Incoherent light control through scattering media based on machine learning and its application to multiview stereo displays

Ryoichi Horisaki^{1,2} · Yuki Mori¹ · Jun Tanida¹

Received: 13 March 2019 / Accepted: 8 October 2019 / Published online: 19 October 2019
© The Author(s) 2019

Abstract

In this paper, we present a method for controlling incoherent light through scattering media based on machine learning and its potential application to multiview stereo displays. The inverse function between input and output light intensity patterns through a scattering medium is regressed with a machine learning algorithm. The inverse function is used for calculating an input pattern for generating a target output pattern through a scattering medium. We demonstrate the proposed method by assuming a potential application to multiview stereo displays. This concept enables us to use a diffuser as a parallax barrier, a cylindrical lens array, or a lens array on a conventional multiview stereo display, which will contribute to a low-cost, highly functional display. A neural network is trained with a large number of pairs of displayed random patterns and their parallax images at different observation points, and then a displayed image is calculated from arbitrary parallax images using the trained neural network. In the experimental demonstration, the scattering-based multiview stereo display was composed of a diffuser and a conventional liquid crystal display, and it reproduced different handwritten characters, which were captured by a stereo camera.

Keywords Light control through scattering media · Incoherent light · Machine learning · Neural network · Three-dimensional display · Multiview stereo display

1 Introduction

Light control through scattering media is a long-standing issue in the fields of optics and photonics for biomedical imaging, optical communication, and so on [1, 2]. Various methods have been proposed with this aim, and they are mainly categorized into feedback-based and inversion-based approaches. The first approach uses an iterative feedback process for optimizing a light pattern behind or inside scattering media [3, 4]. The second one takes the inverse of the transmission matrix for non-iterative wavefront shaping through scattering media [5, 6].

A machine learning technique has recently been introduced to the inversion-based approach for light control

through scattering media to simplify the optical setup and extend the range of its applications [7, 8]. In this method, the relationship between the input and output of light through scattering media is regressed with a large number of training input and output pairs. In the previous studies, a coherent light source and a spatial light modulator were used, but these devices were costly. Here we show that the machine learning approach can be applied to control incoherent light through scattering media. This reduces the hardware cost and extends the impact of light control through scattering media to various fields. As an example of its application, we demonstrate a multiview stereo display (MSD) with scattering media based on the machine learning approach described in this study.

An MSD is one category of three-dimensional displays, which presents a three-dimensional scene to viewers. MSDs are promising for glasses-free viewing and commercial products [9, 10]. This display technique employs an optical modulation element on a two-dimensional display or a stack of multiple displays for angle control of light rays toward observers [10–12]. In this study, we took the first approach because it uses only a single display and is cost-effective

✉ Ryoichi Horisaki
r.horisaki@ist.osaka-u.ac.jp

¹ Department of Information and Physical Sciences, Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

² JST, PRESTO, 4-1-8 Honcho, Kawaguchi-shi, Saitama 332-0012, Japan

compared to the second one. Parallax barriers, cylindrical lens arrays, and microlens arrays are typically used as the optical modulation elements in MSDs based on the first approach. An image on the display using this approach is geometrically calculated to produce arbitrary parallax images based on a pinhole array model. Misalignments and/or aberrations (especially when the viewing angle is large) of the optical modulation elements affect the reproduced image quality due to the mismatch between the model and the actual optical system [13, 14].

In our experimental demonstration, a diffuser was used as a novel optical modulation element of the MSD to show the general versatility of this approach, although it is readily applicable to conventional MSDs for compensating the model mismatch based on machine learning. Our display reproduced different images at just two viewing positions for simplicity, but it is straightforward to extend the system to MSDs with more viewing positions. Our study relaxes the requirements for optical control through scattering media and its applications. Also, the demonstrated display technique may contribute to cost reduction of the optical modulation element by, for example, allowing use of a low-quality lens array and diffuser, as well as improved performance of MSDs.

2 Methods

A schematic diagram of the proposed method is shown in Fig. 1. A diffuser is attached to a liquid crystal display to serve as the optical modulation element. The scattered light field is captured as a parallax pair by a stereo camera with two viewing positions. In the training stage, randomly generated input patterns are displayed on the display, and their output images are captured by the camera through the diffuser. The inverse function from the output to the input is regressed with an artificial neural network. The optical process is linear because incoherent light is emitted from the liquid crystal display. Thus, we employ a simple perceptron without any hidden layers or activation layers [15]. After the training stage, a target image is provided to the network, and an input pattern is calculated. The calculated input pattern is displayed on the display, and the target image is reproduced on the camera through the diffuser.

3 Experimental demonstration

The proposed method was experimentally demonstrated. A diffuser composed of five acrylic ground plates (Template acry manufactured by AcrySunday, thickness: 3 mm, color number: 811, diffusion angle of the five stacked plates: 26°) was directly attached to the liquid crystal display (XSP-04

Fig. 1 Schematic diagram illustrating the proposed method

manufactured by Luckyster, pixel count: 1920×1080 , screen size: 13.3 in.). The stereo camera was implemented by a camera lens (Ai Nikkor 85 mm F2 manufactured by Nikon), a monochrome image sensor (PL-B953 manufactured by PixeLink, pixel count: 768×1024 , pixel pitch: $4.65 \mu\text{m}$), and an iris stop (diameter: 1 cm) located in front of the camera lens with a motorized stage (OSMS20-85 manufactured by OptoSigma). The exposure time was set to 700 ms. The left view of the parallax pair was captured by moving the iris stop to the left edge of the camera lens, and the right view was captured by moving the iris stop to the right edge. The distance between the diffuser and the camera lens was 54.9 cm. The distance between the left and right viewing positions was 3.2 cm. This distance may be controllable by choosing the scattering angle of the diffuser.

Examples of the random input patterns and their parallax pairs are shown in Fig. 2. The pixel count of the images was 64×64 . The image size was $1.5 \text{ cm} \times 1.5 \text{ cm}$ on the surface of the diffuser. The simple perceptron was trained with 100,000 random input patterns and parallax pairs. The perceptron was composed of two layers (input and output

View	L	R	L	R	L	R
Input pattern
Parallax pairs

Fig. 2 Examples of the random input patterns and their parallax pairs

layers) without any hidden or activation layer. The Adam optimizer was used for optimizing the network with an initial learning ratio of 0.001, a batch size of 32, and a number of epochs of 20 [16].

First, incoherent light control through the diffuser was performed with only the right viewing position as an individual demonstration. In this case, the numbers of nodes in the input and output layers of the neural network were 64×64 and 64×64 , respectively. The inputs to the network in the training stage were the captured images in the right viewing position. The network calculated an input pattern from the target image. Then, the calculated input pattern was displayed on the display, and the image was reproduced through the diffuser. The reproduced image was captured with the right viewing position of the stereo camera. The experimental result is shown in Fig. 3, where the target images were handwritten numbers randomly selected from the MNIST database [17]. The target images were reproduced through the diffuser. The reproduced images were evaluated using structural similarity (SSIM) [18]. The average SSIM of five hundred reproduced images was 0.32.

Next, reproduction of the parallax pairs through the diffuser was demonstrated. In this case, images on both the right and left viewing points were provided to the network, where the numbers of nodes in the input and output layers were $64 \times 64 \times 2$ and 64×64 , respectively. The target parallax pair consisted of two different handwritten numbers randomly selected from the MNIST database. The network calculated an input pattern from the target parallax pair. Then, the calculated input pattern was displayed on the

View	R	R	R
Target image
Calculated input pattern
Reproduced image

Fig. 3 Experimental results of image transmission through scattering media with the right viewing position

display, and the parallax pair was reproduced through the diffuser. The reproduced parallax pair was captured by the stereo camera. Examples of the target parallax pairs, the calculated input patterns, the reproduced parallax pairs, and the reproduced images at the center viewing position are shown in Fig. 4. The handwritten numbers were reproduced at the left and right viewing positions. The average SSIM of five hundred parallax pairs was 0.15. The crosstalk between the left and right viewing positions is shown in the reproduced images at the center viewing position.

4 Conclusion

We presented an extension of light control through scattering media based on machine learning to the incoherent case. The relationship between an image on the liquid crystal display and that captured by the camera was learned using the perceptron with randomly generated training patterns. The image on the display was calculated with the network for reproducing a target image on the camera. We showed a potential application of the concept to MSDs. In this case, the network was trained with parallax pairs captured by a stereo camera. Target images and target parallax pairs were reproduced in the experimental demonstration in both cases of single-view and multiview, respectively.

Our method simplifies the optical setup for light control through scattering media using a conventional display. The demonstrated application for MSDs here may be used for compensating optical processes, such as aberrations of lens arrays, that cannot be expressed with the pinhole array model used in conventional MSDs, and can be also extended to secure information displays [19, 20]. The crosstalk and the viewing zone of the scattering-based MSD should be investigated further in future work.

View	L	R	L	R	L	R
Target parallax pairs
Calculated input pattern
Reproduced parallax pairs
Reproduced image at the center position

Fig. 4 Experimental results of reproduction of parallax pairs through scattering media

Funding This work was supported by JSPS KAKENHI Grant numbers JP17H02799 and JP17K00233, and JST PRESTO Grant number JPMJPR17PB.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Mosk, A.P., Lagendijk, A., Leroosey, G., Fink, M.: Controlling waves in space and time for imaging and focusing in complex media. *Nat. Photonics* **6**, 283–292 (2012)
2. Horstmeyer, R., Ruan, H., Yang, C.: Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue. *Nat. Photonics* **9**, 563–571 (2015)
3. Vellekoop, I.M., Mosk, A.P.: Focusing coherent light through opaque strongly scattering media. *Opt. Lett.* **32**, 2309–2311 (2007)
4. Katz, O., Small, E., Bromberg, Y., Silberberg, Y.: Focusing and compression of ultrashort pulses through scattering media. *Nat. Photonics* **5**, 372–377 (2011)
5. Popoff, S.M., Leroosey, G., Carminati, R., Fink, M., Boccaro, A.C., Gigan, S.: Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. *Phys. Rev. Lett.* **104**, 100601 (2010)
6. Jang, M., Horie, Y., Shibukawa, A., Brake, J., Liu, Y., Kamali, S.M., Arbabi, A., Ruan, H., Faraon, A., Yang, C.: Wavefront shaping with disorder-engineered metasurfaces. *Nature Photonics* **12**, 84–90 (2018)
7. Horisaki, R., Takagi, R., Tanida, J.: Learning-based focusing through scattering media. *Appl. Opt.* **56**, 4358–4362 (2017)
8. Turpin, A., Vishniakou, I., Seelig, J d: Light scattering control in transmission and reflection with neural networks. *Opt. Express* **26**, 30911–30929 (2018)
9. Geng, J.: Three-dimensional display technologies. *Adv. Opt. Photon.* **5**, 456–535 (2013)
10. Stern, A., Javidi, B.: Three-dimensional image sensing, visualization, and processing using integral imaging. *Proc. IEEE* **94**, 591–607 (2006)
11. Loukianitsa, A., Putilin, A.N.: Stereodisplay with neural network image processing. *Proc. SPIE* **4660**, 207–211 (2002)
12. Wetzstein, G., Lanman, D., Hirsch, M., Raskar, R.: Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graphics* **31**, 1–11 (2012)
13. Arai, J., Okui, M., Kobayashi, M., Okano, F.: Geometrical effects of positional errors in integral photography. *J. Opt. Soc. Am. A* **21**, 951–958 (2004)
14. Zhang, W., Sang, X., Gao, X., Yu, X., Yan, B., Yu, C.: Wavefront aberration correction for integral imaging with the pre-filtering function array. *Opt. Express* **26**, 27064–27075 (2018)
15. Rosenblatt, F.: The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **65**, 386–408 (1958)
16. Kingma, D. P., Ba, J.: Adam: a method for stochastic optimization. In: *International Conference on Learning Representations (ICLR)* (2015)
17. Cohen, G., Afshar, S., Tapson, J., van Schaik, A.: EMNIST: an extension of MNIST to handwritten letters, arXiv preprint p. 1702.05373 (2017)
18. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004)
19. Yamamoto, H., Hayasaki, Y., Nishida, N.: Secure information display with limited viewing zone by use of multi-color visual cryptography. *Opt. Express* **12**, 1258–1270 (2004)
20. Yamamoto, H., Hayasaki, Y., Nishida, N.: Secure information display with two limited viewing zones using two decoding masks based on visual secret sharing scheme. *Jpn. J. Appl. Phys.* **44**, 1803–1807 (2005)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.