Shack–Hartmann wavefront sensor with array of phase coded masks

NITIN DUBEY, RAVI KUMAR, JOSEPH ROSEN

School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 8410501, Israel
*Corresponding author: nitinyog@post.bgu.ac.il

Shack–Hartmann wavefront sensors (SHWS) are generally used to measure the wavefront curvature of light beams. Measurement accuracy and the sensitivity of these sensors are important factors for better wavefront sensing. In this study, we demonstrate a new type of SHWS with better measurement accuracy than the regular SHWS. The lenslet array in the regular SHWS is replaced with an array of coded phase masks and the principle of coded aperture correlation holography (COACH) is used for wavefront reconstruction. Sharper correlation peaks achieved by COACH improve the accuracy of the estimated local slopes of the measured wavefront and consequently improve the reconstruction accuracy of the overall wavefront. Experimental results confirm that the proposed method provides a lower mean square wavefront error by one order of magnitude in comparison to the regular SHWS.

Shack–Hartmann wavefront sensor (SHWS) was developed for optical metrology purposes [1] and later has been used in various other fields, such as adaptive optics [2], microscopy [3], retinal imaging [4], and high-power laser systems [5]. Usually, in SHWS a microlens array is used to measure the wavefront local gradients. For this, the incident wavefront is spatially sampled such that each lenslet focuses the local sub-aperture into an imaging sensor located at the focal plane of the lenslet. The position of the focused spot corresponds to the average slope of the local wavefront of each microlens. The local slope of the wavefront can be evaluated by calculating the spot displacement from the optical axis. Afterward, a wavefront reconstruction algorithm is utilized to estimate the complete wavefront using these estimated slopes [6].

Recently, an incoherent digital holography technique called coded aperture correlation holography (COACH) [7,8] has been developed. In this study, we present a new technique of COACH-based SHWS that increases the accuracy of wavefront sensing. Instead of using the regular lenslet array, a coded phase mask (CPM) is introduced in each sub-aperture of the array. Based on several recently published studies [9–11] about interferenceless COACH, we conclude that using a certain CPM can yield at the end of the process a spot that is narrower than the focused spot of the regular SHWS. In any SHWS, the size of each spot constraints the sensitivity and accuracy of the measured slope of the local wavefront. Thus, by use of the CPM array, one can achieve narrower correlation peaks which improve the accuracy of the measured wavefront.

Modified Gerchberg-Saxton algorithm (GSA) is utilized for synthesizing the CPMs [9–11]. The CPMs displayed on a spatial light modulator (SLM) generate an ensemble of sparse randomly distributed intensity dots on the sensor plane. The ensemble of dots obtained from the measured wavefront are cross-correlated with a reference dot pattern created by illuminating a single CPM with a plane wave. The intensity response in each cell of the array is shifted according to the average slope of the wavefront in the corresponding cell. Composing the entire slopes yields a three-dimensional curvature used as an approximation to the wavefront incident on the SHWS.

The optical scheme of COACH-based SHWS is shown in Fig. 1(a). A plane wave passes through a phase object and the emitted wavefront is sampled by an array of identical phase elements each of which is a product of the CPM and a lens transmittance. The lens performs a Fourier transform which according to the GSA is needed for the response of the dots on the camera plane. As in conventional SHWS, in front of each CPM, the local wavefront emitted from the phase object is approximated to a tilted plane wave. The goal of the system is to measure the tilt angle of each local planar wavefront and to estimate the global curvature by fusing all the local angles. As in ordinary SHWS, estimating the tilt angle is done by measuring the shift of the intensity response recorded on the camera. Since in COACH-based SHWS, this intensity response is the ensemble of dots, their common shift is measured by cross-correlation with a reference pattern. As much as the cross-correlation peak is sharper, the accuracy of the measurement is higher. Therefore, to narrow the peak as much as possible, the cross-correlation is done by a non-linear process with two parameters optimized to yield the sharpest cross-correlation peak.

We start the formal analysis by calculating the distribution of the reference pattern. Each cell of the size $a\times a$ in the CPM designed by the GSA has the distribution of $t(x,y)=\exp\{\phi(x,y)\} \text{Rect}(x/a,y/a)$, where $\phi(x,y)$ is a chaotic phase function and $\text{Rect}(x/a,y/a)=1$ for all $|x|,|y|<a/2$ and 0 otherwise. $t(x,y)$ is attached to a diffractive microlens with focal length. It is well-known that illuminating these combined phase elements at a single cell of the array with a plane wave yields the following intensity [12],

$$I(x,y) = \left| \int \frac{1}{\lambda f} \mathcal{F} \{t(x,y)\} \right|^2,$$  \hspace{1cm} (1)
where $\mathcal{F}$ is 2D Fourier transform and $v_f$ is the scaling operator such that $v_f(1/\lambda f) = g(\lambda x, \lambda y)$, $l_n(x,y)$ plays the role of the reference pattern in the cross-correlation. As mentioned above, $g(x,y)$ is designed by the GSA to yield an ensemble of sparse randomly distributed dots as a result of a Fourier transform of $\exp[i\phi(x,y)]$ [9]. Therefore, 

$$I_g(x,y) = v_f(1/\lambda f) \mathcal{F} \{ \exp[i\phi(x,y)] \}$$

$$= \sum_{p=1}^P C_p \sin^2 \left( \frac{x-x_p}{\lambda f} + y-y_p}{\lambda f} \right) \right),$$

(2)

where

Cross-correlation between $I_{m,n}(x,y)$ and $I_n(x,y)$ yields a correlation peak at the distance $f \tan \theta_m f \tan \theta_n$ from the center of the $(m,n)$-th subarea. Measuring this shift of the peak gives an estimation on the value of the tilt angle of $(\theta_m, \theta_n)$. Therefore, to further increase the sharpness of the correlation peaks. Therefore, to further increase the accuracy of the tilt angle measurement, we tested the method. The average slope of the tested wavefront in each cell is directly assessed by calculating the deviation of the reconstructed wavefront from the digitally simulated wavefront. The MSE is calculated as,

$$\text{MSE} = \frac{1}{L \cdot K} \sum_{l=1}^{L} \sum_{k=1}^{K} \left( \tilde{I}_{m,n} - W_{m,n} \right)^2,$$

(6)

where $\tilde{I}_{m,n}$ and $W_{m,n}$ are the digital and experimental reconstructed wavefronts, respectively. $L$ are the numbers of data points in the wavefront. In case the measured wavefront is a-priory unknown, two strategies can be followed. First, the values of $\alpha$ and $\beta$ that yield the sharpest peak are the values that satisfy the equation $\alpha + \beta = 0$. However, the noisy experimental environment usually yields different optimal parameters and may not follow this equation. In this study, the optimal parameters $\alpha$ and $\beta$ of the NCC are found by minimizing the mean square error (MSE), because minimum MSE is the measure of accuracy which we aim to increase. The MSE assesses the reconstruction error by calculating the deviation of the reconstructed wavefront from the digitally simulated wavefront. The MSE is calculated as,
option of using bipolar functions by recording two camera shots with two independent CPMs and with two different dot responses, where one response is subtracted from the other. In this scheme, the bipolar reference pattern is,

\[ I_{a}(x, y) = \nu \left[ \frac{1}{\lambda f} \right] \mathcal{F} \left\{ t_1(x, y) \right\} - \nu \left[ \frac{1}{\lambda f} \right] \mathcal{F} \left\{ t_2(x, y) \right\}, \] (7)

where \( t_1(x,y) = \exp(i \phi(x,y)) \text{Rect}(x/a, y/a) \), and \( \phi(x,y) \) is the phase distribution of the \( j \)th CPM \((j=1,2)\). \( I_{a}(x, y) \) is also obtained by two exposures with the same two CPMs so that its distribution is,

\[ I_{a,n}(x, y) = \nu \left[ \frac{1}{\lambda f} \right] \mathcal{F} \left\{ \exp \left( \frac{2\pi}{\lambda} \left( x \sin \theta_n + y \sin \theta_n \right) \right) t_1(x, y) \right\} - \nu \left[ \frac{1}{\lambda f} \right] \mathcal{F} \left\{ \exp \left( \frac{2\pi}{\lambda} \left( x \sin \theta_n + y \sin \theta_n \right) \right) t_2(x, y) \right\} \]

\[ \geq \sum_{p=1}^{m} C_{p} \sin^2 \left( x - x_{p} - f \tan \theta_n \right) + \frac{x - x_{p} - f \tan \theta_n}{\lambda f / a} x - y_{p} - f \tan \theta_n \] (8)

\[ \frac{f}{\lambda} \left( x - x_{p} - f \tan \theta_n \right) + \frac{x - x_{p} - f \tan \theta_n}{\lambda f / a} y - y_{p} - f \tan \theta_n \]

where the two random series \( \{x_{1,1}, \ldots, x_{b}\} \) and \( \{y_{1,1}, \ldots, y_{b}\} \) are independent. The rest of the process is identical to the above-mentioned description of the process of the unipolar patterns.

The schematic of the experimental setup for COACH-based SHWS is shown in Fig. 1(b). A HeNe laser with a beam size of 8mm is used as the source, where a pair of lenses \((f_{1}=10cm; f_{2}=48cm)\) is used to expand the beam diameter from 8mm to about 38mm. A pinhole of 20μm diameter used as a lowpass filter is placed at the common focal plane of the two lenses. The lowpass filter makes the wavefront after the beam expander as close as possible to a plane wave. This plane wave propagates through a phase object such that its phase distribution is represented by the wavefront emitted from the object. In the experiment, a positive lens of 30cm focal length is used as the phase object. The beam is polarized to the active orientation of the SLM (Holoeye PLUTO, 1920x1080 pixels, 8μm pixel pitch). The array of coded phase patterns displayed on the SLM is obtained by modulo-2π phase addition of the CPMs with the digital wavefront. The digital wavefront is created by numerical Fresnel propagation from the phase object to the SLM plane using the digital wavefront. The digital wavefront is created by numerical Fresnel propagation from the phase object to the SLM plane using MATLAB with the same parameters as the experimental setup. The MSE was calculated between the digital wavefront and the tested wavefront reconstructed by the regular and by the COACH-based SHWS systems. For COACH-based SHWS, both techniques, unipolar and bipolar, were tested. To find the optimal number of dots of the sparse response, numbers of 3, 5, 10, and 12 dots were tested.

Figures 2(a) and 2(a) present the intensity response for two different CPMs when only the central cell is activated in the array. The dot structure is clearly seen in both figures and in the bipolar pattern of Fig. 2(a) obtained as the difference of the intensities of Figs. 2(a) and 2(a). Figures 2(b) and 2(b) show the intensity response in the center of the COACH-based array equipped with two different CPMs, where the tested wavefront is introduced into the system. The corresponding bipolar response is shown in Fig. 2(b). Figure 3 shows the experimental results of the NCC in comparison with the focal spot of a single microlens shown in Fig. 3(a). The autocorrelation of Fig. 2(a) by the NCC technique with \((\alpha=0.0; \beta=1.0)\) and \((\alpha=0.1; \beta=0.9)\) is shown in Figs. 3(b) and 3(c), respectively. The autocorrelation of Fig. 2(a) by NCC technique with \((\alpha=0.0; \beta=1.0)\) and \((\alpha=0.1; \beta=0.9)\) is shown in Figs. 3(d) and 3(e), respectively. The NCC parameters \((\alpha, \beta)\) can vary between -1 to +1 and are optimized by searching the minimum MSE between the experimental and the digital wavefront. With the parameters \((\alpha=0.0; \beta=1.0)\) and \((\alpha=0.1; \beta=0.9)\), the wavefront reconstruction result was the best with the lowest MSE in comparison to other values. Figure 3(f) shows the intensity response of regular SHWS for the test wave. Figure 3(h) shows the NCC of Fig. 2(a) with Fig. 2(b) [when the scale of 2(a) is the same as one cell of 2(b)]. The central part of two intensity patterns on the sensor for the two CPM arrays and for the input of the test wavefront; (b) the corresponding bipolar pattern obtained as the difference between (b) and (b).
bipolar (overall 6 dots), the accuracy of COACH-based SHWS is calculated. MSE values can be noted that for the best case of 3-dot patterns are compared with the regular SHWS and the results confirm the validity and effectiveness of our method. From the experimental results with a different number of background noise. The experimental results with a different number of sparse dots are given in Table 1. From the error maps in Figs. 4(a) and 4(b), it can be seen that the boundary region of the wavefront extremely deviated from the digital shape resulting in higher MSE values because of the peripheral areas. Therefore, the 3D error maps for the centrally cropped 800×800 data points for both cases are shown in Figs. 4(a) and 4(b), respectively. From the flat error map in Fig. 4(b), it is clear that the reconstructed wavefront of the bipolar method has a closer fit to the digital wavefront. MSE values of the centrally cropped wavefront are also calculated and given in Table 2. The MSE of 5-dot patterns for the unipolar case and most of the 3-dot patterns for the bipolar method yield minimum MSE among the tested cases and these minimum values are bolded in the tables. The case of 3-dot bipolar COACH-based SHWS has a smaller MSE by order of magnitude in comparison to the regular SHWS. The lower MSE values confirm the improvement in the wavefront reconstruction accuracy of the COACH-based SHWS. Furthermore, it should be noted that all the tested bipolar reconstruction results are better than the 5-dot bipolar but still the 5-dot unipolar gives better reconstruction results in comparison to the regular SHWS and to unipolar cases with other dot numbers.

In conclusion, we have presented a new SHWS technique with improved measurement accuracy. Coded phase masks are synthesized using the modified GSA algorithm to produce sparse dot patterns. Further, the bipolar sparse dot patterns are generated to reduce background noise. The experimental results with a different number of dot patterns are compared with the regular SHWS and the results confirm the validity and effectiveness of our method. From the calculated MSE values it can be noted that for the best case of 3-dot bipolar (overall 6 dots), the accuracy of COACH-based SHWS is improved by one order of magnitude in comparison with the regular SHWS. The cost of such improvement is the need for two camera shots. In case the speed of the measurement is more critical than the accuracy, the single-shot unipolar COACH-based SHWS can offer approximately double the accuracy than of the regular SHWS.

### Table 1. MSE values of SHWS for different number of dots

| No. of Sparse Dots | Regular SHWS | Unipolar (α=0.0; β=1.0) | Unipolar (α=0.1; β=0.9) | Bipolar (α=0.0; β=1.0) | Bipolar (α=0.0; β=1.0) |
|--------------------|--------------|--------------------------|--------------------------|------------------------|------------------------|
| 3                  | 0.01089      | 0.10920                  | 0.07920                  | 0.00075                | 0.00150                |
| 5                  | 0.01089      | 0.00944                  | 0.00889                  | 0.00181                | 0.00196                |
| 10                 | 0.01089      | 0.08654                  | 0.05396                  | 0.00563                | 0.00313                |
| 12                 | 0.01089      | 0.06623                  | 0.04522                  | 0.00528                | 0.00612                |

### Table 2. MSE value of centrally cropped wavefront

| No. of Sparse Dots | Regular SHWS | Unipolar (α=0.0; β=1.0) | Unipolar (α=0.1; β=0.9) | Bipolar (α=0.0; β=1.0) | Bipolar (α=0.0; β=1.0) |
|--------------------|--------------|--------------------------|--------------------------|------------------------|------------------------|
| 3                  | 0.00705      | 0.12840                  | 0.08820                  | 0.00045                | 0.00130                |
| 5                  | 0.00705      | 0.00379                  | 0.00378                  | 0.00087                | 0.00056                |
| 10                 | 0.00705      | 0.00169                  | 0.05426                  | 0.00310                | 0.00192                |
| 12                 | 0.00705      | 0.00564                  | 0.04280                  | 0.00246                | 0.00165                |

### References
1. S. Goelz, J. J. Persoff, G. D. Bittner, J. Liang, C.-F. T. Hsueh, and J. F. Bille, Active. “New wavefront sensor for metrology of spherical surfaces,” Proc. SPIE 1542, Active and Adaptive Optical Systems, (1 December 1991)
2. D. Dayton, J. Gonglewski, B. Pierson, and B. Spielbusch, “Atmospheric structure function measurements with a Shack–Hartmann wave-front sensor,” Optics Letters 17, 1737-1739 (1992).
3. X. Tao, B. Fernandez, O. Azucena, M. Fu, D. Garcia, Y. Zuo, D. C. Chen, and J. Kubby, “Adaptive optics confocal microscopy using direct wavefront sensing,” Optics Letters 36, 1062-1064 (2011).
4. Y. Zhang, C. A. Girkin, J. L. Duncan, and A. Roorda, “Adaptive optics scanning laser ophthalmoscopy,” Advanced Biophotonics: Tissue Optical Sectioning 10, 507-557 (2013).
5. A. Talmi and E. N. Ribak, “Wavefront reconstruction from its gradients” Journal of the Optical Society of America A 23, 288-297 (2006).
6. V. Yanovsky, “Generation and characterization of the highest laser intensities (10^{21} W/cm^2),” Optics Letters 29, 2837-2839 (2004).
7. V. Yanovsky, “Generation and characterization of the highest laser intensities (10^{21} W/cm^2),” Optics Letters 29, 2837-2839 (2004).
8. J. Rosen, “Interferenceless coded aperture correlation holography—a new technique for recording incoherent digital holograms without two-wave interference” Optics Express 25, 13883-13886 (2017).
9. J. Rosen, A. Vijayakumar, M. R. Rai, S. Mukherjee, and A. Bulbul “Review of 3D Imaging by Coded Aperture Correlation Holography (COACH),” Appl. Sci. 9(3), 605 (2019).
10. M. R. Rai and J. Rosen, "Noise suppression by controlling the sparsity of the point spread function in interferenceless coded aperture correlation holography (I-COACH)," Optics Express 27, 24311-24323 (2019).
11. M. R. Rai, A. Vijayakumar, and J. Rosen, “Non-linear adaptive three-dimensional imaging with interferenceless coded aperture correlation holography (I-COACH),” Optics Express 26, 18143 (2018).
12. N. Dubey, J. Rosen, and I. Gannot, “High-resolution imaging system with an annular aperture of coded phase masks for endoscopic applications,” Optics Express 28, 15122-15137 (2020).
13. J. W. Goodman, “Introduction to Fourier Optics,” Roberts & Co. (2005).
13. C. Liu, T. Man, and Y. Wan, "Optimized reconstruction with noise suppression for interferenceless coded aperture correlation holography," \textit{Appl. Opt.} \textbf{59}, 1769-1774 (2020).

14. G. Dai, "Wavefront Optics for Vision Correction," (SPIE Press, 2008, n.d.).

15. W. H. Southwell, "Wave-front estimation from wave-front slope measurements" \textit{Journal of the Optical Society of America} \textbf{70}, 998-1006 (1980).