The following publication Lina Zhou, Yin Xiao, and Wen Chen "Image recovery through turbid water under wide distance ranges", Proc. SPIE 11205, Seventh International Conference on Optical and Photonic Engineering (icOPEN 2019), 112051D (16 October 2019) is available at https://doi.org/10.1117/12.2542212.

Image Recovery Through Turbid Water under Wide Distance Ranges

Lina Zhou,[a,b] Yin Xiao,[a,b] and Wen Chen[a,b,*]
[a]Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hong Kong, China
[b]The Hong Kong Polytechnic University Shenzhen Research Institute,
Shenzhen 518057, China
*Email: owen.chen@polyu.edu.hk

ABSTRACT

Imaging through scattering media is a long-standing problem which has been extensively studied to promote the development of imaging in complex environments. Extant techniques for image reconstruction in scattering media face with the disadvantages of limited ranges of applications, high sensitivity to environmental changes and huge computational load. The scattering media commonly used in practical applications are more complicated due to unknown perturbations. One of the most outstanding problems is the uncertainty of the object position which obstructs progressive development of image recovery techniques. Therefore, it is meaningful to explore a feasible method to bypass additional requirements of precision measuring instruments. Here, we present a method based on convolution neural network (CNN) for optical image reconstruction. The targets are placed in the scattering media which are composed of a certain volume of water and milk, and their diffraction patterns are recorded by using a camera. The learning model demonstrated in this paper is tolerant to uncertainty of object positions. It is foreseeable to be a promising substitute for imaging objects in harsh environments.

Keywords: Imaging through turbid water, underwater imaging, machine learning, image reconstruction.

1. INTRODUCTION

Image recovery in turbid media is one of the hottest research areas ranging from macroscopic imaging to microscopic imaging [1,2]. It is also an important step in extensive applications, such as marine exploration of living creatures. The ubiquitous problem of imaging in scattering media is inhomogeneity of the media which hampers delivery of useful information, and accordingly increases difficulties of image retrieval [3]. The existing methods for imaging in scattering media include enhancement of contrast [4], calculation of transmission matrix [5], and complicated mathematical algorithms [6,7]. However, these methods have some inevitable disadvantages, i.e., narrow applicability, susceptible to small perturbations, and large computational load. Hence, it is necessary to seek for a relatively easy approach to deal with complicated scattering media. In previous studies, a learning model for retrieving useful information from turbid media is proposed [8,9]. However, it is still a huge challenge to process the diffraction patterns recorded at an uncertain position. Usage of precision measuring equipments is a feasible solution to this problem but causes extra requirements of sophisticated instruments.

This paper is to discuss the circumstance that the objects are placed at uncertain distances. Situations are more complicated due to the following reasons: (1) in most applications for imaging, objects are usually fixed in a constant distance to ensure an exact model for image reconstruction. (2) labile factors which greatly increase computational load and model complexity need to be taken into consideration. The existing methods for imaging in strongly scattered media are ineffective. Here, we attempt to implement imaging of objects placed at unknown distance using more advanced methods, i.e., machine learning.

2. EXPERIMENTAL SETUP AND THEORETICAL ANALYSES

2.1 Experimental Setup

The turbid medium used in our experimental setup is a mixture emulsion of pure water and milk. A He-Ne laser beam is first expanded and then collimated to transmit through a water tank illuminating on the spatial light modulator

(SLM) as shown in Fig. 1. The SLM serves as objects. The light reflected from the SLM propagates through a linear polarizer. After that, the reflected light passes through the murky medium again. Then, the strongly scattered light propagates through a 4f system, and the speckle patterns are captured by using a CCD camera. The camera is placed at wide ranges of distances away from the water tank. The turbid water is mixed by wide ranges of volumes of milk. Under a fixed density of the water solution, we will move the image sensor away from focal plane to diffraction plane until nothing can be recognized in the CCD camera. As known, penetration depth of the light strongly depends on the density of the scattering media. Hence, using the higher turbidity of the water solution, there will be the shorter propagation length. When the volume of milk is $40\ ml$, $50\ ml$, $60\ ml$, $70\ ml$ and $80\ ml$, their corresponding farthest distances to recognize the speckle pattern are about $90\ mm$, $80\ mm$, $60\ mm$, $30\ mm$ and $10\ mm$, respectively. In addition, with the larger distance between the camera and the object, active area of the recorded diffraction pattern is larger.

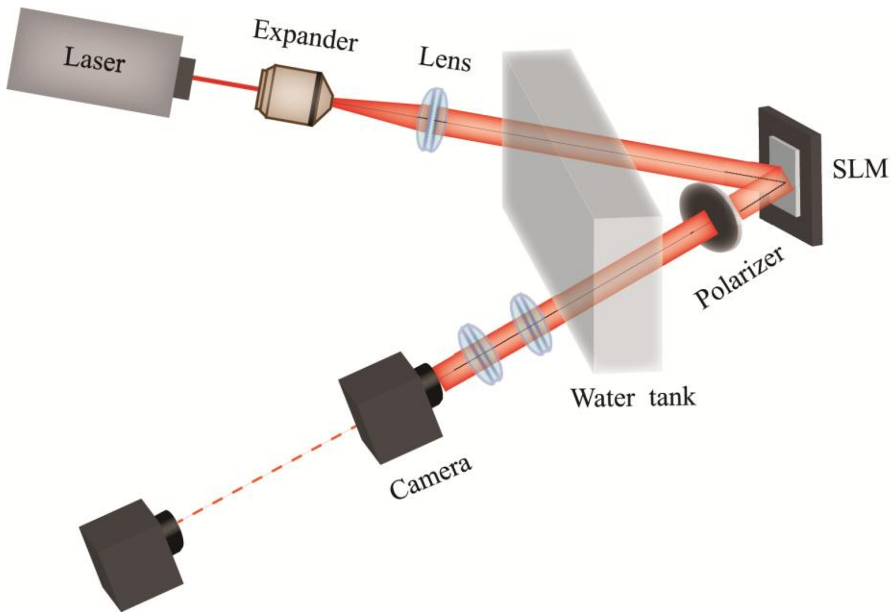

Figure 1. Experimental setup for imaging the objects in turbid media using the camera placed under a wide distance range. SLM: spatial light modulator.

2.2 Theoretical Analyses

Machine learning for image recovery consists of two consecutive steps: the first step is to train a neural network by using a given dataset, and then the second step is to predict an unknown datum (recorded in the same environment) using the trained network. The presented CNN architecture for underwater imaging in our experiments also includes two phases. In the training phase, speckle patterns are fed to a designed learning model. Since the diffraction patterns are recorded at different distances, dimensions of the speckles are varied as the changes of the position. At the beginning, the input images are downsized to $d1 \times d1$ images which contain adequately useful information. Then, the input image convolves with 20 kernels (size of 5×5) forming the first convolution layer. The activation function used in the convolution layer is sigmoid function. Then the first convolution layer (size of $(d1-4) \times (d1-4) \times 20$) is downsampled to $[(d1-4)/2] \times [(d1-4)/2] \times 20$ to generate the first pooling layer. Pooling is an important action to be taken to reduce computational load. Then, the first pooling layer continues to convolve with 20 kernels which are of the same size with that used in the first convolution layer. The second convolution layer is of size $\{[(d1-4)/2]-4\} \times \{[(d1-4)/2]-4\} \times 20$ followed by the second pooling layer (size of $(\{[(d1-4)/2]-4\}/2) \times (\{[(d1-4)/2]-4\}/2) \times 20$). After two rounds of convolution and pooling processing, the second pooling layer is reshaped to one-dimension vector. To establish a relationship between the input image and the ground truth (28×28), it is essential to fully connect the reshaped one-dimension vector to a vector of size 1×784. Then the second transformation is to reshape the fully connected layer (1×784) to a 28×28 image which is viewed as the final estimation of the input image. With pairs of diffraction patterns and corresponding ground truths sent to the designed model, the CNN structure can be considered to have learned a specific environment but cannot provide exact representations of explicit values of the parameters. Although precise representation of the environment cannot be given, the CNN model is capable to recover the object from the raw intensity speckle pattern. An example is displayed in Fig. 2 to verify the effectiveness of the trained CNN structure in the testing phase. It is obvious that the testing image is successfully recovered by using the trained CNN model and quality of the reconstruction is satisfactory. The CNN architecture is implemented by using Matlab 2009 with Nvidia Geforce GTX1080Ti.

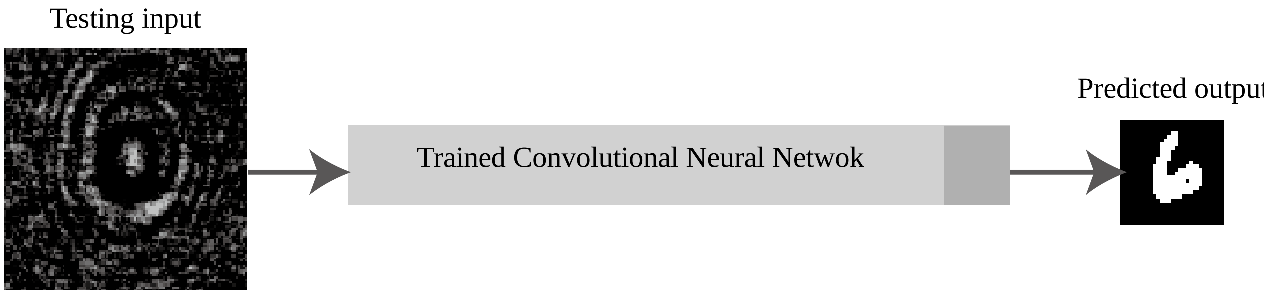

Figure 2. Testing phase: a testing result of the trained learning model for image reconstruction from the speckle pattern.

It is worth mentioning that the speckle pattern is pre-processed to remove its DC component through subtracting its mean value. Mean squared error (MSE) is used to evaluate quality of the recovered image which is given by

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2, \tag{1}$$

where n denotes pixel number of the original one, and Y_i and $\hat{Y}_i$ represent the pixel value of the ground truth and the predicted image, respectively. To continuously update the parameters (including weights and bias) for well-matched CNN model, stochastic gradient descent (SGD) is used as the optimization function to minimize the MSE value. After 5 epochs of iterations, the value of MSE decreases gradually, and then stabilizes at an optimal value. Finally, the designed CNN model is trained to learn the relationship between speckle patterns and ground truths, and then can be applied to make predictions of unknown objects from speckles.

3. EXPERIMENTAL RESULTS AND DISCUSSION

At the beginning, the murky medium is made by mixing $1200\ ml$ of water with $40\ ml$ of milk. The camera is placed at $1\ mm$ to $90\ mm$ away from the focal plane. 2000 handwritten digits from the MNIST database [10] are sequentially sent to the SLM. Consequently, 2000 diffraction patterns are recorded at each distance, where 1900 speckle patterns and their corresponding ground truths are used as the training dataset and another 100 pairs are used to test performance of the trained CNN model. Without sophisticated equipments to ascertain precise position of the camera, the camera is moved backward every $10\ mm$ until nothing can be recognized. Reconstruction examples at different distances are given in Fig. 3. The speckle patterns in the left column are recorded under the turbid medium with $40\ ml$ of milk and at different distances. The right column displays the reconstructed images. To reduce the computational load, the speckle patterns are cropped into appropriate sizes for image retrieval. For speckles recorded at the distance of $1\ mm$, $10\ mm$, $20\ mm$, $30\ mm$, $40\ mm$, $50\ mm$, $60\ mm$, $70\ mm$, $80\ mm$ and $90\ mm$, corresponding dimensions of the cropped patterns are 200×200, 200×200, 240×240, 240×240, 300×300, 400×400, 400×400, 400×400, 460×460 and 480×480, respectively. Dimension of the recorded speckle pattern is increased with the larger distance between the focal system and the camera. Time used to train the CNN model is varied and increases greatly with the enlargement of the dimension. It is obvious that quality of the reconstructed objects decreases sharply as the increase of the distance. It is worth mentioning that there is the existence of low-quality and wrong reconstructions. An explanation for this phenomenon can be ascribed to the strongly scattered media which severely degrades discrimination capability of the speckles.

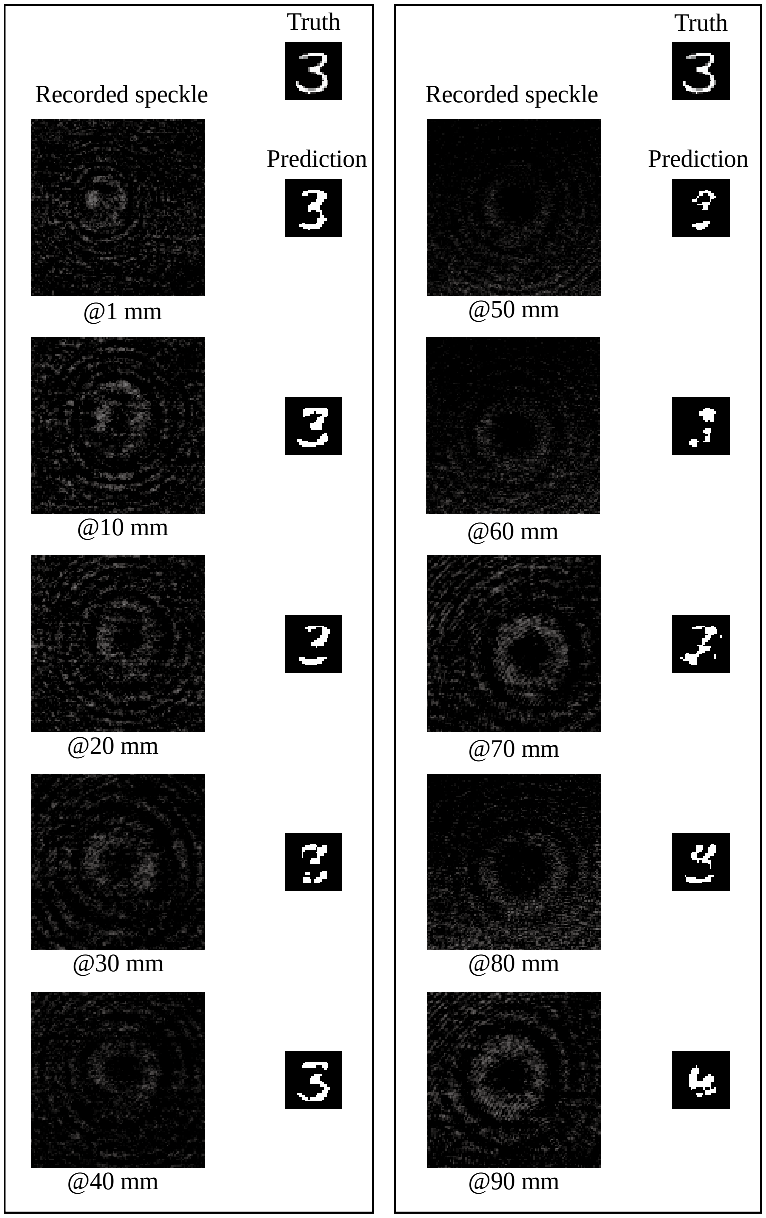

Figure 3. Reconstructed images with the camera placed at different axial distances.

However, training a CNN model for each distance is time-consuming. To build a learning model applicable for multiple distances, we try to combine the dataset recorded at each distance together to form a wide-ranging dataset to train a CNN model. However, difficulties are that size of the speckles is not uniform, and there are vast differences among the sizes. It is not realistic to crop all the speckles to the same dimension. Otherwise, the time used to train the model can be extremely large. In total, there are 5 groups of learning models for the objects recorded at uncertain distances. As shown in Fig. 4, the speckle patterns are recovered by general models and each separate model, respectively. It can be seen that quality of the reconstruction recovered by using the general CNN model is comparable with that given by each separate CNN model. Each generalized CNN model is trained by 2 groups of datasets, since the diffracted patterns recorded under nearby distances share some similarities in dimensions and shapes. Without precision techniques to measure the precise distance, it is possible to extract the objects from the speckles by using one of the learning CNN models, but it is at the cost of image quality. Here, it is experimentally verified that our CNN model is robust to different axial distances.

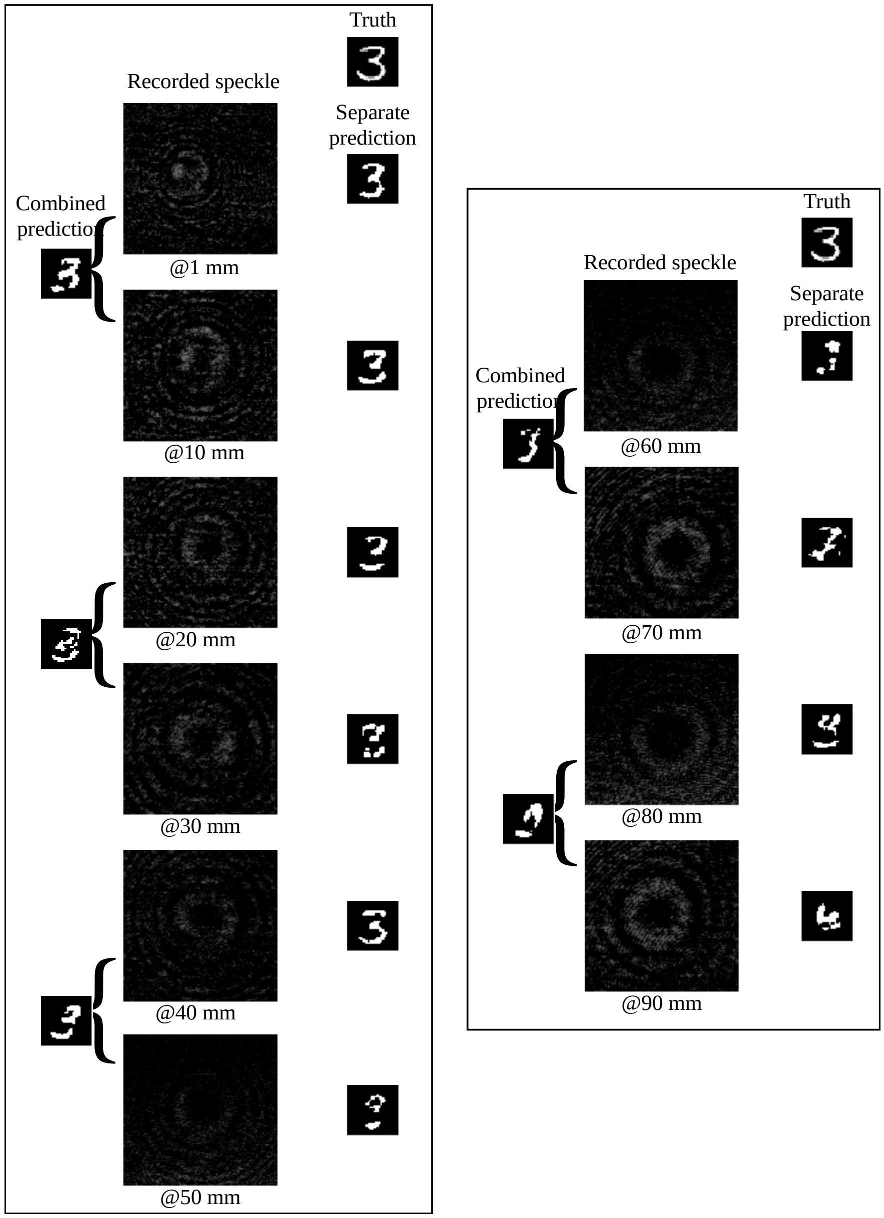

Figure 4. Comparisons of the predictions respectively obtained from combined models and separate models.

The learning models which are trained by using several groups of data recorded at different distances work effectively for recovering the objects recorded at uncertain positions.

4. CONCLUSIONS

We exploit optical imaging through turbid water at different axial distances. The general CNN model for image reconstruction can be utilized to bypass the variables aroused by unknown distances. It is experimentally demonstrated that the CNN models are feasible and effective. Though the precise distance of the camera is unknown, the presented method is advanced to extract the objects from the speckles. Quality of the reconstructions is satisfactory. The learning methods for image retrieval lead to decreased requirements of cutting-edge precision measurement techniques.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (NSFC) (61605165) and Hong Kong Research Grants Council (25201416).

REFERENCES

[1] A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, “Controlling waves in space and time for imaging and focusing in complex media,” Nature Photon. 6(5), 283, (2012).

[2] V. Ntziachristos, “Going deeper than microscopy: the optical imaging frontier in biology,” Nat. Methods 7(8), 603–614, (2010).

[3] S. M. Popoff, G. Lerosey, M. Fink, A. C. Boccara, and S. Gigan, “Image transmission through an opaque material,” Nat. Commun. 1(6), 81, (2010).

[4] C. D. Mobley, Light and Water. Radiative Transfer in Natural Waters, (Academic Press, New York, 1994).

[5] J. W. Goodman, W. H. Huntley, D. W. Jackson, and M. Lehmann, “Wavefront-reconstruction imaging through random media,” Appl. Phys. Lett. 8(12), 311–313, (1966).

[6] A. V. Kanaev, K. P. Judd, P. Lebow, A. T. Watnik, K. M. Novak, and J. R. Lindle, “Holographic imaging through extended scattering media under extreme attenuation,” OSA Imaging Systems, ITu3E, San Francisco CA, June 26–29, (2017).

[7] A. Liutkus, D. Martina, S. Popoff, G. Chardon, O. Katz, G. Lerosey, S. Gigan, L. Daudet, and I. Carron, “Imaging with nature: compressive imaging using a multiply scattering medium,” Sci. Rep. 4, 5552, (2014).

[8] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” IEEE Trans. Pattern Analysis Machine Intelligence 35(8), 1798–1828, (2013).

[9] L. Zhou, Y. Xiao and W. Chen, “Imaging through turbid media with vague concentrations based on cosine similarity and convolutional neural network,” IEEE Photon. J. 11(4), 7801315, (2019).

[10] L. Deng, “The MNIST database of handwritten digit images for machine learning research [best of the web],” IEEE Signal Process. Mag. 29(6), 141–142, (2012).