover many pixels. By increasing the contrast of the SH patterns (e.g., by raising it by a power larger than one), the highly aberrated SAs are filtered out. This helps the TIP algorithm to find the correct object. (a) Simulated SH image. (b) Simulated image. (c) Estimated SH pattern. (d) True SH pattern. (e) Estimated SH pattern (logarithmic scale). (f) True SH pattern (logarithmic scale).

## 3. DEEP LEARNING WAVEFRONT SENSING

The CNN architecture used to reconstruct the wavefront given the processed SH image is, like [7,8], based on the *U-net* [16].

### A. Network Architecture

There are several key differences in SH patterns used as input for the CNN between literature [7,8] and this research. Most notably, previous research has used the true SH pattern, whereas the SH pattern used in this research is an estimate of the true SH pattern found through a blind-deconvolution step. The blind-deconvolution step not only introduces discrepancies in the shape of the individual PSFs, but also results in a lower signal-to-noise ratio.

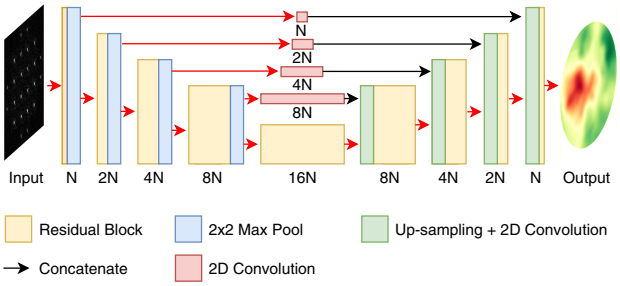


**Fig. 4.** Visualization of the general U-net architecture modified for the application in this paper, with $N = 12$. The convolution between the concatenation prevents noise present in the input image from propagating into deeper layers. The residual block used is displayed in Fig. 5.

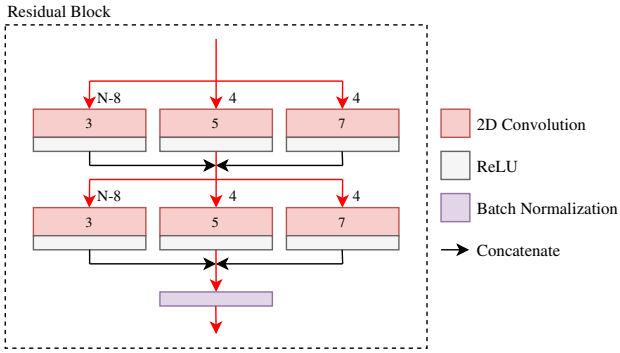


**Fig. 5.** Residual block used in the proposed CNN architecture. Similar to Hu *et al*. [7], different sizes of convolutional layers are used. The amount of larger convolutional layers is limited, and the concatenation of the layers happens twice in the residual block.

The architecture used by Hu *et al.* [7] is designed such that information can be preserved throughout the network without being filtered by introducing a *bypassing* layer in the residual block. This can be very useful when there is limited noise present in the SH PSF, but limits the architecture's ability to extract information from the pre-processed images used in this paper. For the same reasons, additional convolution layers are placed between the concatenations of input to output layers compared to the original U-net architecture, as shown in Fig. 4.

Figure 5 shows the residual block used for this proposed architecture. This residual block has features based on the architecture of Hu *et al.* [7], where different sizes of convolutional kernels are used. The amount of larger size kernels is, however, significantly reduced and does not depend on $N$. As a result of this reduction, the size of the network is much lower than the architecture of Hu *et al.* Additionally, there is another concatenation action introduced in the new residual block, located between the two convolutional layers. This concatenation action mixes the information of the differently sized convolutional layers and allows information to pass through two different sizes of kernels within the residual block.

### B. Training

For the training of the CNN, a numerical simulation of a DFWS system was constructed in Python. $2 * 10^5$ Kolmogorov phase screens were simulated with a turbulence strength ranging from $D/r_0 = 0$ to $D/r_0 = 18$ with removed piston, tip, and tilt modes. Details on the practical implementation of this process are described in [17].

The training of the CNN was done on a desktop PC with a 12 core Intel Xeon E5-2630 DUAL CPU with 64 GB of memory and a NVIDIA GeForce GTX 970 GPU with 4 GB of memory. The Python package Keras was used with a Tensorflow backend. The used loss function is *mean absolute error* (MAE) with the *Adam* optimizer and a batch size of 32. MAE was chosen rather than the more standard *mean squared error* (MSE) because of its better observed performance in this use case. The reason for this is likely the relatively lower penalty MAE puts on large errors, which occur mostly in higher-level wavefront modes, which the CNN cannot reconstruct due to sensor noise, normalization errors, and physical sensor limitations.

## 4. RESULTS

To testing the wavefront sensing capabilities of the newly developed method, the developed SH image pre-processing combined with three different DLWS networks is tested on a turbulence strength ranging from $D/r_0 = 0$ to $D/r_0 = 18$ in 19 discrete steps. The three different DLWS architectures are the architecture proposed in Section 3, the architecture proposed by Hu *et al.* [7], and the U-net based architecture used by Bekendam [8]. For each step in turbulence strength, 100 unique wavefronts were generated that were not included in the training data set. The tests were performed using the software simulation of the optical setup for easier calculation of the wavefront reconstruction performance.

To compare the developed methods to the conventional way of extended scene wavefront sensing, zonal and modal methods were implemented, too. For conventional methods, slope detection was done in accordance with Zhou *et al.* [18], using the absolute difference function as a correlation function combined with sub-pixel interpolation to retrieve SA shifts with an accuracy higher than one pixel. The modal method represents the estimated wavefront using the first 24 Zernike modes, equal to the number of effective SAs.

The experiment was conducted on a software simulation of the optical system as shown in Fig. 6, with a MLA of $6 \times 6$ SAs. The atmospheric turbulence follows a Kolmogorov spectrum, and it is approximated as a single phase screen located at the entrance pupil.

Figure 7 [19] shows the performance of the different methods. In this figure, the 1 rad RMS wavefront estimation error is highlighted, indicating the minimum required performance.

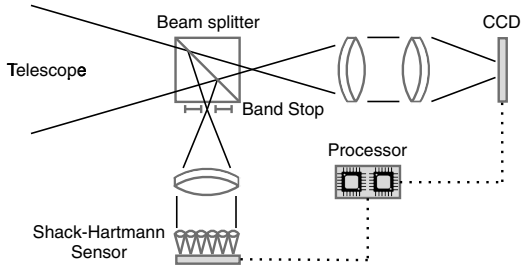


**Fig. 6.** Schematic overview of the optical setup of the used software simulation.

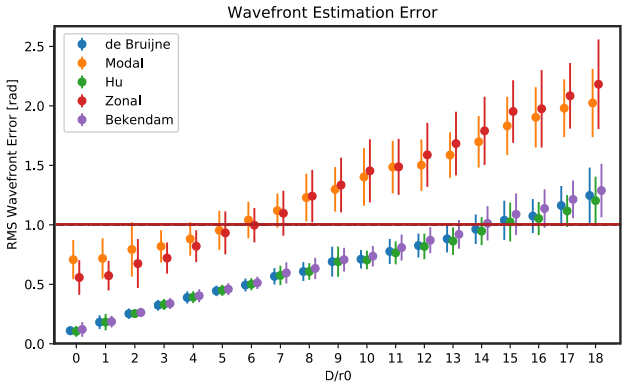


**Fig. 7.** Overview of the wavefront estimation performance of the different developed DLWS methods, compared to the traditional zonal and modal wavefront sensing methods. The wavefronts used for evaluation are 1900 unique wavefronts that were not included in the training data for the CNNs. For each discrete step in turbulence between $D/r_0 = 0$ and $D/r_0 = 18$, 100 wavefronts were generated. As objects, a variety of space objects with black backgrounds [19] were used that were converted to gray scale and covered between roughly 20% and 90% of the aperture. The dot represents the mean value over the 100 tested wavefronts, and the error line represents the standard deviation of the error. For readability, some data points are placed slightly before or after the integers on the $x$ axis. This does not reflect a difference in tested wavefront strength between the different methods.

The RMS wavefront error is calculated by

$$e_{\mathrm{RMS}} = \sqrt{\sum_{x=1}^{X}\sum_{y=1}^{Y}\left(\phi(x,y) - \tilde{\phi}(x,y)\right)^2}, \tag{5}$$

where $\tilde{\phi}$ is the reconstructed wavefront, and $\phi$ is the true wavefront of size $[X, Y]$.

As expected, the zonal method manages to sufficiently estimate the wavefront up to a turbulence strength of roughly $D/r_0 = 6$, which is equal to the linear amount of SAs in the SH sensor. Since the wavefront is generated using Kolmogorov statistics rather than Zernike polynomials, the performance of the modal method is lower than that of the zonal method. The performance of the tested CNNs is very similar, all estimating wavefront sufficiently up to a turbulence strength of $D/r_0 = 14$, which is a very significant improvement compared to the modal and zonal methods.

Up to $D/r_0 = 12$, the newly proposed architecture performs slightly, but not significantly, better than Hu's architecture, while both perform roughly 4% better than Bekendam's architecture. At turbulence strengths of $D/r_0 > 12$, Hu's architecture performs roughly 2% better than the proposed architecture.

Table 1 shows the computational efficiency of the evaluated architectures. Even though the proposed architecture reconstructs the wavefront with accuracy similar to Hu's architecture, the evaluation time and memory requirements are much lower, which is beneficial for real-time integration of the DLWS techniques.

Figure 8 shows an example of the wavefront reconstruction performance of the tested CNN architectures. It can be seen that the DLWS method is able to reconstruct the global shape of the wavefront accurately, but fails to retrieve the pixel-to-pixel fluctuations in the turbulent wavefront.

**Table 1. Computational Efficiency Evaluation of the Three Tested DLWS CNNs[a]**

| CNN | Number of Parameters | Evaluation Time [s] | Memory [MB] |
|---|---|---|---|
| Hu | $2.14 \cdot 10^7$ | 0.145 | 252 |
| Bekendam | $4.86 \cdot 10^5$ | 0.092 | 6.29 |
| de Bruijne | $1.45 \cdot 10^6$ | 0.099 | 18.4 |

[a]The evaluation times depend greatly on the amount of background processes active. These tests were performed while the other parts of the DFWS system were also running.

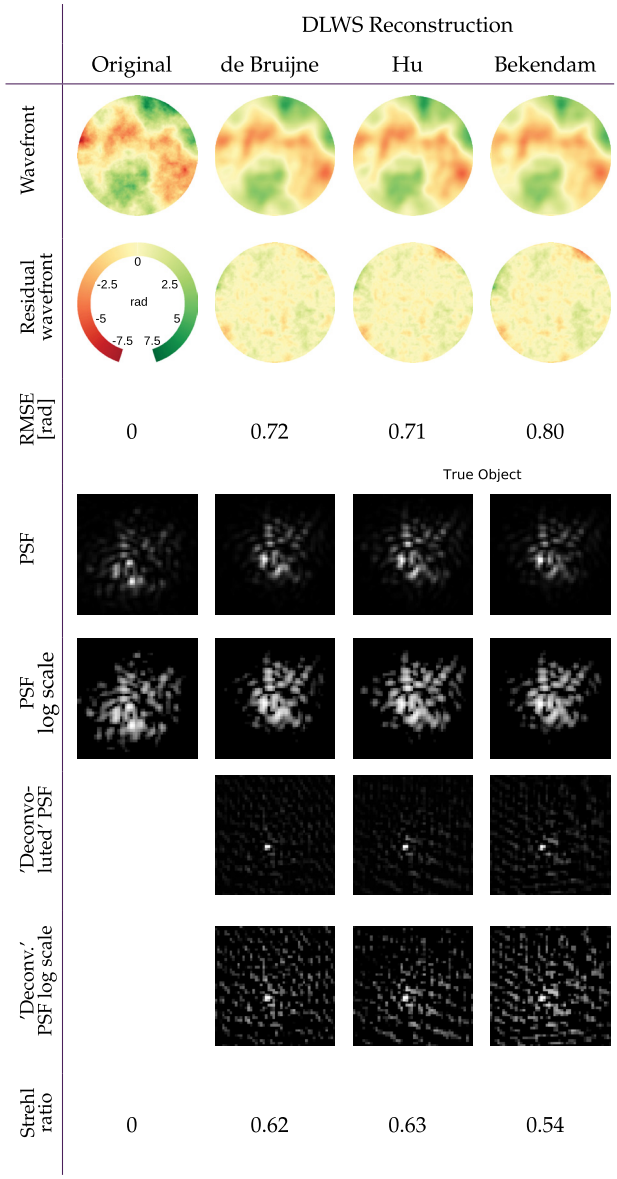


**Fig. 8.** Example of the wavefront sensing performance of the three tested DLWS architectures. The tested wavefront has a turbulence strength of $D/r_0 \approx 12$, and all networks reconstruct the wavefront sufficiently. The DLWS architectures perform overall very similar, and it does not appear that one particular architecture fails or succeeds to recognize wavefront features that the others do not.

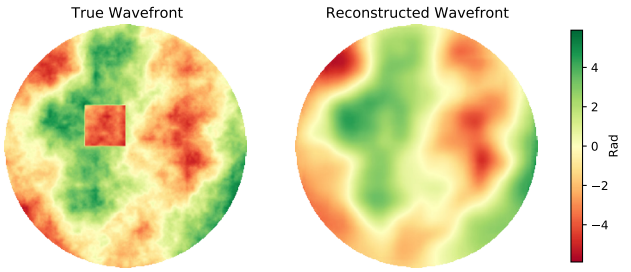


**Fig. 9.** Visualization of the phase unwrapping behavior of the DLWS methods. The DLWS methods have been trained to reconstruct a wavefront with a continuous phase; if a discontinuous jump in phase of $2\pi$ rad is added to the wavefront, this will not be present in the reconstructed wavefront.

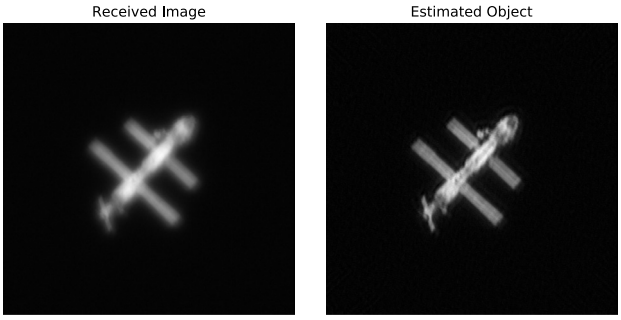


**Fig. 10.** Simulated performance of the proposed wavefront sensing system used in a DFWS system as described in [17]. The simulated turbulence has a strength of $D/r_0 = 15$, and the used SH sensor has a MLA of $6 \times 6$ SAs. Original image sourced from [20].

Figure 8 also shows the PSFs of reconstructed wavefronts, as well as the deconvolution of the true PSF with that of reconstructed wavefronts. This "deconvoluted" PSF depends on the used deconvolution technique but gives a rough idea of the quality of the corrected image in a DFWS setting. The noise in these PSFs makes the calculation of the Strehl ratio unreliable. Instead, the Strehl ratios of the PSFs resulting from the residual wavefronts are shown.

Interestingly, the DLWS methods show perfect phase unwrapping behavior. The training data for the CNNs consisted of continuous wavefronts (i.e., no large steps in phase delay), which led to the DLWS reconstructing only continuous wavefronts. This effect is highlighted in Fig. 9, where a jump in phase is manually added to the true wavefront, but the DLWS method reconstructs a wavefront with a phase wrap that corresponds to the turbulent wavefront without the added jump in phase.

Figure 10 [17,20] shows an example of the performance of the proposed system integrated in a DFWS setup. The simulated turbulence has a strength of $D/r_0 = 15$, which is on the limit of what the proposed system is capable of correcting using a SH sensor with a MLA of $6 \times 6$ SAs.

As part of the experiment, the developed method was implemented in an optical laboratory setup, as shown in Fig. 11. The physical experiment was used to verify the real-world performance of the system, as well as test its resilience to anisoplanatic wavefront aberrations. The turbulence simulator (custom-made

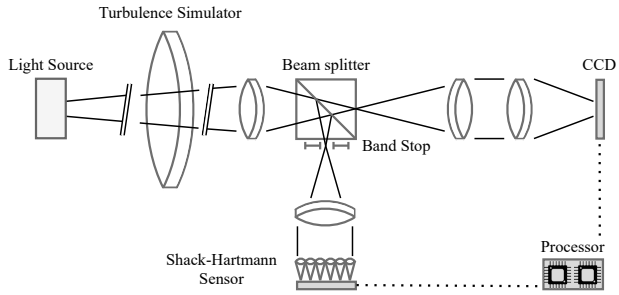


**Fig. 11.** Schematic overview of the optical setup used for real-world testing of the developed methods.

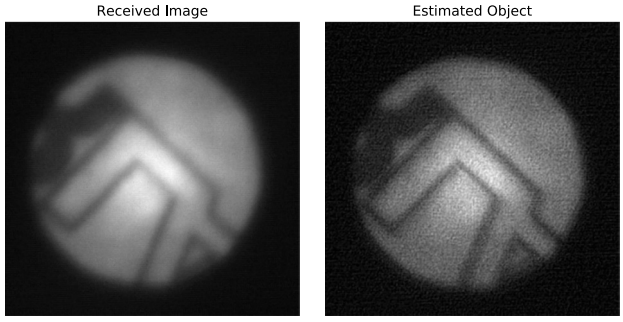


**Fig. 12.** Real-world performance of the proposed wavefront sensing system used in a DFWS system as described in [17]. The simulated turbulence has a strength of $D/r_0 \approx 18$, and the used SH sensor has a MLA of $6 \times 6$ SAs.

by Lexitek, Inc.) in the setup introduces known aberrations in the wavefront.

Figure 12 showcases the real-world performance of the system. As the turbulence strength is $D/r_0 \approx 18$, a diffraction limited correction is not expected, and the estimated image contains high-frequency error component.

## 5. CONCLUSION AND DISCUSSION

In this paper, it is shown that blind deconvolution methods have the potential to open up DLWS methods to extended-scene observations. The proposed system outperforms conventional extended-scene wavefront sensing methods significantly in terms of wavefront reconstruction resolution. Using this approach, a $6 \times 6$ microlens SH sensor was shown to be able to reconstruct the wavefront up to a turbulence strength of $D/r_0 = 14$. Using a conventional wavefront sensing method, a $14 \times 14$ microlens SH sensor would be necessary to achieve this accuracy. As a result of this reduction in microlenses, more than five times the amount of light becomes available per SA, allowing for the wavefront correction in low-light scenes.

The proposed wavefront sensing method is still in early development and currently has a number of limitations and questions.

It was observed that the TIPs algorithm used for preprocessing the SH image occasionally introduces shifts in the estimated SH pattern. As a result of this, the proposed DFWS system was not able to correct the tip and tilt modes of the wavefront. Additional constraints on the PSF and object could be added in

the TIP implementation to eliminate the occurrence of these shifts.

The developed system is currently limited to extended objects surrounded by a dark background, which limits it employment in, for example, satellite or surveillance imaging. It is expected that using a field stop to introduce a dark background into the SA images can elevate this limitation, but this needs to be verified in future research.

This paper investigated the performance of the system using a $6 \times 6$ MLA. Bekendam [8] showed promising wavefront reconstruction capabilities for larger MLA sizes using point-source DLWS techniques. This trend is believed to be true for the proposed system as well, but future research will have to verify this assumption.

More research is needed to explore the potential of the proposed wavefront sensing method in other areas of AO. For example, extended-scene DLWS can be extended to anisoplanatic imaging or combined with the frozen flow hypothesis to utilize temporal correlations of turbulence for more accurate wavefront reconstruction.

**Disclosures.** The authors declare no conflicts of interest.

**Data availability.** Data underlying the results presented in this paper are not publicly available but can be generated using the resources available in [19].

## REFERENCES


1. D. R. Neal, J. Copland, and D. A. Neal, "Shack-Hartmann wavefront sensor precision and accuracy," Proc. SPIE **4779**, 148–160 (2002).
2. M. Viegers, E. Brunner, O. Soloviev, C. C. de Visser, and M. Verhaegen, "Nonlinear spline wavefront reconstruction through moment-based Shack-Hartmann sensor measurements," Opt. Express **25**, 11514–11529 (2017).
3. H. Guo, N. Korablinova, Q. Ren, and J. Bille, "Wavefront reconstruction with artificial neural networks," Opt. Express **14**, 6456–6462 (2006).
4. R. Swanson, M. Lamb, C. Correia, S. Sivanandam, and K. Kutulakos, "Wavefront reconstruction and prediction with convolutional neural networks," Proc. SPIE **10703**, 481–490 (2018).
5. S. L. Suárez Gómez, C. González-Gutiérrez, E. Díez Alonso, J. D. Santos Rodríguez, M. L. Sánchez Rodríguez, J. Carballido Landeira, A. Basden, and J. Osborn, "Improving adaptive optics reconstructions with a deep learning approach," in *Hybrid Artificial Intelligent Systems*, F. J. de Cos Juez, J. R. Villar, E. A. de la Cal, Á. Herrero, H. Quintián, J. A. Sáez, and E. Corchado, eds. (Springer International Publishing, 2018), pp. 74–83.
6. L. Hu, S. Hu, W. Gong, and K. Si, "Learning-based Shack-Hartmann wavefront sensor for high-order aberration detection," Opt. Express **27**, 33504–33517 (2019).
7. L. Hu, S. Hu, W. Gong, and K. Si, "Deep learning assisted Shack–Hartmann wavefront sensor for direct wavefront detection," Opt. Lett. **45**, 3741–3744 (2020).
8. J. Bekendam, "Deep learning wavefront sensing via raw Shack-Hartmann images," Master's thesis (Delft University of Technology, 2020).
9. F. Sánchez-Lasheras, C. Ordóñez, J. Roca, and F. de Cos Juez, "Real-time tomographic reconstructor based on convolutional neural networks for solar observation," Math. Methods Appl. Sci. **43**, 8032–8041 (2019).
10. O. Soloviev, H. Thao Nguyen, V. Bezzubik, N. Belashenkov, G. Vdovin, and M. Verhaegen, "Shack-Hartmann sensor as an imaging system with a phase diversity," arXiv:2011.00891 (2020).
11. D. Wilding, O. Soloviev, P. Pozzi, G. Vdovin, and M. Verhaegen, "Blind multi-frame deconvolution by tangential iterative projections (tip)," Opt. Express **25**, 32305–32322 (2017).
12. G. R. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," Opt. Lett. **13**, 547–549 (1988).
13. L. P. Yaroslavsky and H. J. Caulfield, "Deconvolution of multiple images of the same object," Appl. Opt. **33**, 2157–2162 (1994).
14. F. Sroubek and P. Milanfar, "Robust multichannel blind deconvolution via fast alternating minimization," IEEE Trans. Image Process. **21**, 1687–1700 (2012).
15. S. Chaudhuri, R. Velmurugan, and R. Rameshan, *Blind Image Deconvolution* (2014).
16. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," arXiv:1505.04597 (2015).
17. V. de Bruijne, "Extended scene deep learning wavefront sensing for real time image deconvolution," Master's thesis (Delft University of Technology, 2021).
18. H. Zhou, L. Zhang, L. Zhu, H. Bao, Y. Guo, X. Rao, L. Zhong, and C. Rao, "Comparison of correlation algorithms with correlating Shack-Hartmann wave-front images," Proc. SPIE **10026**, 196–207 (2016).
19. V. de Bruijne, "Extended scene deep learning wavefront sensing for real time image deconvolution, code repository," Master's thesis (Delft University of Technology, 2021).
20. NASA, *14 Years Ago: The First Crew Moves into Space Station* (2014).