Aberration measurement from confocal axial intensity response using neural network

Y. Yasuno and T. Yatagai

Institute of Applied Physics, University of Tsukuba, Tennôdai 1-1-1, Tsukuba, Ibaraki, 305-8573, Japan

yasuno@optlab2.bk.tsukuba.ac.jp

http://optlab2.bk.tsukuba.ac.jp/

T. F. Wiesendanger, A. K. Ruprecht and H. J. Tiziani

Institut für Technische Optik, Universität Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany

http://www.uni-stuttgart.de/ito/

Abstract: We propose a high-speed, parallel system for lens aberration measurement employing a confocal optical setup. This method uses a non-interferometric, conventional confocal axial response to determine the spherical aberration coefficient of a confocal objective. The aberration coefficients are successfully calculated from the intensity axial response by employing a neural network. It is estimated that the system can find out the aberration coefficients of 10,000 microlenses in 20 seconds of measurement and 1 second of calculation time. Our experimental results also demonstrate the practicality of this system.

OCIS codes: (180.1790) Confocal microscopy; (200.4260) Neural networks

References and links

1. H. J. Matthews and D. K. Hamilton and C. J. R. Sheppard, "Aberration Measurement by Confocal Interferometry," J. Mod. Opt. **36**, 233–250 (1989).
2. H. Zhou and M. Gu and C. J. R. Sheppard, "Investigation of Aberration Measurement in Confocal Microscopy," J. Mod. Opt. **42**, 627–638 (1995).
3. H. Zhou and C. J. R. Sheppard, "Aberration measurement in confocal microscopy: phase retrieval from a single intensity measurement," J. Mod. Opt. **44**, 1553–1561 (1997).
4. Hans J. Tiziani, Hans-Martin Uhde, "Three dimensional analysis by a microlens-array confocal arrangement," Appl. Opt. **33**, 567–572 (1994).
5. H. J. Tiziani and R. Achi and R. N. Kramer and L. Wiegers, "Theoretical analysis of confocal microscopy with microlenses," Appl. Opt. **35**, 120–125 (1996).
6. H. J. Tiziani and T. Haist and S. Reuter, "Optical inspection and characterization of microoptics using confocal microscopy," Opt. and Laser Eng. **36**, 403–415 (2001).
7. T. R. Corle and C.-H. Chou and G. S. Kino, "Depth Response of Confocal Optical Microscopes," Opt. Lett. **11**, 770–772 (1986).
8. T. Wilson and A. R. Carlini, "The Effect of Aberrations on the Axial Response of Confocal Imaging System," J. Microscopy **154** 243–256 (1989).
9. Joseph M. Geary and Phil Peterson, "Spherical aberration: a possible new measurement technique," Opt. Eng. **25** 286–291 (1986).
10. Qian Gong and Smiley S. Hsu, "Aberration measurement using axial intensity," Opt. Eng. **33** 1176–1186 (1994).
11. Todd K. Barrett and David G. Sandler, "Artificial neural network for determination of Hubble Space Telescope aberration from stellar images," Appl. Opt. **32**, 1720–1727 (1993)

#1499 - $15.00 US
Received July 25, 2002; Revised November 28, 2002

12. Tae-Seok Yang and Jun Ho Oh, "Identification of primary aberrations on a lateral shearing interferogram of optical components using neural network," Opt. Eng. **40**, 2771–2779 (2001).
13. Developed at University of Stuttgart, Maintained at University of Tübingen, "Stuttgart Neural Network Simulator," http://www-ra.informatik.uni-tuebingen.de/SNNS/
14. D. E. Rumelhart and G. E. Hinton and R. J. Williams, "Learning representation by back-propagation errors," Nature **323**, 583–586 (1986).

1. Introduction

Like VCSEL arrays, microlens arrays are also becoming increasingly important for applications requiring small and spatially parallel optical devices. For industrial applications, optical communications, etc., it is becoming of crucial importance to be able to evaluate the aberrations of great numbers of micro lens arrays. Conventionally, interferometric methods have been extensively employed for measurement of aberration. However, the interferometric method requires a high-resolution image detector, making it extremely difficult and time-consuming to achieve simultaneous detection of large numbers of lenses.

A confocal interferometer has also been employed to measure the aberration of the measured object and the objective itself using a confocal microscope[1, 2, 3]. On the other hand, a parallel confocal microscope with a microlens array objective has been studied[4, 5]. These suggest that a parallel confocal interferometer which can simultaneously determine the aberration of a microlens array can be constructed by combining these two techniques[6].

However, when using an interferometric method such as this, it is necessary to detect the carrier frequency of the confocal interferogram, which requires high-accuracy, high-density axial scanning. Furthermore, depending on the type of interferometer, the system needs to be insulated from any vibration. However, non-interferometric aberration measurement system based on axial intensity response measurement would be both rapid and robust.

Although it is well known that non-interferometric axial intensity response does not contain enough information to retrieve the complete phase information[3], any aberrations in the objective markedly affect the axial intensity response[7, 8] and non-confocal axial intensity response[9, 10].

In this paper, we propose the use of a neural network to determine the spherical aberration coefficient of the objective from the axial intensity response. Neural networks have often been employed to estimate optical parameters using optical signals with insufficient information[11, 12]. To retrieve this information, the neural network compares numerically calculated reference signals with the signals to be measured. In this study, we created axial intensity response signals with several known spherical aberrations as a reference set, and compared them with measured axial intensity response signals. Although this system cannot determine the entire phase distribution of the objective, the use of a neural network enables the spherical aberration coefficients to be determined rapidly and flexibly.

2. Optical setup

Figure 1-(a) is a schematic diagram of a confocal setup for measuring confocal axial intensity response, which includes the properties of the objective. This setup is a simple and exemplary reflective confocal setup based on mirror scanning, except an objective lens is replaced by a microlens array whose aberration coefficient will be measured. We obtain the several axial intensity responses, like that shown in Fig. 1-(b), by only one axial scanning operation of the reference mirror. Each axial response corresponds to each microlens of the microlens array.

#1499 - $15.00 US Received July 25, 2002; Revised November 28, 2002

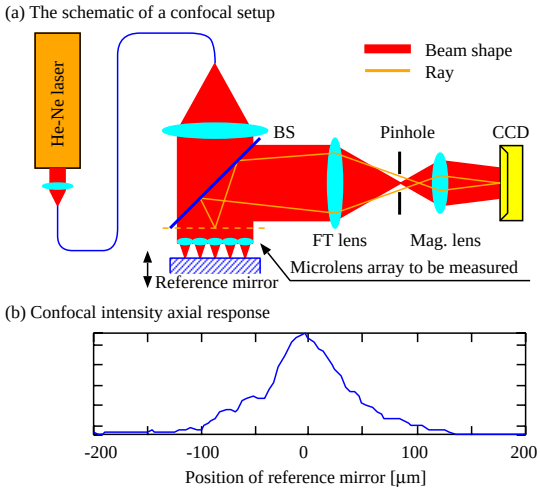

Fig. 1. (a) Schematic diagram of the confocal setup employed to measure the axial intensity response, where FT lens is a Fourier transform lens, Mag. lens is a magnification lens, and BS is a beam splitter. The pupils of the microlenses are imaged on the CCD camera, and the reference mirror is driven axially by a stepping motor. The response includes the aberration information of the objective. (b) One example of axial intensity response.

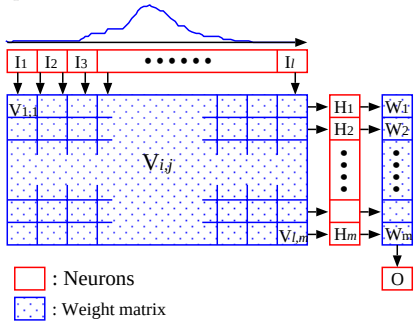

Fig. 2. The schema of a neural network for determining aberration coefficient from confocal intensity-axial-response.

In this setup, the pinhole in a conventional single pinhole, not a pinhole array. It is due to that the pinhole is placed on the focal plane of a Fourier transform lens[4]. And the diameter of the pinhole is selected not to block ±1st order and to block higher order diffraction of the array structure of the microlens. This correct diameter- pinhole separates each axial response of each microlens and, at the same time, works as a confocal pinhole.

In conventional confocal measurement, we place the object to be measured in the place of the mirror; the axial response signal contains information of both of the object and the objective. In this paper, we use an optically flat mirror as the object, so that the axial response contains only the information of the objective.

3. Neural network design

To estimate the aberration coefficient from the axial intensity response, we constructed a neural network which inputs the axial response and outputs the spherical aberration coefficient of the objective. We employed a SNNS neural network simulator[13] on a Pentium III personal computer as the actual implementation of the neural network.

#1499 - $15.00 US
Received July 25, 2002; Revised November 28, 2002

This neural network has 99 neurons in the input layer, 50 neurons in the hidden layer, 1 neuron in the output layer, and a fully connected network structure, as shown in Fig. 2. In this network, the input vector I_i is the axial response itself, and the output O is a linearly compensated spherical aberration coefficient. This compensation strategy uses a bias and scale factor, which is necessary because the output value of this neural network is permitted to range only between 0 and 1. Here the aberration coefficient [-0.7λ ~0.3λ] is calibrated into [0.05 ~ 0.95] with the operating wavelength, λ, of 632.8 nm.

The value of the hidden neurons are calculated as

$$H_j = f\left(\sum_i^l V_{i,j} I_i\right) \tag{1}$$

where $V_{i,j}$ is the weighted matrix between the input layer and the hidden layer, and function of f represents a sigmoid function[14]. The output is then calculated from the hidden vector as

$$O = f\left(\sum_j^m W_j H_j\right) \tag{2}$$

where $W_{i,j}$ is the weighted matrix between the hidden layer and the output layer. Finally, this value is linearly calibrated into an actual spherical aberration coefficient.

We trained this neural network to estimate coefficients correctly using several training signals. We used numerically calculated axial intensity responses as the training set. The training set is calculated using the following parameters. The focal length of the micro lenses is 400 μm, ten spherical aberration coefficients from -0.7λ to +0.3λ are considered, and the scanning range of the reference mirror is 400 μm. Every axial response is shifted to make its center of gravity at FWHM the center of the input vector. This shifting operation reduces the required accuracy of absolute positioning of the scanning mirror in the experiment. We also used $\pm$4 μm-shifted training signals to make the neural network more robust. Hence, the training signal set contains a total of 33 axial responses, namely, 11 spherical aberration coefficient; {-0.7λ, -0.6λ, -0.5λ, -0.4λ, -0.3λ, -0.2λ, -0.1λ, -0.0λ, 0.1λ, 0.2λ, 0.3λ} times 3 shifting variations; {-4 μm, 0 μm, +4μm}.

We employed back-propagation algorithm[14], one of the most commonly used algorithms to train the multilayer neural network. The initial values of the weight matrices V and W are randomly determined between -0.1 to 0.1. According to the back-propagation algorithm, the weight matrices are changed as

$$W_j^{n+1} = W_j^n + \alpha(O^n - T)O^n(1 - O^n)H_j^n \tag{3}$$

$$V_{i,j}^{n+1} = V_{i,j}^n + \alpha(O^n - T)O^n(1 - O^n)W_j^n H_j^n(1 - H_j^n) \tag{4}$$

where the superscript n denotes that the particular value of the weight is taken at n-th epoch, T is a ideal output value namely, a calibrated spherical aberration coefficient. The learning coefficient α determines the learning speed. Although the learning speed is faster with bigger α and slower with smaller α, the learning accuracy is more with smaller α and less with bigger α. In learning process, thus, we dynamically changed α as, $\alpha = 0.5$ for 1st to 25000th epoch, $\alpha = 0.3$ for 25001st to 35000th epoch, and $\alpha = 0.1$ for 35001st to 45000th epoch. Thus by changing α, the network has been trained rapidly in the primary stage of learning process, then trained accurately in the later stage. At each epoch, although, we use all of the 33 training signals to train the network, the order of the training set has been randomly changed in every epoch to enable to find the best solution, not any local solution.

#1499 - $15.00 US Received July 25, 2002; Revised November 28, 2002

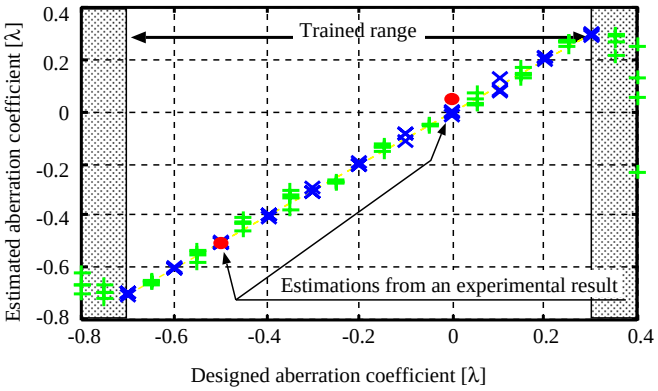

Fig. 3. Estimated spherical aberration coefficients. The lateral axis and the height axis respectively represent designed aberration coefficients and estimated coefficients. $\times$, $*$, and $\circ$ respectively denote the results from trained, untrained, and experimental axial responses. The solid line represents an ideal estimation.

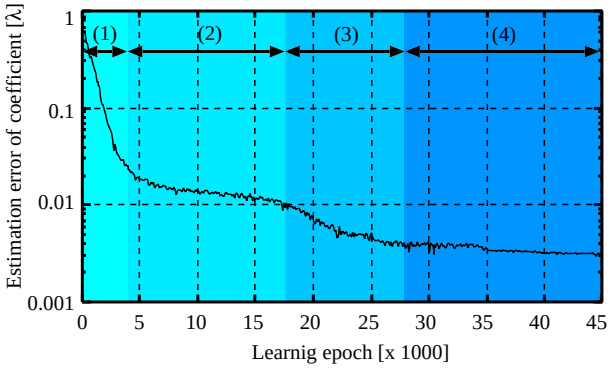

Fig. 4. The learning curve of the neural network. The learning process can clearly be classified into four processes.

4. Aberration estimation using neural network

This trained neural network was evaluated with several axial responses, namely, (1) axial responses contained in the training set, (2) numerically calculated axial responses not contained in the training set, and (3) experimentally measured axial responses. In the last case, we used two diffractive objective lenses, one with zero aberration and the other with a spherical aberration coefficient of -0.5λ. Figure 3 shows the estimated aberration coefficients, where $\times$, $+$ and $\circ$ respectively denote the cases of (1), (2) and (3). The three results for one designed aberration coefficient represent the results of ± 4 μm and 0 shifted inputs. The results of experimental ones are averages of results of three different shift- value 0 and ± 4 μm. Figure 3, it is evident that the neural network correctly estimates the aberration coefficient within the training range. The RMS error is 0.01λ, 1% of the dynamic range of the system.

This system estimates closely similar values for the inputs with the same aberration coefficient but different shift quantity. On the other hand, it estimates different coefficient values for the inputs with the same aberration but outside of the training range. This shows that the network is operating correctly within the range.

#1499 - $15.00 US
Received July 25, 2002; Revised November 28, 2002

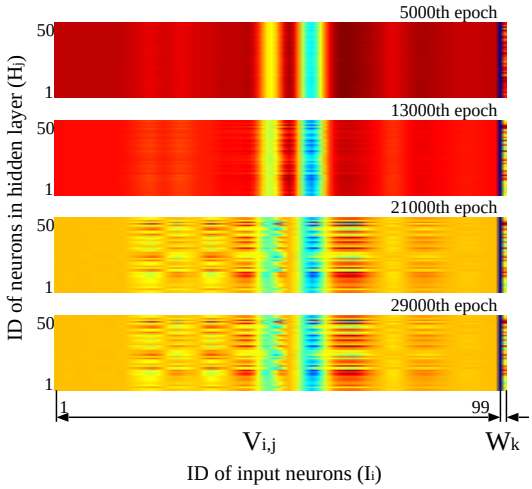

Fig. 5. Various activities of the weight in different processes, in the second and third phases. The weight matrices are aligned as the same as Fig. 2, V is the weight matrix between the input and the hidden layer, and W is the weight matrix between the hidden and the output layer. The horizontal axis of V represents the ID of input neurons i, and vertical axes of V and W represent the ID of hidden neurons j.

5. Analysis of neural network

Here we discuss the learning process of the neural network. First, the weight matrix $V_{i,j}$ and $W_{i,j}$ are randomly initialized. Then we iterate the calculation to optimize the weight matrix. This operation consists in calculating the estimated output by equations (1) and (2), and calculating the error between the estimated and ideal output, and modifying the weight matrices by equations (3) and (4). Figure 4 shows the learning curve and the history of errors for each iteration. The learning curve shows four phases. The first phase is simply for adjusting the total energy of the weight matrices. The important phases are the second and third phases. The absolute value, namely activity, of the weight matrix $V_{i,j}$ after several iterations is shown in Fig. 5. Here the weight matrices are aligned as the same as Fig. 2, weight matrices V are the matrices between the input and the hidden layer, and weight matrices W are the matrices between the hidden and the output layer. The horizontal axis of $V_{i,j}$ represents the ID of input neurons i, and vertical axes of $V_{i,j}$ and W_j represent the ID of hidden neurons j. In the second phase, the weight matrix has only a one-dimensional structure, whereas in the third phase, it has a two-dimensional structure. The lateral axis of $V_{i,j}$ in this figure corresponds to the scanning depth of the confocal setup. Hence we conclude that in the second phase, the neural network mainly selects the important part of the axial response; then in the third phase, it creates a network structure that can calculate actual aberration coefficients from key signals. The fourth phase is for fine adjustment of the weight matrices.

6. Working speed

Finally, we discuss the working speed of this system. Here we use 99 neurons in the input layer. The scanning range of the reference mirror is 400 μm. This means that 99 slices for each 4 μm are needed in one axial scanning. The number of slices, at 99, is fewer than in the case of a confocal interferometer, allowing high-speed measurement. In our prototype system, this operation takes only 20 seconds. This long scanning time is due to the slow driving and control time of the stepping motor employed to scan the reference mirror. We can easily reduce this measurement time by employing a piezo-

#1499 - $15.00 US
Received July 25, 2002; Revised November 28, 2002

electric transducer or other high speed scanning stages.

The evaluation time of a neural network for one input is estimated as

$$\tau = L(M+N)\delta_m + [L(M-1) + N(L-1)]\delta_a \tag{5}$$

where L, M, and N are respectively the numbers of neurons in the input, hidden, and output layer; and δ_m and δ_a are the calculation times of a CPU for one floating point multiplication and addition. We also have to take into account the time required to calculate the center of gravity, and the time for shifting operation of an axial response. This needs to be done three times for one signal to create three signals with different shifts. For this reason, Equation (5) is modified to

$$\tau' = (2M + NL + LM + 1)\delta_m + (LM - NL + N - L + 3M)\delta_a. \tag{6}$$

For example, by substituting $\delta_a \sim \delta_m \simeq 1.4$ ns, the actual benchmark value of the Athlon 700 MHz processor used and $M = 99$, $N = 50$, $L = 1$, the operation time of the neural network is estimated as $\tau' \sim 3$ μs. It suggests that, even if the actual calculation time is 10 times slower than this estimation, due to slow memory access or other reasons, our system is able to estimate the aberration coefficients of 10,000 microlenses within 1 second. This system requires only a few tens of seconds of operation time, including time for both measurement and computation, to determine the aberrations of a 100 $\times$ 100 microlens array.

7. Conclusion

In conclusion, we have proposed and constructed a high-speed aberration measurement system which determines the spherical aberration coefficient of the objective of a confocal setup by means of a neural network. Our prototype system can determine spherical aberration coefficients in the range from -0.7λ to 0.3λ with less than 1% RMS error. This measurement range can be adjusted to suit other situations.

In this paper, we only consider spherical aberration, because odd aberrations are mostly canceled by reflection type confocal setup, and defocus is also canceled in the signal shifting process. We believe that this system also can estimate the aberration coefficients of spherical aberration and astigmatism at the same time by using training signals which contain both of the aberrations.

We estimate that this system can rapidly determine the aberration coefficients of a 100 $\times$ 100 microlens array, with 20 second measurement time and less than 1 second calculation time.

#1499 - $15.00 US Received July 25, 2002; Revised November 28, 2002