Optimized Binary Patterns by Gradient Descent for Ghost Imaging

IKUO HOSHI, TOMOYOSHI SHIMOBABA, TAKASHI KAKUE, (Member, IEEE), AND TOMOYOSHI ITO
Graduate School of Engineering, Chiba University, Chiba 263-8522, Japan
Corresponding author: Ikuo Hoshi (aeka2345@chiba-u.jp)
This work was supported in part by the Yazaki Memorial Foundation for Science and Technology.

ABSTRACT Ghost imaging reconstructs images using a single-element photodetector; it performs imaging by illuminating an object with binary modulation patterns. This technique has various advantages, including a wide wavelength, noise robustness, and high measurement sensitivity. However, one challenge is the low image quality in the undersampling. The examination of modulation patterns is intended to solve this issue. In ghost imaging, randomly generated or basis patterns have been studied as modulation patterns; however, in undersampling, random patterns exhibit noise robustness but low image quality, whereas basis patterns exhibit high image quality but are sensitive to noise and low resolution. Thus, ghost imaging requires patterns that simultaneously achieve high image quality, resolution, and noise robustness. This study proposes a method of pattern optimization using gradient descent and a binarization method to further improve image quality. Numerical simulation and experimental results show that the proposed approach offers high image quality, high resolution, and robustness to noise.

INDEX TERMS Ghost imaging, gradient descent, optimization, single-pixel imaging.

I. INTRODUCTION

Conventional imaging records an image using a sensor composed of several light detectors, such as charge-coupled devices. However, ghost imaging [1], [2] is a unique technique that uses only a single-element photodetector. Figure 1 shows a schematic of ghost imaging, first, random binary patterns on a digital micromirror device (DMD), which is high speed binary display, modulate a target object. Then, the modulated light from the target object is collected by a lens, and its intensity is recorded with a single-element photodetector. This intensity is obtained by sequentially modulating the patterns; subsequently, we reconstruct the target object’s image by performing a correlation calculation between the measured data and the known modulation patterns. With the features obtained using a single-element photodetector, ghost imaging offers several advantages, such as the ability to operate in a low-light environment, as well as high-speed and broad-wavelength imaging. Hence, ghost imaging can be applied to various fields, such as remote sensing [3], object tracking [4], high-speed cytometry [5], and terahertz imaging [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Md. Abdur Razzaque.
In this study, we propose a two-step optimization of modulation patterns to improve image quality by offering high resolution and robustness against environmental noise. We apply a sophisticated optimization technique on the basis of deep learning to better optimize modulation patterns. In recent years, deep learning has achieved great successes in image noise reduction [16] and super-resolution [17]. The most common examples of applying deep neural networks to ghost imaging in previous studies are both refining the image reconstructed by correlation calculation with deep neural network [18], [19], and directly reconstructing the image with networks without using correlation calculations [20]. However, these methods have some disadvantages, such as additional and high computational load and decreased noise robustness which is a feature of ghost imaging.

In this study, we performed the two-step pattern optimization using Keras framework with Tensorflow backend as an optimization solver, not as a deep learning framework. We implemented correlation calculation into the framework to obtain optimized patterns. Thus, similar to conventional ghost imaging, the proposed method can obtain reconstruction images with better image quality to use correlation calculations in the reconstruction process without using deep learning network and noise reduction methods [21]. Optimization is performed in two steps: first, optimized grayscale patterns are acquired; then, the optimized grayscale patterns are binarized with as little negative effect as possible. Our concept follows that of a previous study [22], whereby the learning of modulation patterns is incorporated into a deep neural network by considering them as network parameters. Unlike spatial frequency-based patterns [7], [8], the proposed modulation patterns improve image quality while retaining robustness against noise and high resolution, even for undersampling. We can therefore achieve ghost imaging with an improved image quality while retaining high resolution and noise robustness using the proposed method.

The remainder of this paper is organized as follows. First, we describe the proposed method; then, we present the results of a simulation and an optical experiment. In simulation, we compare noiseless and noisy cases. Furthermore, we compare the proposed binarization method with simple binarization. Finally, we present the conclusions.

II. PROPOSED METHOD

The starting point of the proposed method is differential ghost imaging (DGI) [23], which is expressed as

\[ O(x, y) = \langle S_i (x, y) \rangle - \frac{\langle S_i \rangle}{\langle R_i \rangle} \langle R_i I_i(x, y) \rangle, \]

where \( I_i(x, y) \) is the modulation pattern with gradation, the subscript \( i = \{1, 2, \ldots, N\} \) indicates the index of the modulation patterns, \( N \) is the number of patterns, \( \langle \rangle \) represents the ensemble average for all the measurements, and \( O(x, y) \) is the reconstructed image. \( S_i \) and \( R_i \) represent the \( i \)-th total light intensities of the object and the patterns, as expressed by

\[ S_i = \sum_{y=1}^{Y} \sum_{x=1}^{X} T(x, y)I_i(x, y), \]
\[ R_i = \sum_{y=1}^{Y} \sum_{x=1}^{X} I_i(x, y), \]

where \( X \) and \( Y \) represent the width and height of an image and \( T(x, y) \) represents the target object. The above calculations are differentiable, hence, to optimize the modulation patterns via gradient descent, we obtain gradients using the error between the reconstructed image and the original target object. Afterward, we update the modulation patterns according to the gradients. By repeating these updates, we can obtain the optimized patterns. The optimized patterns have nonbinary values; thus, our proposed method only performs well using spatial light modulators such as liquid crystal displays, which can display grayscale patterns with low refresh rates. Although the use of binary spatial light modulators such as DMDs can project binary patterns at high speed, simple binarization of the optimized grayscale patterns may degrade the quality of the reconstructed images. Hence, we need pattern optimization based on our binarization assumption. We propose a two-step optimization scheme with scaling factors. Let us begin by explaining the scaling factor.

A. SCALING FACTOR

The scaling factor is used to improve the accuracy of binary neural networks [24]. One binary illumination pattern has a scaling factor \( \alpha_i \), which is a scalar value. By approximating \( I_i(x, y) \) as \( I(x, y) \approx \alpha_i B_i(x, y) \) (where \( B_i(x, y) \) is the desired binary pattern), we can express \( S_i \) and \( R_i \) as

\[ S_i = \sum_{y=1}^{Y} \sum_{x=1}^{X} T(x, y)I_i(x, y) \]
\[ \approx \alpha_i \sum_{y=1}^{Y} \sum_{x=1}^{X} T(x, y)B_i(x, y), \]
\[ R_i = \sum_{y=1}^{Y} \sum_{x=1}^{X} I_i(x, y) \approx \alpha_i \sum_{y=1}^{Y} \sum_{x=1}^{X} B_i(x, y). \]

If optimal scaling factors \( \alpha_i \) can be obtained, it will reduce the binarization error of the patterns. The scaling factors are decided to minimize the differences between the binary patterns with \( \alpha_i \) and the grayscale patterns; thus, we optimize \( \alpha_i \) as

\[ \alpha_i = \arg \min_{\alpha_i} ||I_i - \alpha_i B_i||^2 \]
\[ = \arg \min_{\alpha_i} \alpha_i^2 ||B_i||^2 - 2\alpha_i I_i^T B_i + ||I_i||^2. \]
where the grayscale and binary patterns are represented as $I_i$ and $B_i$, which are $M \times N$ vectors, with each vector element denoted by $I_{i,k}$ or $B_{i,k}$ with the index $k = \{1, 2, \ldots, X \times Y\}$, respectively. To select $B_i$ such that Eq. (6) is minimized, we use $B_{i,k} = 1$ if $I_{i,k} \geq 0.5$ or $B_{i,k} = 0$ otherwise, because an actual optical system cannot represent negative values. In this study, we initialize $I_{i,k}$ ranging from 0 to 1. Since the optimal scaling factors $\alpha_i$ in Eq. (6) can be simply found by $\partial \parallel I_i - \alpha_i B_i \parallel^2 \parallel \alpha_i \parallel = 0$, the scaling factors $\alpha_i$ are expressed as

$$\alpha_i = \frac{I_i^T B_i}{\|B_i\|^2}. \quad (7)$$

Finally, the proposed reconstruction calculation with a scaling factor $\alpha_i$ is expressed as

$$O(x, y) = \langle S, \alpha_i B_i(x, y) \rangle - \langle S, (R_i \alpha_i B_i(x, y)) \rangle. \quad (8)$$

**B. SUMMARY OF THE PROPOSED METHOD**

Algorithm 1 shows the overall flow of the proposed method combining two-step optimization with the scaling factor. In the first step, the optimized grayscale patterns $I_i(x, y)$ with floating point precision are obtained via stochastic gradient descent (SGD) with the state-of-the-art optimizer, Adam [25]. Adam (adaptive moment estimation) is one of the most common optimization approaches in neural networks. It is one of the derivatives of gradient descents, which allows for more efficient optimization. The feature of Adam is that it can apply an adaptive learning rate to each parameter by using the previous gradient. We used the default parameters of the Adam optimizer. The gradients are obtained by performing the DGI calculation of Eq. (8). The normalize function in Algorithm 1 normalizes the reconstruction results ranging from 0 to 1, which is expressed as

$$O_{\text{new}}(x, y) = \frac{O(x, y) - O_{\text{min}}(x, y)}{O_{\text{max}}(x, y) - O_{\text{min}}(x, y)}. \quad (9)$$

where $O_{\text{max}}(x, y)$ and $O_{\text{min}}(x, y)$ are the maximum and minimum values of $O(x, y)$, respectively. SGD proceeds with optimization using mini-batches that divide entire datasets into subsets. In the second step, we obtain binary patterns optimized from grayscale patterns using SGD with Adam, as well as a binarization method with scaling factors $\alpha_i$ to reduce binarization’s negative effects. One set of modulation patterns and scaling factors is obtained per one training dataset.

**III. SIMULATION AND EXPERIMENT**

We conduct simulations and optimal experiments to verify the effectiveness of the proposed method. Using Python 3.5.4, Tensorflow 1.8.0, and Keras 2.2.4 as our software environment, we applied STL-10 [26] as a training dataset. The number of training data is 100,000; the numbers of validation and test data are both 4,000. The test data are used to evaluate numerical simulations. The original and reconstructed images in the figures are chosen from the test data.

**Algorithm 1 Proposed Two-Step Optimization. Blue Texts Denote Comments**

**Input:** Random modulation patterns $I = I_{i(i=1,2,\ldots,N)}$, a minibatch of ground truth images $T$  

**Output:** Optimized binary patterns $B = B_{i(i=1,2,\ldots,N)}$, scaling factor $\alpha = \alpha_{i(i=1,2,\ldots,N)}$.

**First step optimization:**

1. for Number of epoch do
2. \( O \leftarrow \text{DGI}(T, I) \)
   \hspace{1cm} // Equation (1) is used
3. \( O \leftarrow \text{Normalize}(O) \)
   \hspace{1cm} // Equation (9) is used
4. \( I_{\text{new}} \leftarrow \text{Optimizer}(T, O) \)
   \hspace{1cm} // Adam [25] is used as optimizer
5. \( I \leftarrow I_{\text{new}} \)
6. end for

**Second step optimization:**

8. \( B \leftarrow \text{Binarization}(I) \)
   \hspace{1cm} // $B_{i,k} = 1$ if $I_{i,k} \geq 0.5$ else $B_{i,k} = 0$
9. \( \alpha_i \leftarrow \frac{I_i^T B_i}{\|B_i\|^2} \)
   \hspace{1cm} // Equation (6)
10. for Number of epoch do
11. \( O \leftarrow \text{Reconstruction}(T, B, \alpha) \)
    \hspace{1cm} // Equation (8) is used
12. \( O \leftarrow \text{Normalize}(O) \)
13. \( E \leftarrow \text{ErrorFunction}(T, O) \)
14. \( B_{\text{new}} \leftarrow \text{Optimizer}(B, \partial E / \partial B) \)
    \hspace{1cm} // Adam [25] is used as optimizer
15. \( I_{\text{new}} \leftarrow I + (B_{\text{new}} - B) \)
16. \( B \leftarrow \text{Binarization}(I) \)
17. \( \alpha_i \leftarrow \frac{I_i^T B_i}{\|B_i\|^2} \)
18. end for

**A. EVALUATION OF NOISELESS ENVIRONMENT**

Table 1 and Fig. 2 show the reconstructed images using each modulation pattern. Table 1 presents the evaluated image qualities for the reconstructions.

Figure 2 shows the reconstructed images when the number of measurements is 1,024. Random modulation patterns are generated with random numbers; for the basis pattern, we used a state-of-the-art pattern called Origami, which was proposed in [8]. These patterns are binary. The insets in Fig. 2 show the enlarged parts of the reconstructed images. It can be seen that the proposed method has smooth tones and high resolution.

Table 1 shows the average peak-signal-to-noise ratios (PSNRs) for the 4,000 test data. The percentage indicates the ratio of the number of modulation patterns to the size (128 × 128[pix]) of the original images. From Table 1, the proposed method resulted in an improved image quality compared with that via a random pattern. The image quality...
obtained via the proposed method was confirmed to be lower than that via the basis patterns without noise. From Fig. 2, we can see that the resolution of the proposed method is higher than that of the basis pattern. Since noise affects measurements in a real environment, our next step is to investigate the noise tolerances of these methods.

### B. EVALUATION OF NOISY ENVIRONMENT

Figures 3 and 4 compare between the basis and proposed patterns in a noisy environment. We added white Gaussian noise to modulation patterns because we assume time-varying noise such as smoke, rain, and fog as a noise environment. The addition of noise intensity is expressed via an SNR, which is defined as the ratio of the mean squared total values of the modulation pattern and added noise, as expressed by $10 \log_{10}(S_{\text{pattern}}/N_{\text{noise}})$, where $S_{\text{pattern}}$ and $N_{\text{noise}}$ are the mean squared total values. The noise increases as the SNR decreases; a SNR of 10 dB is 1/10 of the intensity of the modulation pattern, and a SNR of 0 dB represents the same noise as the intensity of the modulation pattern. The number of patterns is 1,024. Figure 3 shows the reconstructed image and its quality when white Gaussian noise is applied in 2 dB steps from 10 to $-10$ dB. Figure 4 shows the image quality evaluation that the noise intensity to the image quality of reconstruction in the basis and proposed patterns. The evaluation presents the average PSNR of 4,000 test data when white Gaussian noise is applied from 10 to $-10$ dB.

Figure 3 shows the deterioration of the image quality for the basis patterns. However, the proposed patterns maintain image quality, even with the increase in noise intensity.

From Fig. 4, we can see that there is already no significant difference in image quality of approximately 10 dB, and that the image qualities of the basis and proposed patterns reverse at 8.64 dB. On the basis of these results, we find that the proposed method has higher noise robustness than the basis patterns.

Because the basis pattern is a method of projecting a spatial frequency pattern, reconstructed images with sufficient resolution cannot be obtained with few measurements. In addition, when noise is applied to the spatial frequency-based basis pattern, the noise is strongly reflected in the reconstructed images. On the other hand, the proposed method is based on random patterns, so it has high noise immunity and high resolution.

### C. EVALUATION OF PROPOSED BINARIZATION

To confirm the effectiveness of the proposed binarization, we compared it with a simple binarization scheme using a threshold value. In the simple thresholding method, the grayscale patterns were binarized according to $B_{i,k} = 1$ if $I_{i,k} \geq 0.5$ or $B_{i,k} = 0$ if otherwise, where $I_{i,k}$ is obtained from the first optimization step in Algorithm 1.
Figure 5 and Table 2 shows the results. The evaluation is the average PSNR for the 4,000 test data in Table 2. These results show that the proposed binarization is superior to the simple thresholding binarization in the subjective and quantitative evaluations of all measurement times.

**Table 2.** Comparison of the binarization methods. The PSNRs are averages in the test data.

| The number of measurements | Threshold | Proposed |
|----------------------------|----------|----------|
| 256 (1.6%)                 | 13.47[dB]| 16.7[dB] |
| 512 (3.1%)                 | 13.62[dB]| 17.69[dB]|
| 1024 (6.3%)                | 14.22[dB]| 18.73[dB]|

**D. OPTICAL EXPERIMENT**

Finally, we conducted an experiment using an actual optical system. Figures 6 and 7 show the optical system and reconstructed images with and without noise, respectively. The insets in Fig. 7 show the enlarged parts of the reconstructed images. As sources of noise, we used smoke from incense sticks placed between a target object and the DMD. As a result of noise measurement, the noise intensity was ranging from $-5$ to $-10$ dB approximately. The PSNRs in Fig. 7 were measured by comparing a reconstructed image in a noisy environment with one in a noiseless environment. The number of modulation patterns is 1,024. In a noiseless environment, the proposed patterns improve the contrast of the reconstructed images compared with that of the random patterns. Although the proposed patterns exhibit less contrast with reconstructed images than do the basis patterns,
We found that the proposed method has higher image quality against noise and high resolution, even for undersampling.

**IV. CONCLUSION**

In conclusion, we have proposed binary modulation patterns with scaling factors optimized using SGD for ghost imaging. We found that the proposed method has higher image quality than that obtained using random patterns. We also showed that the proposed method has a higher resolution and noise robustness than those obtained using basis patterns. From these results, we could conclude that the proposed method could improve image quality while retaining a high resolution and noise robustness. In this study, we used a universal dataset to optimize binary patterns. However, in practice, we will obtain better results if we use an application-specific dataset. In future work, it will not be easy to increase the number of pixels in the current Keras and Tensorflow-based framework because of the lack of memory in a graphics processing unit (GPU). Hence, we intend to increase the number of pixels via full-scratch GPU programming.

**REFERENCES**

[1] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, “Optical imaging by means of two-photon quantum entanglement,” *Phys. Rev. A*, Gen. Phys., vol. 52, no. 5, pp. R3429–R3432, Nov. 1995.

[2] J. H. Shapiro, “Computational ghost imaging,” *Phys. Rev. A*, Gen. Phys., vol. 78, no. 6, Dec. 2008, Art. no. 061802.

[3] B. I. Erkmen, “Computational ghost imaging for remote sensing,” *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 29, no. 5, pp. 782–789, 2012.

[4] D. Shi, K. Yin, J. Huang, K. Yuan, W. Zhu, C. Xie, D. Liu, and Y. Wang, “Fast tracking of moving objects using single-pixel imaging,” *Opt. Commun.*, vol. 440, pp. 155–162, Jun. 2019.

[5] Z. Wen, J. Guan, T. Zeng, and Y. Li, “Hyperspectral terahertz microscopy via nonlinear ghost imaging,” *Optica*, vol. 7, no. 2, pp. 1816–191, 2020.

[6] I. Hoshi, T. Shimobaba, T. Kakue, and T. Ito, “Single-pixel imaging using a universal dataset,” *IEEE Trans. Opt. Sci.*, vol. 19, no. 23, p. 5135, Nov. 2019.

[7] S. Fekri-Ershad, S. Fakhrahmad, and F. Tajeripour, “Impulse noise reduction for texture images using real word spelling correction algorithm and local binary patterns,” *Int. Arab. J. Antimicrob. Agents*, vol. 15, p. 1024, Nov. 2018.

[8] C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, “Deep learning for real-time single-pixel video,” *Sci. Rep.*, vol. 8, no. 1, p. 2369, Feb. 2018.
[23] F. Ferri, D. Magatti, L. A. Lugiato, and A. Gatti, “Differential ghost imaging,” *Phys. Rev. Lett.*, vol. 104, no. 25, Jun. 2010, Art. no. 253603.

[24] M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi, “XNOR-Net: ImageNet classification using binary convolutional neural networks,” in *Proc. Eur. Conf. Comput. Vis.* (ECCV). New York, NY, USA: Springer, 2016, pp. 525–542.

[25] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2014, arXiv:1412.6980. [Online]. Available: https://arxiv.org/abs/1412.6980

[26] A. Coates, A. Ng, and H. Lee, “An analysis of single-layer networks in unsupervised feature learning,” in *Proc. 14th Int. Conf. Artif. Intell. Statist.*, 2011, pp. 215–223.

IKUO HOSHI received the B.E. and M.E. degrees in electrical and electronic engineering from Chiba University, Japan, in 2018 and 2020, respectively. His research interests include signal processing, deep learning, high-performance computing, and single-pixel imaging.

TAKASHI KAKUE (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electronics and information science from the Kyoto Institute of Technology, Japan, in 2006, 2008, and 2012, respectively. Since 2012, he has been an Assistant Professor with the Graduate School of Engineering, Chiba University, Japan. His research interests include holography, digital holography, computer holography, holographic interferometry, 3D imaging, high-speed imaging, and ultrafast optics. He is currently a member of OSA, OSJ, and SPIE.

TOMOYOSHI SHIMOBABA received the B.E. and M.E. degrees in electrical engineering from Gunma University, Japan, in 1997 and 1999, respectively, and the Ph.D. degree from Chiba University, Japan, in 2002. From 2002 to 2005, he was a Special Postdoctoral Researcher with RIKEN. From 2005 to 2009, he was an Associate Professor with the Graduate School of Science and Engineering, Yamagata University. From 2009 to 2019, he was an Associate Professor with the Graduate School of Engineering, Chiba University, where he has been a Professor, since 2019. His research interests include computer holography and its applications. He is currently a member of ITE, IEICE, OSA, OSJ, and SPIE.

TOMOYOSHI ITO received the B.S. degree in pure and applied sciences and the M.S. and Ph.D. degrees in earth science and astronomy from The University of Tokyo, Japan, in 1989, 1991, and 1994, respectively. From 1992 to 1994, he was a Research Associate with Gunma University, Japan, where he was an Associate Professor, from 1994 to 1999. From 1999 to 2004, he was an Associate Professor with the Graduate School of Engineering, Chiba University, Japan, where he has been a Professor, since 2004. His research interests include high-performance computing and its applications, such as electronic holography for 3D TV. He is currently a member of ACM, ASJ, ITE, IEICE, IPSJ, and OSA.