Modal Amplitude and Phase Estimation of NFP of Six-Mode FMF Based on Artificial Neural Network with the Help of Grey-Wolf-Optimizer

N. Sugawara[(1)], T. Fujisawa[(1)], K. Nakamura[(1)], Y. Sawada[(1)],
T. Sakamoto[(2)], T. Matsui[(2)], K. Nakajima[(2)], and K. Saitoh[(1)]

[(1)] Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan, sugawara@icp.ist.hokudai.ac.jp
[(2)] NTT Access Network Service Systems Laboratories, NTT Corporation, Tsukuba 305-0805, Japan

Abstract A simple and efficient method for estimating modal amplitude and phase of few-mode-fiber based on artificial-neural-network with the help of grey-wolf-optimizer is proposed. Experimentally obtained six-mode near-field-patterns of recently proposed six-mode mixer are analyzed and successfully reproduced for the first time.

Introduction

In recent years, mode division multiplexing (MDM) transmission attracts a lot of attention for expanding transmission capacity. In MDM transmission, so-called differential mode delay (MDM) and mode dependent loss (MDL) deteriorate the receiver performance. To overcome these problems, a mode exchanging or mixing technique at a relay point between two FMFs is useful[1]. In a mode-exchanger or mode mixer, input modes are converted or mixed with other modes, and various devices were proposed for 3 and 6 modes[2-5].

In the mode mixers, the output is usually the mixture of multiple modes. To evaluate the experimental performance of the device, the estimation of modal amplitude and phase of the fabricated device is necessary. To estimate these quantities, various methods have been developed. Among them, a numerical estimation method based on near field patterns (NFP) of the beam is simple and fast compared with fully experimental methods. Recently, a numerical method for estimating the modal amplitude and phase from the NFPs of FMF based on artificial neural network (ANN) was reported[6]. In [6], an ”elaborated” NN based on convolutional NN, VGG-16[7], which is specially designed for large scale image recognition with various numerical techniques, was used to learn mixed field patterns in FMF. A successful NFP reconstruction was demonstrated theoretically for five modes and experimentally for three modes.

In this work, we, for the first time, experimentally demonstrate successful NFP reconstruction of recently proposed six-mode mixer[5]. We propose an efficient estimation method of modal and phase coefficients based on simple ANN with the help of grey-wolf optimizer [8]. In the method, we just use usual ANN, which is significantly simpler than the “elaborated” NN. To increase the accuracy of the NFP reconstruction, a grey-wolf optimizer (GWO)[8] is used to help the phase estimation. For the measured NFPs of six-mode mixer, the successful NFP reconstruction is demonstrated with the average accuracy of 0.96 in terms of correlation of images.

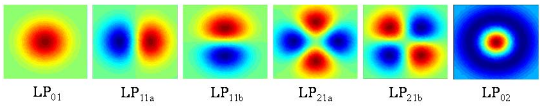

Fig. 1: Six mode fields of FMF.

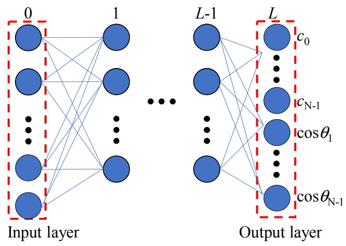

Fig. 2: A schematic of ANN.

Learning with ANN

We consider the mixture of six modes in FMF as shown in Fig. 1. The mixed trasnverse field $\phi(x,y)$ can be expressed as

$$\phi(x,y)=\sum_{k=0}^{N-1} c_k \phi_k(x,y) e^{-j\theta_k} \qquad (1)$$

where, N is the number of modes, c_k is the k-th modal amplitude with $\Sigma|c_k|^2 = 1$, $\phi_k(x,y)$ is the electric field of k-th each eigenmode, and θ_k is the modal phase, whose range is $[-\pi, \pi]$. Here, we set the phase of the fundamental mode $\theta_0 = 0$ and relative phases to the fundamental mode is treated. Therefore the goal of this paper is to estimate c_k and θ_k from given $\phi(x,y)$.

We use the ANN as shown in Fig. 2 for this purpose. The ANN is composed of input layer, output layer and multiple hidden layers. By renewing the weights of the nodes by error back propagation, the relationship between input and

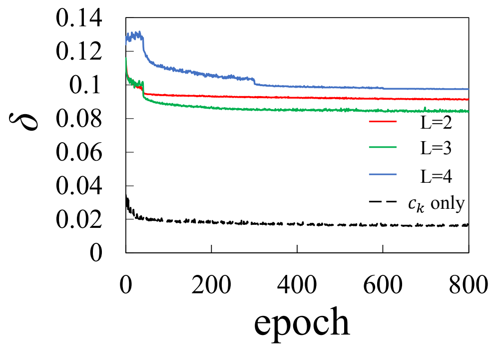

Fig. 3: Learning curves of c_k and θ_k (solid lines) and (b) c_k only (dashed line).

output characteristics are obtained. Here, the input layer is composed of the intensity (luminance) of NFP, in other words, pixel values of NFP picture. In this paper, we use images with the resolution of 64 × 64. The output layer contains c_k and $\cos\theta_k$, and there are $2N$-1 nodes. In the six-mode case, there are 11 nodes (six c_k and five $\cos\theta_k$). The reason for using $\cos\theta_k$ instead of θ_k is that there are multiple combinations of the phase that give the same NFP. This deteriorates the learning of ANN. As shown later, the learning is more efficient if only the modal amplitude is treated. In this case, the output nodes are six. The number of hidden layers is L-1, where L is the number of all layers except for the input layer. The number of nodes of i-th hidden layer is $N_{hd,i}$ (i = 1 to L-1). A sigmoid function is used for all the activation functions. No other special techniques, such as ReLU function, drop out, normalization, etc, are used.

We generate 100,000 NFP data from the field of Fig. 1 by randomly generating c_k and θ_k. 90,000 data were used for training and remaining 10,000 data were used for testing. The learning is evaluated by the average error given by

$$\delta = \sum_{t_N} \sum_{2N-1} \left| Z_o - Z_t \right| / t_N \left(2N-1 \right) \qquad (2)$$

where t_N is the number of testing data, Z_0 is the output value of ANN, and Z_t is the answer value of testing data. Also, the correlation value C of two images is defined as[6]

$$C = \frac{\sum_{j=1}^{s}\sum_{i=1}^{s}\left(I_r(i,j)-\overline{I}_r\right)\left(I_m(i,j)-\overline{I}_m\right)}{\sqrt{\left(\sum_{j=1}^{s}\sum_{i=1}^{s}\left(I_r(i,j)-\overline{I}_r\right)^2 \cdot \sum_{j=1}^{s}\sum_{i=1}^{s}\left(I_m(i,j)-\overline{I}_m\right)^2\right)}} \qquad (3)$$

where, s = 64 is the size of images in x and y directions, I_m, $\overline{I_m}$ is the luminance and average luminance values of inputted images, I_r, $\overline{I_r}$ is the luminance and average luminance values of ANN-reconstructed images, respectively. The range of the C is [0.0, 1.0], and C = 1.0 if two images are perfectly equal. Since we learn $\cos\theta_k$ for the phase, there are two candidates for θ_k. Therefore, there are 2^{N-1} candidates for the phase combination. We calculate C for each set of the phase and choose the phase set having the largest C.

Solid lines in Fig 3 show the learning curves for c_k and θ_k for different number of hidden layers. The error is reduced with the epoch and ANN with L = 3 (2 hidden layers) shows the best performance. However, there are still 8% errors. This is due to the phase ambiguity described above. A Dashed line in Fig 3 shows the learning curve for only c_k for L = 3. The error is significantly reduced compared with solid lines, showing the difficulty of estimating the phase from the intensity pattern.

Phase estimation based on GWO

To estimate the phase of the mixed mode fields with simple ANN shown in the previous section, we use a GWO[8]. The GWO is one of the optimizing algorithms mimicking the hunting of grey wolves. There are N_w wolves and the position of i-th wolf is denoted by vector $\boldsymbol{X}_i$. The dimension of $\boldsymbol{X}_i$ corresponds to the search space of wolves. Each wolf has a fitness value according to the position, and large fitness means that the wolf is near the prey. The position is renewed by the positions of three wolves, α, β, and δ, who have top 3 fitnesses. At the iteration t, random vectors $\boldsymbol{A}_i$ and $\boldsymbol{C}_i$ are initialized as

$$\boldsymbol{A}_i = 2a\boldsymbol{r}_i - a \qquad (4)$$

$$\boldsymbol{C}_i = 2\boldsymbol{r}_i \qquad (5)$$

where $a = 2\text{-}2t/N_{it}$, N_{it} is a total iteration count and $\boldsymbol{r}_i$ is a random vector. Then, the fitnesses of all wolves are evaluated by their positions $\boldsymbol{X}_i$. In this paper, the fitness and the position correspond to the correlation value, C, and the phase set of mixed mode fields, θ_k. Therefore, the dimension of $\boldsymbol{X}$ is N-1. From the fitness values, top 3 wolves are selected and their positions are $\boldsymbol{X}_\alpha$, $\boldsymbol{X}_\beta$, $\boldsymbol{X}_\delta$. From these positions, following vectors are calculated.

$$\boldsymbol{X}_{1,i} = \boldsymbol{X}_\alpha - \boldsymbol{A}_i \circ \boldsymbol{D}_{\alpha,i} \qquad (6a)$$

$$\boldsymbol{X}_{2,i} = \boldsymbol{X}_\beta - \boldsymbol{A}_i \circ \boldsymbol{D}_{\beta,i} \qquad (6b)$$

$$\boldsymbol{X}_{3,i} = \boldsymbol{X}_\delta - \boldsymbol{A}_i \circ \boldsymbol{D}_{\delta,i} \qquad (6c)$$

where ∘ denotes Hadamard product. $\boldsymbol{D}_n$ is given by

$$\boldsymbol{D}_{n,i} = \left| \boldsymbol{C}_i \circ \boldsymbol{X}_n - \boldsymbol{X}_i \right| \qquad (7)$$

where n = α, β, and δ. Then, the position of i-th wolf is renewed as

$$\boldsymbol{X}_i^{t+1} = \left(\boldsymbol{X}_{1,i} + \boldsymbol{X}_{2,i} + \boldsymbol{X}_{3,i} \right) / 3 \qquad (8)$$

where superscript shows the iteration count.

Here, we choose N_w = 32 and the initial positions of wolves are set to all phase sets given by the combinations of θ_k. The left row of Fig. 4 show the intensity distributions of three test data, No. 1, 2, and 3. The correlation values between test data and ANN-generated images are 0.986,

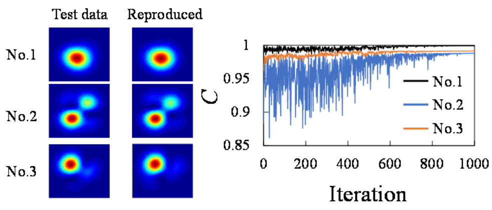

Fig. 4: Examples of three reproduced images of test data and phase estimations by GWO.

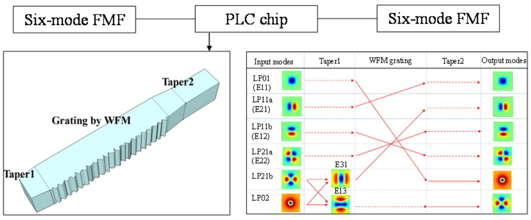

Fig. 5: A schematic of six-mode mixer.

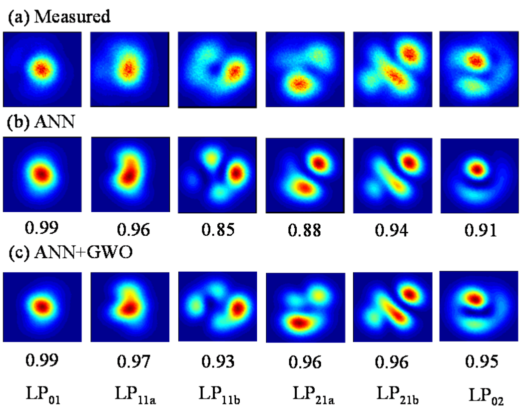

Fig. 6: (a) Measured, (b) ANN-reproduced, and (c) ANN+GWO-reproduced NFPs of six-mode mixer.

0.883, and 0.962. By using the phase set given by ANN, we estimate the true phase set by using the GWO, and *C* as a function of iterations are shown in the right panel of Fig. 4. The correlation values are increased and converged to larger values compared with the initial phase set given by ANN for all examples by increasing the iteration count. The final values of C are 0.999, 0.988, and 0.991 for three test data, No. 1, 2, and 3. The center panels of Fig. 4 shows the intensity distributions of three test data, generated by ANN-speculated power coefficients, and phase set estimated by the GWO. Almost identical images are obtained, showing the usefulness of the proposed approach.

NFP reconstruction of six-mode mixer

We use above described method to estimate power and phase coefficients of six-mode mixer[5]. Figure 5 shows the schematic of the device. Two six-mode FMFs[9] are pigtailed to a planar lightwave circuit (PLC) chip based on silica waveguide. In the PLC chip, a grating-like waveguide designed by a wavefront matching (WFM) method[10] is fabricated as shown in the Figure, and an input mode is converted to

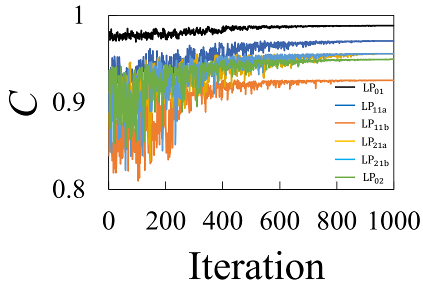

Fig. 7: *C* as a function of iteration for each mode input.

different modes. A schematic of designed mode exchanging operation is also shown in the Figure. The detail can be found in [5]. Since the waveguide has complex geometry, the fabrication imperfection affects the device performance. Furthermore, due to the degeneracy of fiber modes, LP_{11a} and LP_{11b} modes, and LP_{21a}, LP_{21b}, and LP_{02} modes are easily mixed in pigtailed FMFs. Therefore, to estimate the performance of the device, the power and phase coefficient estimations are very important.

Figure 6 (a) shows one example of measured output NFPs of the module for each input mode. Mixed mode fields are obtained. Figure 6 (b) shows the reproduced images of the six-mode mixer based on ANN-speculated power and phase coefficients. The values below the images indicate the correlation values. The values are below 0.9 for LP_{11b} and LP_{21a} mode inputs, and the images are not well-matched to the measured ones. This is because the phase set speculated by ANN has some errors as shown in Fig. 3. The phase set values can be improved by using the GWO and Fig. 7 shows the *C* as a function of iteration for each input mode. The *C* values are improved with the iteration, especially for LP_{11b} and LP_{21a} mode inputs. Figure 6 (c) shows the reproduced images and the *C* values of the six-mode mixer based on ANN-speculated power coefficients and phase set estimated by the GWO. Clearly, the images, especially for LP_{11b} and LP_{21a} mode inputs, are better matched to measured ones. The averaged *C* value is 0.96. Remaining errors may come from the existence of radiation modes in the NFPs, which is not taken into account in the theory.

Conclusion

Successful NFP reconstruction of six-mode FMF based on ANN approach is demonstrated for the first time. We propose the efficient estimation method of power and phase coefficients based on simple ANN with the help of GWO. Measured NFPs of recently proposed six-mode mixer is successfully reproduced with the average accuracy of 0.96 in terms of correlation of images.

References

[1] Y. Wakayama et al., "Intermediate mode interchange for reduction of differential mode-group delay in weakly-coupled 6-mode fiber transmission line," Proc of OFC2016, M3E.6, 2016.

[2] W. Jin, and K.S. Chiang, "Mode converters based on cascaded long-period waveguide gratings," Opt. Lett., vol. 41, pp.3130-3133, July 2016.

[3] T. Fujisawa *et al.*, "One chip, PLC three-mode exchanger based on symmetric and asymmetric directional couplers with integrated mode rotator," Proc of OFC2017, W1B.2, 2017.

[4] Y. Zhao et al., "Broadband and low-loss mode scramblers using CO2-laser inscribed long-period gratings," Opt. Lett., vol. 43, pp. 2868-2871, June 2018.

[5] T. Fujisawa et al., "Wavefront-matching-method-designed six-mode-exchanger based on grating-like waveguide on silica-PLC platform," Proc of OFC2020, Th1A.5, 2020.

[6] Yi An et al., "Learning to decompose the modes in few-mode fibers with deep convolutional neural network," *Opt. Express*, vol. 27, no. 7, pp. 10127–10137, Apr. 2019.

[7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition,", https://arxiv.org/abs/1409.1556.

[8] S. Mirjalili et al., "Grey wolf optimizer," Advances in Engineering Software, vol. 69, pp.46-61, Jan. 2014.

[9] T. Mori et al., "Few-mode fibers supporting more than two LP modes for mode-division-multiplexed transmission with MIMO DSP", J. Lightwave Technol., vol. 32, pp. 2468-2479, July 2014.

[10] Y. Yamashita et al., "Design and fabrication of broadband PLC-based 2-mode multi/demultiplexer using wavefront matching method," J. Lightwave Technol., vol. 35, pp. 2252- 2017, June 2017.