



Communication

Improvement in Signal Phase Detection Using Deep Learning with Parallel Fully Connected Layers

Michito Tokoro¹ and Ryushi Fujimura^{1,2,3,*}

- ¹ Graduate School of Regional Development and Creativity, Utsunomiya University, 7-1-2 Yoto, Utsunomiya 321-8585, Japan
- ² Center for Optical Research and Education (CORE), Utsunomiya University, 7-1-2 Yoto, Utsunomiya 321-8585, Japan
- ³ Institute of Industrial Science, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
- * Correspondence: fujimura_r@cc.utsunomiya-u.ac.jp

Abstract: We report a single-shot phase-detection method using deep learning in a holographic data-storage system. The error rate was experimentally confirmed to be reduced by up to three orders of magnitude compared with that in the conventional phase-determination algorithm by learning the light-intensity distribution around a target signal pixel. In addition, the output speed of a signal phase could be shortened by devising a network and arranging the fully connected layers in parallel. In our environment, the phase-output time of a single-pixel classification was approximately 18 times longer than that in our previous method, with the minimum-finding algorithm. However, it could be reduced to 1.7 times or less when 32 pixels were simultaneously classified. Therefore, the proposed method can significantly reduce the error rates and suppress the phase-output time to almost the same level as that in the previous method. Thus, our proposed method can be a promising phase-detection method for realizing a large-density data-storage system.

Keywords: holographic data storage; single-shot detection; interpixel crosstalk; deep learning

1. Introduction

Holographic data storage [1] is an optical memory system with architecture that is different from conventional optical discs such as Blu-ray, which uses the principle of holography [2] for the recording and readout of stored information. Because the twodimensionally arranged bit data (page data) are handled as signal information, the datatransfer rate is extremely high [3], and a large recording density can be expected owing to multiplex recording [4–12] in which the page data are overwritten at the exact location of the recording medium. Reports have been presented on the experimental demonstrations of 2 TB capacity and 1 Gbit data-transfer rates [13]. The recording density and data-transfer rate can be improved by increasing the code rate. The code rate denotes the amount of information (number of bits) carried by one pixel. Many researchers have employed and investigated multilevel phase-modulated signals instead of the conventional binary intensity-modulated signal to increase the code rate [14-17]. However, in this case, because an imager such as a charge-coupled device (CCD) camera can only acquire the light-intensity information, interference measurement such as the four-step phase-shifting method [18] is required to detect the phase signal. In general, interference measurement requires another light wave (phase-detection reference wave) to interfere with the diffracted light on the imager surface, which complicates the optical system and makes signal detection susceptible to mechanical vibration.

To address this issue, we have recently proposed a new method that can stably determine the signal phase in a single shot without using a phase-detection reference wave [19]. In this method, known phase-reference pixels are embedded in the signal image. The signal phase is determined from the light-intensity information at the pixel



Citation: Tokoro, M.; Fujimura, R. Improvement in Signal Phase Detection Using Deep Learning with Parallel Fully Connected Layers. *Photonics* **2023**, *10*, 1006. https:// doi.org/10.3390/photonics10091006

Received: 14 July 2023 Revised: 25 August 2023 Accepted: 31 August 2023 Published: 3 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). boundaries between the known phase pixel and signal pixel. So far, we have demonstrated the effectiveness of this method in both simulation and experiment. A sufficiently low phase-detection error rate can be obtained in a four-level phase-modulated signal. However, our recent investigation has revealed that the leakage of light waves from nonadjacent distant pixels affects the phase detection and causes errors. To reduce detection errors, we need to introduce a broader range of pixel information into the phase-determination algorithm. However, building such an algorithm is difficult because of its complexity.

In the present study, deep learning [20] is introduced to develop a phase-determination algorithm that considers information on distant and adjacent pixels. Deep learning is a method that has achieved excellent results in the field of image classification [21], language translation [22], drug discovery [23], and holographic data storage [24–28]. Our proposed algorithm utilizes deep learning to determine the phase signal from the intensity pattern. A significant reduction in the phase-detection errors is experimentally demonstrated in an algorithm with deep learning compared with the conventional phase-determination algorithm based on boundary intensity. Furthermore, we significantly improve the computational speed of the phase determination. By devising a network structure, we obtain the signal phase at the same computational speed as the conventional method even with deep learning, which requires a substantial computational cost.

2. Methods

2.1. Principle of Phase Detection Using Interpixel Crosstalk

The concept of phase detection is based on a recently proposed method [19]. This method takes advantage of the interpixel crosstalk produced by spatial filtering on the Fourier plane of an input image. This interpixel crosstalk causes the light wave of each pixel to leak into the other pixels and interfere near the pixel boundaries. We assume that reference pixels with known phases are appropriately introduced in the page data. In this case, the signal phase can be determined from the intensity distribution at the pixel boundaries. For example, in the case of a four-level phase signal, four different known phase pixels are placed around the signal pixel, as shown in Figure 1. Among the light intensities at the pixel boundaries above, below, left, and right of the signal pixel, we determine the known phase pixel that exhibits the minimum boundary intensity. Because such a minimum boundary intensity results from a destructive interference, the signal phase can be determined as the opposite of the known phase. This conventional phase-determination algorithm is called the minimum-finding algorithm.



Figure 1. Arrangement of the known phase-reference pixel in a four-level phase signal. ϕ_s represents the signal phase.

However, not only do the light waves that reach the pixel boundaries come from adjacent pixels, but sufficient light waves may also leak from distant pixels, depending on the size of the aperture for spatial filtering. Because this light wave leakage from distant pixels is not considered in the previous minimum-finding algorithm, signal-detection errors may occur.

In the present study, supervised machine learning is employed to consider the influence of these distant pixels. Figure 2 shows that the intensity distributions of the 3×3 pixels that surround the target signal are extracted from the intensity image and trained using the signal phase as the label. Because the extracted light-intensity distribution contains information on the light waves coming from the surrounding pixels as well as the adjacent known phase pixels, this learning corresponds to learning the influence of distant pixels.

Detected intensity image

Labeling of the signal phase $\pi/2$

Figure 2. Preparation of the training data from the intensity image detected by the imager. The image with 3×3 pixels that surround the target signal pixel is extracted and added with the label of the corresponding signal phase.

2.2. Acquisition of Experimental Training Data

Input phase image

Figure 3 shows the experimental optical setup to validate our proposed method. A single-mode semiconductor laser (LM405-PLR40, ONDAX, Inc., California, United States) that operates at a wavelength of 405 nm is first passed through a spatial filter to form a clean spherical wave. Subsequently, it is phase-modulated using a spatial light modulator (SLM) (X10468-05, Hamamatsu Photonics K.K., Shizuoka, Japan) to add a phase signal to the laser beam. The phase-modulated signal light is focused by a lens with a focal length of 200 mm and then low-pass-filtered using a variable square aperture (SLX-1, SIGMAKOKI Co., Ltd., Tokyo, Japan) inserted into the Fourier plane. Finally, the filtered signal light is imaged by a CCD camera (PL-B953U, Pixelink, Ontario, Canada) that detects the intensity distribution with an 8-bit resolution. Note that we directly detect the intensity pattern passing through the aperture, omitting the holographic recording material. This means that we neglect the influence of holographic reconstruction on the spatial frequency components of the signal image. However, it is a reasonable assumption if the aperture size determines the in-plane hologram size. In this case, even though holographic reconstruction will change the spatial frequency components owing to the off-Bragg diffraction, its influence is identical to that caused by the aperture. The pixel pitch of the SLM and CCD camera is 20 and 4.65 mm, respectively. Each data pixel of the input page data is composed of 4×4 pixels in SLM. Thus, the pixel pitch of the data pixel is 80 μ m. Nyquist size *w* [29,30] is expressed by Equation (1).

U

$$v = \frac{f\lambda}{a},\tag{1}$$

where f is the focal length of the lens, λ is the wavelength of the light source, and *a* is the pixel pitch. The Nyquist size of our experimental setup is approximately 1 mm. The aperture size normalized by this Nyquist size is defined as the Nyquist ratio. The experiment is conducted at various Nyquist ratios using apertures that can be adjusted in $10 \mu m$ increments. The optical lens system magnifies the signal image by 1.5. One data pixel corresponds to 26×26 pixels in the CCD camera. After the data acquisition, one data pixel is resampled from 26×26 pixels to 4×4 pixels. Each page data sample consists of 20 \times 20 data pixels in which 162 data pixels are assigned as signal pixels. A total of 900 page data pixels are acquired at each Nyquist ratio: 300 for training, 300 for validation, and 300 for test data. The acquired output images are cropped to prepare the input image for deep learning, whose pixel size is changed according to the number of signal pixels simultaneously output in a single classification process. For example, when one signal pixel is output in a single classification process, the acquired output image is divided into 3×3 data pixels and input into a neural network to predict the original signal phase. Note that the relative phase between the signal and known phase reference pixels represent crucial information in this study, and each image is rotated or flipped so that all input images have the same arrangement of the known phase reference pixel.



Figure 3. Schematic diagram of the optical setup used in the experiment.

2.3. Structure of the Neural Network

This study performs the training using a convolutional neural network that consists of three convolutional layers; two pooling layers; and 1, 2, 4, 8, 16, or 32 parallel fully connected and output layers. Figure 4 shows a network composed of four fully connected layers as an example. The number of output layers corresponds to the number of signal phases that are simultaneously classified, and one output layer corresponds to one signal phase. We note that the outermost signal phases are excluded because they do not have sufficient information for classification. Therefore, the input image size must vary with the number of each output layer, as listed in Table 1. As the number of signal phases to be simultaneously classified increases, the network becomes more complex while the number of input images that are required to recover all signal phases decreases. All convolutional layers use the rectified linear unit (ReLU) function as the activation function. The first convolutional layer has 32 filters, and the remaining two layers have 64 filters. The output layer has a Softmax function set to classify the target signal phase into four values: $0, \pi/2, \pi$, and $3\pi/2$. We also use Dropout [31], which has a 0.5 probability of disabling neurons after each pooling layer, to prevent overfitting of the convolutional neural networks. A categorical cross entropy is used as a loss function, which is minimized using Adam [32] with a batch size of 32.



Figure 4. Network structure with four fully connected layers.

Number of Fully Connected Layers	Input Image Size
1	3×3 data pixels
2	4 imes 4 data pixels
4	4 imes 6 data pixels
8	6×6 data pixels
16	6 imes 10 data pixels
32	10 imes 10 data pixels

Table 1. Relationship between the input image size and number of fully connected layers.

3. Results and Discussions

A pixel error rate (PxER) index is introduced to evaluate the proposed method. PxER is defined as the ratio of the number of pixels of incorrect signal data to the total number of signal-data pixels. Evaluation is performed on 48,600 signal-data pixels in the 300-page data acquired for the test data. The results are shown in Figure 5, which show that the proposed method significantly reduces PxER compared with that in the minimum-finding algorithm in a previous study [19]. In particular, at a Nyquist ratio of 1.4, the PxER drops by three orders of magnitude. Moreover, PxER is almost equivalent independent of the number of pixels to be simultaneously classified, even at 32-pixel simultaneous classification.

If a network is configured using a single output layer to predict multiple signal phases, not the parallel output layers proposed in this paper, the PxER increases with the number of pixels classified simultaneously. The reason can be considered as follows. When the multiple signal phases are classified using a single output layer, the number of classes within the network *C* is expressed as

$$C = Pb^N, (2)$$

where *Pb* is the number of phase levels and *N* is the number of pixels to be predicted simultaneously. Equation (2) implies that the number of classes *C* increases exponentially as the number of pixels *N* increases. Accordingly, the number of parameters of the fully connected layer increases exponentially, and at the same time, the number of training data assigned to each class also decreases exponentially.



Figure 5. Experimental results of the phase determination with deep learning. "Conventional" in the figure represents the minimum-finding algorithm mentioned in Section 2.1.

In contrast, in our proposed network with parallel output layers, since the number of classes classified in each output layer is equal to the number of phase levels *Pb*, the total number of classes *C* in all output layers can be expressed as

$$C = N \times Pb. \tag{3}$$

Therefore, our proposed network can suppress the increase in the number of classes *C* within the network compared to the network with a single output layer. As a result, even if the number of simultaneous signal pixels *N* is increased, a sufficient number of training data can be maintained for each class, and over-fitting due to a large number of parameters can be prevented.

Using another network structure, U-net [33], it is possible to predict all signal phases within a page efficiently. However, even U-net still requires a considerable amount of training data. It has been reported that an error rate of 0.1% was achieved with 9000 training data for 128×128 signal data pixels [28]. In our method, on the contrary, the required training data do not depend on the input image size but on the number of parallel output signals. Therefore, even if a large number of signal data pixels are included within a page to realize a large data transfer rate, our method can moderately suppress the required training data.

Next, we evaluate the computation time to output the signal phase from the input image. The neural network requires a substantial computational cost to output the resultant category. Therefore, deep learning requires a longer phase-evaluation time than the deterministic methods such as the minimum-finding algorithm. Verification is performed 10 times each by detecting 300 pages of test data for seven Nyquist ratios and by measuring the time required for the phase output. Figure 6 shows the results of the phase-output time. The phase-output time is normalized by the output time of the minimum-finding algorithm, which is represented by the red line in Figure 6.

The result shows that the phase-output time decreases with the increase in the number of simultaneous output signal pixels. Whereas the phase-output time for the single-pixel classification is approximately 18 times longer than the conventional minimum-finding algorithm, it is less than 1.7 times when 32 pixels are simultaneously classified. Therefore, the phase-output time can be significantly reduced by increasing the number of pixels to be simultaneously classified. The phase-output time can be further improved by increasing the number of pixels to be simultaneously classified. In fact, when the number of simultaneous output signal pixels is increased to 64 pixels, the normalized output time decreases to 1.3. However, in this case, PxER increases to some extent. For example, at RNyq = 1.4, $Log_{10}PxER$ increases to -3.4 at 64 pixels. This is due to the reduction in the number of training data extracted from one detected image, but it is not a critical defect of our method. If the number of training data can be made large enough, PxER can maintain a small value even at 64 pixels. Therefore, we confirm that PxER in our proposed method can be significantly reduced compared with that in the conventional minimum-finding algorithm. Simultaneously, the phase-output time can be suppressed to almost the same level by employing parallel fully connected layers.



Figure 6. Phase-output time using a network with parallel fully connected layers. The horizontal axis represents the number of signal pixels that are simultaneously output. The vertical axis is the phase-output time normalized by the output time of the minimum-finding algorithm. The dashed line represents the fitted curve.

4. Conclusions

In conclusion, we have investigated and evaluated a single-shot phase-detection method using deep learning. We experimentally confirmed that PxER could be significantly reduced compared with that in the previous studies by learning the light-intensity distribution around the target signal pixel. In addition, the phase-output time could be shortened by devising a network and arranging the fully connected layers in parallel. In our environment, the phase-output time of the single-pixel classification was approximately 18 times longer than that in the previous method; however, it could be reduced to 1.7 times or less when 32 pixels were simultaneously classified. Therefore, we concluded that the proposed method can significantly reduce PxER and suppress the phase-output time to almost the same level as that in the previous method. Usually, errors that occur in signal detection can be entirely corrected using an error-correction code if they are within the limits allowed by the system. Therefore, a reduced PxER can improve the recording density using tuning parameters such as aperture size and multiple recording-angle spacing. This result indicates that the proposed method can increase the recording density while maintaining the transfer rate. Thus, this method can be a promising phase-detection method to realize a large-density data-storage system in the future.

Author Contributions: Conceptualization, M.T. and R.F.; methodology, M.T. and R.F.; software, M.T.; validation, M.T. and R.F.; formal analysis, M.T.; investigation, M.T.; writing—original draft preparation, M.T. and R.F.; writing—review and editing, R.F.; visualization, M.T. and R.F.; supervision, R.F.; project administration, R.F.; funding acquisition, R.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by JSPS KAKENHI, grant number 21K04917.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. van Heerden, P.J. Theory of optical information storage in solids. Appl. Opt. 1963, 2, 393–400. [CrossRef]
- 2. Gabor, D. A new microscopic principle. *Nature* **1948**, *161*, 777. [CrossRef]
- Orlov, S.S.; Phillips, W.; Bjornson, E.; Takashima, Y.; Sundaram, P.; Hesselink, L.; Okas, R.; Kwan, D.; Snyder, R. High-transfer-rate high-capacity holographic disk data-storage system. *Appl. Opt.* 2004, 43, 4902–4914. [CrossRef]
- 4. Denz, C.; Pauliat, G.; Roosen, G.; Tschudi, T. Volume hologram multiplexing using a deterministic phase encoding method. *Opt. Commun.* **1991**, *85*, 171–176. [CrossRef]
- 5. Rakuljic, G.A.; Leyva, V.; Yariv, A. Optical data storage using orthogonal wavelength multiplexed volume holograms. *Opt. Lett.* **1992**, *17*, 1471. [CrossRef]
- 6. Mok, F.H. Angle-multiplexed storage of 5000 holograms in lithium niobate. Opt. Lett. 1993, 18, 915. [CrossRef]
- Curtis, K.; Pu, A.; Psaltis, D. Method for holographic storage using peristrophic multiplexing. *Opt. Lett.* 1994, 19, 993–994. [CrossRef]
- 8. Psaltis, D.; Levene, M.; Pu, A.; Barbastathis, G.; Curtis, K. Holographic storage using shift multiplexing. *Opt. Lett.* **1995**, *20*, 782–784. [CrossRef]
- 9. Barbastathis, G.; Levene, M.; Psaltis, D. Shift multiplexing with spherical reference waves. *Appl. Opt.* **1996**, *35*, 2403–2417. [CrossRef]
- 10. Barbastathis, G.; Psaltis, D. Shift-multiplexed holographic memory using the two-lambda method. *Opt. Lett.* **1996**, *21*, 432–434. [CrossRef]
- 11. Chuang, E.; Psaltis, D. Storage of 1000 holograms with use of a dual-wavelength method. *Appl. Opt.* **1997**, *36*, 8445–8454. [CrossRef] [PubMed]
- 12. Anderson, K.; Curtis, K. Polytopic multiplexing. Opt. Lett. 2004, 29, 1402–1404. [CrossRef] [PubMed]
- 13. Hoshizawa, T.; Shimada, K.; Fujita, K.; Tada, Y. Practical angular-multiplexing holographic data storage system with 2 terabyte capacity and 1 gigabit transfer rate. *Jpn. J. Appl. Phys.* **2016**, *55*, 09SA06. [CrossRef]
- 14. John, R.; Joseph, J.; Singh, K. Holographic digital data storage using phase-modulated pixels. *Opt. Lasers Eng.* **2005**, *43*, 183–194. [CrossRef]
- 15. Yu, C.; Zhang, S.; Kam, P.Y.; Chen, J. Bit-error rate performance of coherent optical M-ary PSK/QAM using decision-aided maximum likelihood phase estimation. *Opt. Express* **2010**, *18*, 12088–12103. [CrossRef]
- 16. Takabayashi, M.; Okamoto, A.; Tomita, A.; Bunsen, M. Symbol error characteristics of hybrid-modulated holographic data storage by intensity and multi phase modulation. *Jpn. J. Appl. Phys.* **2011**, *50*, 09ME05. [CrossRef]
- 17. Nakamura, Y.; Fujimura, R. Wavelength diversity detection for phase-modulation holographic data storage system. *Jpn. J. Appl. Phys.* **2020**, *59*, 012004. [CrossRef]
- Schwider, J.; Burow, R.; Elssner, K.E.; Grzanna, J.; Spolaczyk, R.; Merkel, K. Digital wave-front measuring interferometry: Some systematic error sources. *Appl. Opt.* 1983, 22, 3421. [CrossRef]
- 19. Tokoro, M.; Fujimura, R. Single-shot detection of four-level phase modulated signals using inter-pixel crosstalk for holographic data storage. *Jpn. J. Appl. Phys.* 2021, *60*, 022004. [CrossRef]
- 20. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444. [CrossRef]
- Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* 2017, 29, 2352–2449. [CrossRef] [PubMed]
- Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. In Proceedings of the 27th International Conference on Neural Information Processing Systems 2, Montreal, QC, Canada, 8–13 December 2014; MIT Press: Cambridge, MA, USA, 2014; pp. 3104–3112.
- Chen, H.; Engkvist, O.; Wang, Y.; Olivecrona, M.; Blaschke, T. The rise of deep learning in drug discovery. *Drug Discov. Today* 2018, 23, 1241–1250. [CrossRef] [PubMed]
- 24. Hao, J.; Lin, X.; Lin, Y.; Chen, M.; Chen, R.; Situ, G.; Horimai, H.; Tan, X. Lensless complex amplitude demodulation based on deep learning in holographic data storage. *OEA* 2023, *6*, 220157. [CrossRef]
- 25. Hao, J.; Lin, X.; Chen, R.; Lin, Y.; Liu, H.; Song, H.; Lin, D.; Tan, X. Phase retrieval combined with the deep learning denoising method in holographic data storage. *Opt. Contin.* **2022**, *1*, 51. [CrossRef]
- Shimobaba, T.; Kuwata, N.; Homma, M.; Takahashi, T.; Nagahama, Y.; Sano, M.; Hasegawa, S.; Hirayama, R.; Kakue, T.; Shiraki, A.; et al. Convolutional neural network-based data page classification for holographic memory. *Appl. Opt.* 2017, *56*, 7327–7330. [CrossRef]

- 27. Katano, Y.; Muroi, T.; Kinoshita, N.; Ishii, N.; Hayashi, N. Data demodulation using convolutional neural networks for holographic data storage. *Jpn. J. Appl. Phys.* 2018, *57*, 09SC01. [CrossRef]
- Hao, J.; Lin, X.; Lin, Y.; Song, H.; Chen, R.; Chen, M.; Wang, K.; Tan, X. Lensless phase retrieval based on deep learning used in holographic data storage. *Opt. Lett.* 2021, 46, 4168–4171. [CrossRef]
- 29. Lee, S.H.; Lim, S.Y.; Kim, N.; Park, N.C.; Yang, H.; Park, K.S.; Park, Y.P. Increasing the storage density of a page-based holographic data storage system by image upscaling using the PSF of the Nyquist aperture. *Opt. Express* **2011**, *19*, 12053–12065. [CrossRef]
- Lin, X.; Hao, J.; Wang, K.; Zhang, Y.; Li, H.; Tan, X. Frequency expanded non-interferometric phase retrieval for holographic data storage. Opt. Express 2020, 28, 511–518. [CrossRef]
- 31. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 2014, *15*, 1929–1958.
- Kingma, D.P.; Ba, L.J. Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
- Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Munich, Germany, 5–9 October 2015; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.