室蘭工業大学
学術資源アーカイブ
Muroran Institute of Technology Academic Resources Archive

A Study on Optimal Design of Optical Devices Utilizing Coupled Mode Theory and Machine Learning

メタデータ	言語: eng 出版者: 公開日: 2020-12-23 キーワード (Ja): キーワード (En): neural network (NN), coupled mode theory (CMT), directional coupler type photonic devices, optimal design, hybrid firefly algorithm 作成者: KUDO, Koji, MORIMOTO, Keita, IGUCHI, Akito, 辻, 寧英 メールアドレス: 所属:
URL	http://hdl.handle.net/10258/00010350

IEICE TRANS. ELECTRON., VOL.E103–C, NO.11 NOVEMBER 2020

PAPER Special Section on Recent Advances in Simulation Techniques and Their Applications for Electronics

A Study on Optimal Design of Optical Devices Utilizing Coupled Mode Theory and Machine Learning

Koji KUDO[†a)], **Keita MORIMOTO**[†b)], ***Student Members***, **Akito IGUCHI**[†c)], ***Member***, ***and*** **Yasuhide TSUJI**[†d)], ***Senior Member***

SUMMARY We propose a new design approach to improve the computational efficiency of an optimal design of optical waveguide devices utilizing coupled mode theory (CMT) and a neural network (NN). Recently, the NN has begun to be used for efficient optimal design of optical devices. In this paper, the eigenmode analysis required in the CMT is skipped by using the NN, and optimization with an evolutionary algorithm can be efficiently carried out. To verify usefulness of our approach, optimal design examples of a wavelength insensitive 3 dB coupler, a 1 : 2 power splitter, and a wavelength demultiplexer are shown and their transmission properties obtained by the CMT with the NN (NN-CMT) are verified by comparing with those calculated by a finite element beam propagation method (FE-BPM).

key words: *neural network (NN), coupled mode theory (CMT), directional coupler type photonic devices, optimal design, hybrid firefly algorithm*

1. Introduction

With the rapid spread of the Internet and the development of IoT, the demand for high-speed and large-capacity optical communication systems is increasing. Nowadays, with the development of computer simulation technology, it is possible to analyze and design optical devices required in optical communication systems, on personal computers. The optimization of the optical device is generally done by iterative approach including evolutionary methods based on multipoint search [1], gradient methods based on sensitivity analysis, or other approaches. Above all, evolutionary methods require a large number of numerical simulations. For efficient design, it is necessary to improve computational efficiency of the optical device analysis.

Recently, machine learning with a neural network (NN) has begun to be utilized in optimal design of optical devices [2]–[6]. In [2] and [3], the NN is learning to predict the output characteristics of an arbitrarily given device structure, and inverse design is performed to design the optimal device structure by inputting the desired output. However, in the inverse design, the solution non-uniqueness problem must be considered. It has been shown that this non-uniqueness problem can be overcome using a tandem structured NN in some design problems [4]. On the other hand, in [6], the efficient optical property estimation of five kinds of device models utilizing a relatively simple NN is demonstrated. However, actual design example is not shown. Although the inverse design network is constructed and used to design photonic devices in [2]–[5], the applicability seems to be limited and it seems to be still difficult to apply these inverse design network to arbitrary multimodal problems. On the other hand, it is relatively easy to evaluate device properties of some basic optical components by using NN as demonstrated in [6]. Therefore this kind of NN is able to be used to improve computational efficiency of analyzing more complicated optical devices and it leads to improve the computational efficiency of automatic optimal design.

In this paper, we propose a new design approach to improve the computational efficiency of the optimal design of optical waveguide devices. In our design approach, an evolutionary method as an optimization method and a coupled mode theory (CMT) [7] as an numerical analysis method are employed. Furthermore, in order to achieve an efficient optimal design, the mode coupling coefficient and the propagation-constant mismatch required in the CMT analysis are predicted utilizing a relatively simple NN without time consuming numerical calculation.

In Sect. 2, we describe the behavior of the NN used in our approach. In Sect. 3, we briefly review the CMT. In Sect. 4, we describe an evolutionary method employed in this study. In Sect. 5, the numerical accuracy of the NN-CMT is illustrated. In Sect. 6, we design wavelength a insensitive 3 dB coupler, a 1 : 2 power splitter, and a wavelength demultiplexer as optimal design examples utilizing our approach. In Sect. 7, we conclude this paper.

2. Neural Network

The NN is a mathematical model of neurons in the human brain. As shown in Fig. 1, neurons are connected by synapses, and information is transmitted through the synapses. The output of the NN is determined by the weights, which is the strength of the synaptic connections between each neuron. In the learning process, the weights are adjusted so that various input/output relationships can be expressed. There are hidden layers between the input layer and the output layers, and a complex input/output relationship can be expressed with increasing the hidden layers.

The behavior of a single neuron is shown in Fig. 2. One neuron is weighted with the output from the neurons in the

Manuscript received October 10, 2019.
Manuscript revised February 19, 2020.
Manuscript publicized March 25, 2020.

[†]The authors are with the Information and Electronic Eng., Muroran Institute of Technology, Muroran-shi, 050–8585 Japan.
a) E-mail: 19043027@mmm.muroran-it.ac.jp
b) E-mail: 19096015@mmm.muroran-it.ac.jp
c) E-mail: iguchia@mmm.muroran-it.ac.jp
d) E-mail: y-tsuji@mmm.muroran-it.ac.jp

DOI: 10.1587/transele.2019ESP0002

Copyright © 2020 The Institute of Electronics, Information and Communication Engineers

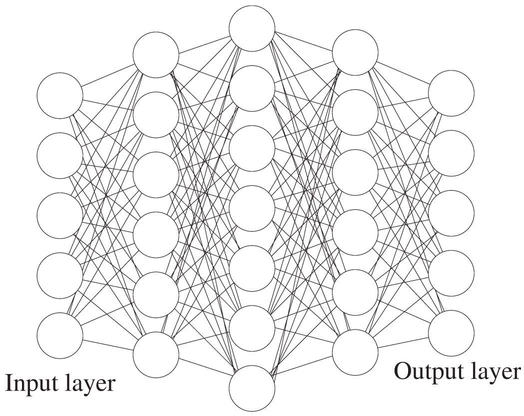


Fig. 1 Basic neural network.

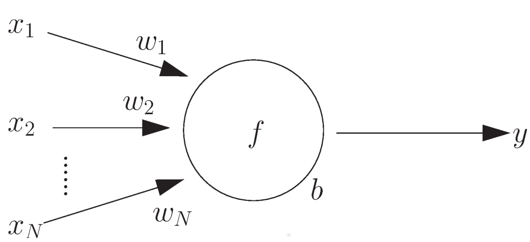


Fig. 2 A neuron and synapses in neural network.

previous layer. The output of the neuron is determined by the sum of these inputs and a bias. The output of the neuron is expressed as follows:

$$y = f(x_1 w_1 + x_2 w_2 + \cdots + x_N w_N + b) \tag{1}$$

where x_i $(i = 1, 2, \cdots, N)$ is an input variable, w_i is a weight, b is a bias, and an output y is obtained through an activation function f. In this study, the hyperbolic tangent function is used as an activation function of the hidden layer and an identity function is used in the output layer.

In order to construct the NN by machine learning, the back-propagation method [8] is used. The cost function E is defined as follows:

$$E = \frac{1}{2}\sum_{k=0}^{N_{L+1}} (y_k - t_k)^2 \tag{2}$$

and these weights are updated based on a gradient descent method. Here, L is the number of hidden layers, y_k is the output from the NN, and t_k is the known output used for learning.

3. Coupled Mode Theory

In this study, we consider a directional coupler in which the waveguide width and waveguide gap vary continuously in the propagation direction. According to the CMT, the evolution of the mode amplitude of each waveguide can be obtained by solving the following mode coupling equation [10]:

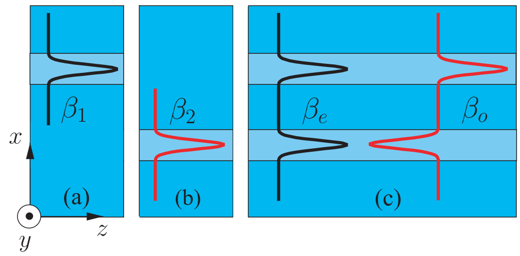


Fig. 3 Isolated and coupled modes in a directional coupler. Isolated system of (a) waveguide 1. and (b) waveguide 2. (c) Coupled system.

$$\frac{d}{dz}\begin{bmatrix} A_1(z) \\ A_2(z) \end{bmatrix} = -j\begin{bmatrix} -\delta(z) & \kappa(z) \\ \kappa(z) & \delta(z) \end{bmatrix}\begin{bmatrix} A_1(z) \\ A_2(z) \end{bmatrix} \tag{3}$$

where A_1 and A_2 represent the modal amplitude in isolated waveguides 1 and 2, and δ and κ are the propagation-constant mismatch and the mode coupling coefficient, respectively. In this study, δ and κ vary along the z-direction according to the longitudinal structural variation, and this equation is solved by a finite difference scheme in the z-direction. In the CMT analysis, it is necessary to know δ and κ according to changes in the waveguide width and gap. δ and κ can be calculated by eigenmode analysis of the waveguide at each cross section point in the z-direction. However, eigenmode analysis is usually time consuming process, especially for three-dimensional waveguides. For this reason, we employ the NN for estimation of δ and κ to improve the computational efficiency of the CMT analysis.

The propagation-constant mismatch δ is defined as follows:

$$\delta = \frac{\beta_1 - \beta_2}{2} \tag{4}$$

where β_1 and β_2 are the propagation-constants of two isolated waveguides, as shown in Fig. 3 (a) and (b).

The mode coupling coefficients which represent the coupling strength in the coupled system can be obtained by

$$\kappa = \frac{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \omega\varepsilon_0 \left(n^2 - n_1^2\right) \boldsymbol{E}_2^* \cdot \boldsymbol{E}_1 dxdy}{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\boldsymbol{E}_2^* \times \boldsymbol{H}_2 + \boldsymbol{E}_2 \times \boldsymbol{H}_2^*) \cdot \boldsymbol{i}_z dxdy} \tag{5}$$

where ω is the angular frequency, ε_0 is the dielectric constant in vacuum, n is the refractive index distribution of the coupled system, n_1 is the refractive index of the isolated system of the waveguide 1, $\boldsymbol{E}_i$ and $\boldsymbol{H}_i$ are the electric and magnetic fields of the eigenmode in the isolated system i, $\boldsymbol{i}_z$ is a unit vector in the propagation direction, and * indicates a complex conjugate.

4. Evolutionary Method

In this section, we illustrate evolutionary methods used in this study. Since evolutionary methods are based on multipoint searching algorithm, it can be expected to find more global optimal solution. In evolutionary approach, first, vectors of design parameter called individual vectors, $\boldsymbol{x}_i$ ($i \in [1, N]$) are generated, where N is population size. By iterating update procedure of the individuals, optimal point of an objective function can be found. In this paper, we adopt the hybrid firefly algorithm (HFA) [9] which is the hybrid optimization approach of the firefly algorithm (FA) and the differential evolution (DE). As an evolutionary approach, many kinds of algorithm have been proposed and applied to various kinds of design problems. Although we cannot clearly mention which algorithm is superior, it is shown that the HFA is relatively robust in the optimal design of photonic devices [1].

4.1 Differential Evolution (DE)

The DE is an algorithm that uses a difference vector to mutate an individual. The mutated individual survives if it is superior to the previous one. In the mutation, three parent individuals $\boldsymbol{x}^n_{pi}$ ($i = 1, 2, 3$) are selected randomly, and a mutant individual $\boldsymbol{x}^n_m$ of the base individual $\boldsymbol{x}^n_{p1}$ is generated using the following formula:

$$\boldsymbol{x}^n_m = \boldsymbol{x}^n_{p1} + F\left(\boldsymbol{x}^n_{p2} - \boldsymbol{x}^n_{p3}\right) \tag{6}$$

where $F \in [0, 1]$ is a scale factor. A child individual x^{n+1}_i is generated between this mutant one and x^n_i, and a superior one among x^n_i and x^{n+1}_i is left in the next generation. Since the difference vector becomes smaller on average as the search progresses, it is possible to continuously shift from a global search to a local search.

4.2 Firefly Algorithm (FA)

The FA is an algorithm inspired by a habit of fireflies. A firefly is attracted by others shining more strongly. Each firefly is attracted to all the other fireflies that are more attractive than itself. The attractive force is determined by their attractiveness and the distance from them. The update formula for the individual $\boldsymbol{x}^n_i$ is given by

$$\boldsymbol{x}^{n+1}_i = \boldsymbol{x}^n_i + \sum_j u\left(\beta_{0,j} - \beta_{0,i}\right)\beta_{0,j}e^{-\gamma r^2_{ij}}\left(\boldsymbol{x}^n_j - \boldsymbol{x}^n_i\right) + \alpha\delta^n\boldsymbol{\varepsilon} \tag{7}$$

where $\beta_{0,j}$ is the attractiveness of the j-th individual, r_{ij} is the distance between $\boldsymbol{x}^n_i$ and $\boldsymbol{x}^n_j$, $\gamma = 1/\sqrt{L}$ is the light extinction coefficient, and L is a quantity related to the search range of each individual. $u(\xi)$ is a unit step function that is 1 when $\xi \geqq 0$ and 0 when $\xi < 0$. The third term on the right-hand side of (7) is a term that adds randomness to the search, where α is a scale factor, δ is an attenuation coefficient, and $\boldsymbol{\varepsilon}$ is a vector consists of random element. Each element of $\boldsymbol{\varepsilon}$, ε_i, is chosen to be $|\varepsilon_i| < 1$. Since random elements decay continuously with iterations, it is possible to gradually switch from global search to local search.

4.3 Hybrid Firefly Algorithm (HFA)

The HFA is a hybrid method of the DE and the FA. In the HFA, the initial population is divided into half, each is evolved by the DE and the FA within one generation, and then individuals in both groups are mixed in the next generation. After that, the population is divided into two groups again and the same evolution process is repeated. It can be expected to find a better solution by utilizing the advantages utilizing the advantages of both search algorithms.

5. Numerical Examples of Analysis for Directional Coupler

To verify the effectiveness of our numerical approach based on the NN and the CMT, we compare the results obtained by the present method with those by more accurate FE-BPM [12]. We consider a directional coupler as shown in Fig. 4, in which the waveguide width, W_1 and W_2, and the waveguide gap, D, vary continuously in the propagation direction. Considering an application to silicon photonics, the material in core and cladding is assumed to be silicon (Si) and silica (SiO_2), respectively, and their refractive indices are set to be $n_{\text{core}} = 3.4$ and $n_{\text{clad}} = 1.45$, respectively. The incident wave is the fundamental TE wave with wavelength of $\lambda = 1.5 \sim 1.6\,\mu$m. The longitudinal variation of W_1, W_2, and D are defined as follows:

$$W_1(z) = W_2(z) = \begin{cases} w_0 & \text{(region 1, 3)} \\ w_1 + (w_0 - w_1)\dfrac{|2z - L|}{L - 2l} & \text{(region 2)} \end{cases}$$

$$D(z) = \begin{cases} d_0 & \text{(region 1, 3)} \\ d_1 + \dfrac{d_0 - d_1}{2}\left\{1 - \cos\left(\dfrac{\pi|2z - L|}{L - 2l}\right)\right\} & \text{(region 2)} \end{cases}$$

where the device length is set to be $L = 100\,\mu$m, $l = 5\,\mu$m, $w_0 = 0.3\,\mu$m, $w_1 = 0.5\,\mu$m, $d_0 = 1.0\,\mu$m, and $d_1 = 0.3\,\mu$m.

First, we construct an NN that outputs δ and κ for the inputs of λ, W_1, W_2, and D. The training data for the NN are created from random combinations in the range: $1.45\,\mu\text{m} \leq \lambda \leq 1.65\,\mu\text{m}$, $0.2\,\mu\text{m} \leq W_1, W_2 \leq 0.6\,\mu\text{m}$,

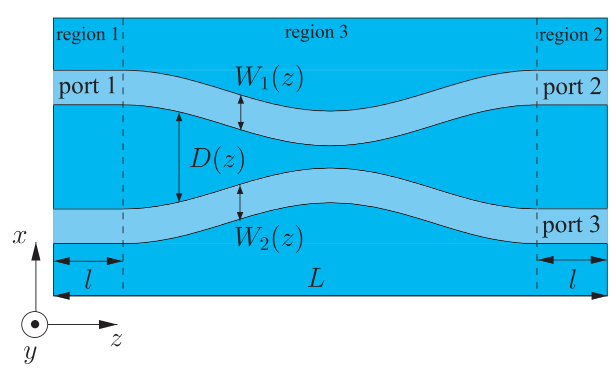
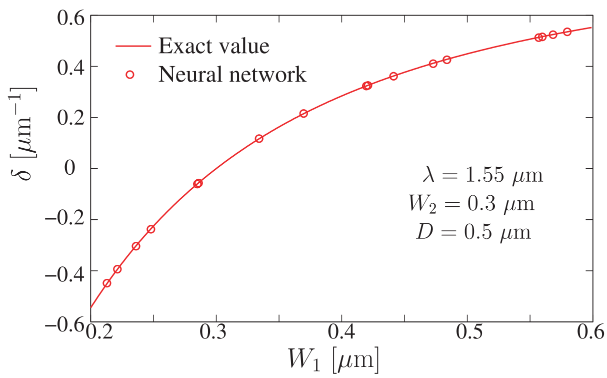
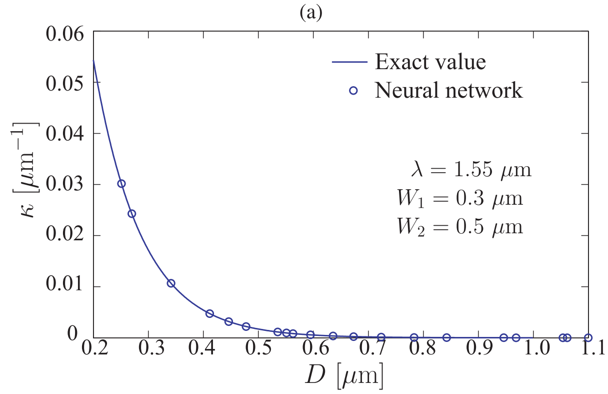


Fig. 4 Structure of a directional coupler for analysis example.



(a)



(b)

Fig. 5 Accuracy of the estimated values by the NN: (a) the propagation-constant mismatch δ and (b) the mode coupling coefficient κ.

$0.2\,\mu\mathrm{m} \leq D \leq 1.1\,\mu\mathrm{m}$. 20 test data which are created independently from the training data is used to verify the constructed NN. This NN consists of a single hidden layer with 7 units. 8,000 training data are used for machine learning, and the learning rate and the epoch are set to be 0.001 and 50,000, respectively.

Figure 5 shows the accuracy of the estimated δ and κ by the constructed NN for 20 test data. We can see that the estimated values using the NN are in good agreement with the exact values. The average relative errors are 0.47% and 0.05% for δ and κ, respectively.

Using this NN and the CMT, we analyze the directional coupler shown in Fig. 4. Figure 6 shows the accuracy of the present NN-CMT. The results by the NN-CMT are compared with those by the FE-BPM. We can see that the relative computational error is about 0.1% within the wavelength range of $\lambda = 1.5 \sim 1.6\,\mu$m and good accuracy to be used in the optimal design is obtained.

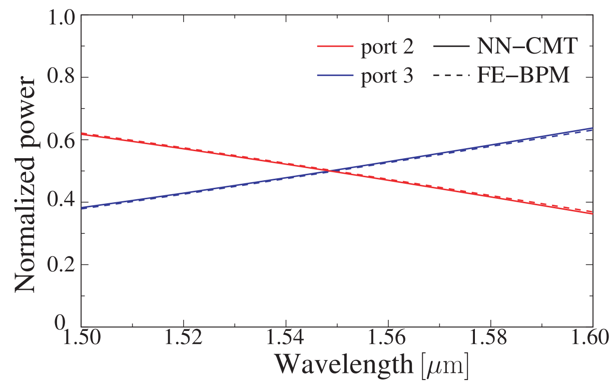
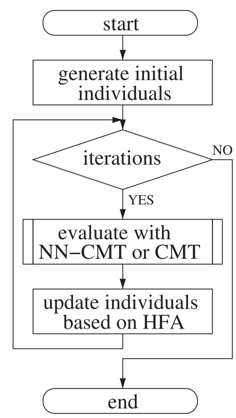
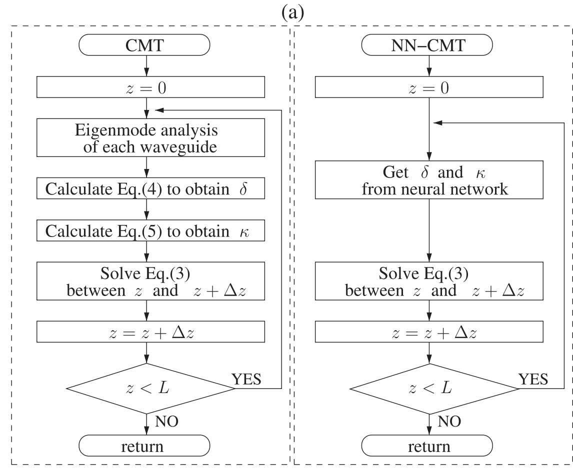


Fig. 6 Normalized power spectrum of directional coupler shown in Fig. 4.



(a)



(b)

Fig. 7 The flowchart of optimal design, (a) main routine and (b) subroutine for evaluating the characteristics of the devices utilizing the CMT and the NN-CMT.

6. Design Examples

The flowchart of the optimal design utilizing CMT and HFA is shown in Fig. 7. In this design process shown in Fig. 7(a), the evaluation of the transmission characteristics of all the given device is the heaviest task and the computational efficiency is desired to be improved. In our design approach, to evaluate transmission characteristics, CMT or NN-CMT is employed. The flowchart of the subroutines of the ordinal CMT and NN-CMT are shown in Fig. 7 (b) and 7 (c). In the ordinal CMT, in order to evaluate δ and κ, eigenmode analysis and overlap integral between the obtained eigenmodes are required at each propagation step. On the other hand, in the NN-CMT, the eigenmode analysis can be skipped and δ and κ are efficiently calculated by using the trained NN. In the NN-CMT, NN has to be trained in advance. However, the training is required only once before the entire optimization process and the design efficiency can be greatly improved. Moreover, the trained NN can be used for an optimal design of other various types of devices.

6.1 Wavelength Insensitive 3 dB Coupler

As a first design example, we consider a wavelength insensitive 3 dB coupler. Although a 3 dB coupler can be designed by a symmetric directional coupler with half coupling length, the splitting ratio strongly depends on the wavelength due to the varying of the coupling length. However, wideband operation is desired in WDM systems and several types of wavelength insensitive 3 dB coupler have been reported [10], [11].

In this example, we consider a design model as shown in Fig. 8 to design a wavelength insensitive 3 dB coupler. The parameter settings are as follows: The device length including input and output waveguides is set to be $L = 100\,\mu$m. The width, gap, and length of the input and output waveguides are $w_{\mathrm{in}} = w_{\mathrm{out}} = 0.3\,\mu$m, $d_{\mathrm{in}} = d_{\mathrm{out}} = 1.0\,\mu$m, and $l = 10\,\mu$m, respectively. The refractive indices of the core and cladding are $n_{\mathrm{core}} = 3.4$ and $n_{\mathrm{clad}} = 1.45$, respectively. The fundamental TE mode operation within $\lambda = 1.5 \sim 1.6\,\mu$m is assumed. The longitudinal structural variation is expressed using the cubic spline function determined by the waveguide width and gap at the sampled points with equal interval. The number of sampling points is set to be $N = 6$. The objective function C is defined as follows:

$$\text{Minimize } C = \sum_{\lambda \in \mathbf{\Lambda}} \left| |S_{21}(\lambda)|^2 - |S_{31}(\lambda)|^2 \right| \tag{8}$$
$$\mathbf{\Lambda} = \{1.50, 1.55, 1.60\,\mu\text{m}\}$$

where S_{21} and S_{31} are the transmission amplitude from port 1 to port 2 and port 1 to port 3, respectively. We chose three wavelength $\lambda \in \Lambda$ in order to obtain wavelength insensitive property within objective operation band. The transmission property of this device is estimated by the NN-CMT and the design variables are optimized by the HFA. The design settings in the HFA are as follows: The number of individuals is 128, the number of iterations is 1,000, and the design variables can be selected within $0.3 \leq w_{1i}, w_{2i} \leq 0.5\,\mu$m, $0.3 \leq d_i \leq 1.0\,\mu$m $(i = 1, 2, \cdots, N)$. Figure 9 shows the structure obtained by this optimization and the propagating fields at $\lambda \in \Lambda$ obtained by the FE-BPM analysis, and it is seen that each optimized structure can split the power equally into two output waveguides. A simple directional coupler with length of the half coupling length can also be used as a 3 dB coupler at a specific wavelength and the device length is considerably shorter than that of the 3 dB coupler optimized here. The propagating field in such a simple 3 dB coupler at $\lambda = 1.55\,\mu$m is shown in Fig. 10 and its structural parameters are as follows: The waveguide width, gap, and length in the coupling region is set to be $0.3\,\mu$m, $0.3\,\mu$m, and $16.2\,\mu$m, respectively. The other parameters are selected for it to operate as a 3 dB coupler at $\lambda = 1.55\,\mu$m. Figure 11 shows the wavelength dependence of the optimized

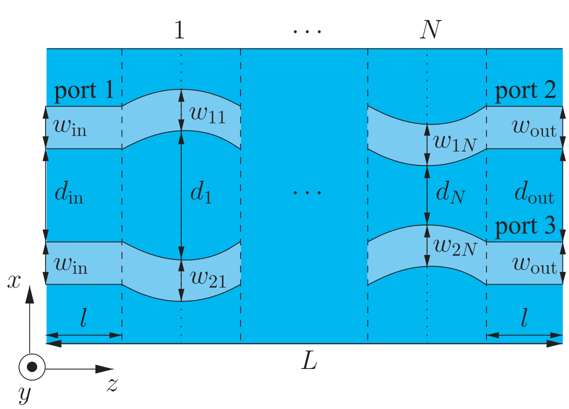


Fig. 8 Design model of a directional coupler type optical device with arbitrary longitudinal variation.

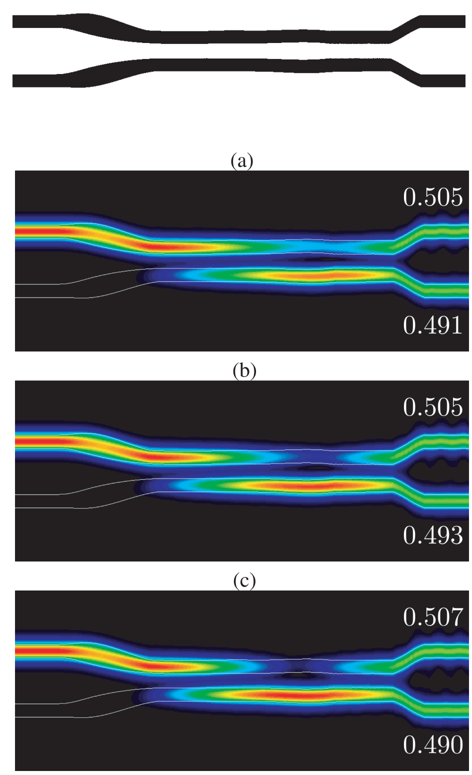


Fig. 9 Optimized results of wavelength insensitive 3 dB coupler with device length of $100\,\mu$m: (a) optimized structure, propagating field with wavelength of (b) $\lambda = 1.50\,\mu$m, (c) $\lambda = 1.55\,\mu$m, and (d) $\lambda = 1.60\,\mu$m.

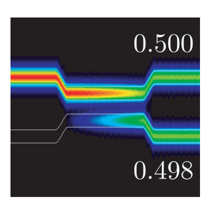


Fig. 10 Propagating field of the simple directional coupler type 3 dB coupler with device length of $42.2\,\mu$m.

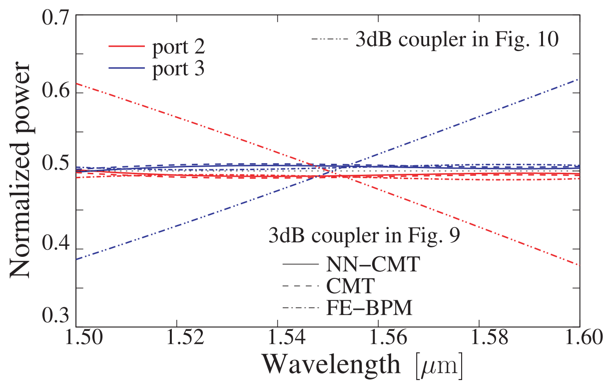


Fig. 11 Normalized power spectrum of 3 dB couplers shown in Fig. 10 and Fig. 9.

3 dB coupler and simple one. For the comparison of the numerical accuracy, the results calculated by the present NN-CMT, the conventional CMT, and the FE-BPM are shown for the optimized 3 dB coupler. We can see that these results are in good agreement. From this result, we can also see that the optimized device can be used as a 3 dB coupler in the wider wavelength range compared with the simple directional coupler.

6.2 Wave length Insensitive 1 : 2 Power Splitter

As a second design example, we consider a wavelength insensitive 1 : 2 power splitter. The design problem settings are the same in the previous subsection except for the splitting ratio and the device length. The device length L is set to be 150 μm and the objective function is defined as follows:

$$\text{Minimize } C = \sum_{\lambda \in \mathbf{\Lambda}} \left| 2 \times |S_{21}(\lambda)|^2 - |S_{31}(\lambda)|^2 \right| \tag{9}$$

$$\mathbf{\Lambda} = \{1.50, 1.55, 1.60\,\mu\text{m}\}$$

Figure 12 shows the optimized structure obtained by our design approach and the propagating fields at $\lambda \in \Lambda$ obtained by the FE-BPM analysis. We can see that the optimized device works as a 1 : 2 power splitter at the designed wavelengths. A simple directional coupler with appropriate device length can also be used as a 1 : 2 coupler at specific wavelength. The propagating field in this simple 1 : 2 coupler at $\lambda = 1.55\,\mu$m is shown in Fig. 13 and its structural parameters are as follows: The waveguide width, gap, and length in the coupling region is set to be 0.3 μm, 0.3 μm, 20.1 μm, respectively. The other parameters are selected for it to operate as a 1 : 2 power splitter at $\lambda = 1.55\,\mu$m. Figure 14 shows the wavelength dependence of the optimized 1 : 2 power splitter and the simple directional coupler based one. For the optimized devices, the results obtained by the NN-CMT, the conventional CMT, and the FE-BPM are compared. We can see that these results are again in good agreement. While the simple directional coupler based 1 : 2 power splitter has strong wavelength dependence, our optimized device overcomes this defect and shows good transparency.

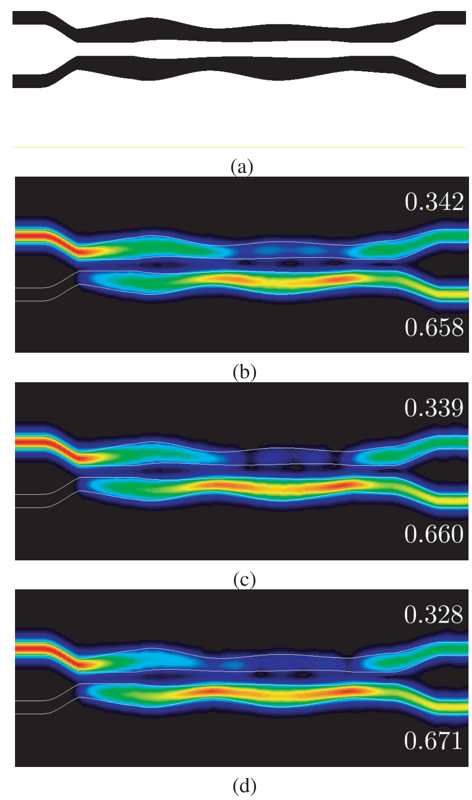


Fig. 12 Optimized results of wavelength insensitive 1 : 2 power splitter with device length of 150 μm: (a) optimized structure, propagating field with wavelength of (b) $\lambda = 1.50\,\mu$m, (c) $\lambda = 1.55\,\mu$m, and (d) $\lambda = 1.60\,\mu$m.

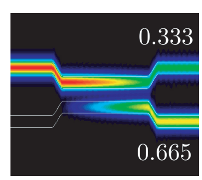


Fig. 13 Propagating field of the simple directional coupler type 1 : 2 power splitter with device length of 46.1 μm.

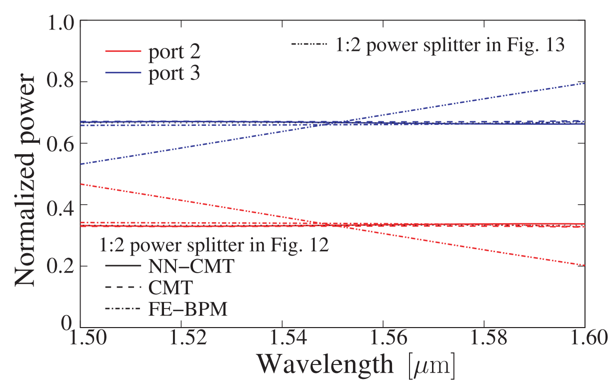


Fig. 14 Normalized spectrum of 1 : 2 power splitters shown in Fig. 13 and Fig. 12 (a).

6.3 Wavelength Demultiplexer

As a last design example, considering a wavelength dependent device, we design wavelength demultiplexer splitting light with $\lambda = 1.52\,\mu$m and $\lambda = 1.58\,\mu$m. The design problem settings are same in the previous subsection except for the device length, and the search range of the design variables. The device length L is set to be $200\,\mu$m and the search range of the waveguide width and gap are set to be $0.25 \leq w_{1i}, w_{2i} \leq 0.5\,\mu$m, $0.25 \leq d_i \leq 1.0\,\mu$m ($i = 1, 2, \cdots, N$). The objective function is defined as follows:

$$\text{Minimize } C = \left|1 - |S_{21}(\lambda_1)|^2\right| + \left|1 - |S_{31}(\lambda_2)|^2\right| \quad (10)$$

where $\lambda_1 = 1.52\,\mu$m and $\lambda_2 = 1.58\,\mu$m. The device is expected to be optimized so that the lights with λ_1 and λ_2 transmit from port 1 to port 2 and port 3, respectively. Figure 15 shows the structure obtained by this optimization and the propagating field obtained by FE-BPM analysis at the wavelength of λ_1 and λ_2. It is seen that wavelength demultiplexing between λ_1 and λ_2 is achieved in the optimized device. Figure 16 shows the wavelength dependence of the optimized device, and the results obtained by the NN-CMT, the conventional CMT, and the FE-BPM are compared. Although these analysis results by different numerical methods are slightly different from each other because of the error accumulation due to longer device length, we can see that the designed device can be used as a wavelength demultiplexer even based on the result by the FE-BPM.

As shown in these design examples, the NN constructed in advance can be reused for design of various directional coupler type optical devices. Since the NN-CMT can skip an eigenmode analysis, it can reduce run time of optimization. It is expected that the reduction is significant especially for 3D design problems.

In order to confirm the efficiency of our design approach using NN-CMT, we compare computational time for analyzing a directional coupler by CMT and NN-CMT because time required in optimization by HFA is almost negligible in the whole optimization process. As a numerical example, we consider a directional coupler with device length of $100\,\mu$m and discretize along propagating direction with propagation step size of $\Delta z = 0.1\,\mu$m. The average computational time of 1,000 times calculations are 17 seconds and 4 seconds in CMT and NN-CMT analysis, respectively, using a PC with a Intel® Core™ i7-3770 3.40 GHz processer. All the programs are written in C-language by ourselves. In the present optimization, $128 \times 1{,}000$ times waveguide evaluation is required, then CMT takes 2,176 seconds for evaluating transmission property. On the other hand, in NN-CMT, although the additional computational time for training NN is required and it takes 365 seconds, the total time including training NN and evaluating transmission property is 877 seconds and the computational time of NN-CMT is reduced to 40% compared with CMT. The training of NN is required only once even if we design a lot of different devices. The computational time of NN-CMT without the process is reduced to 24% compared with CMT. It is expected that this reduction rate becomes greatly higher in a design of more practical three-dimensional devices because a computational cost of eigenmode analysis is greatly higher in three-dimensional waveguide analysis.

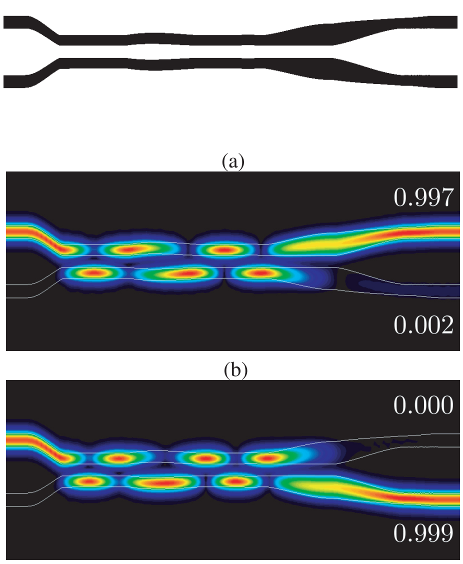


Fig. 15 Optimized results of wavelength demultiplexer: (a) optimized structure, propagating field of (b) $\lambda = 1.52\,\mu$m, and (c) $\lambda = 1.58\,\mu$m.

7. Conclusion

In this paper, we proposed the NN-CMT as a new approach to improve the computational efficiency of optical analysis design. In the proposed the NN-CMT, the mode coupling coefficient and the propagation-constant mismatch are efficiently estimated by the NN without requiring modal analysis. Utilizing this NN-CMT and the HFA, we showed the design examples of a wavelength insensitive 3 dB coupler, a wavelength insensitive 1 : 2 power splitter, and a wavelength demultiplexer and verified the effectiveness of our

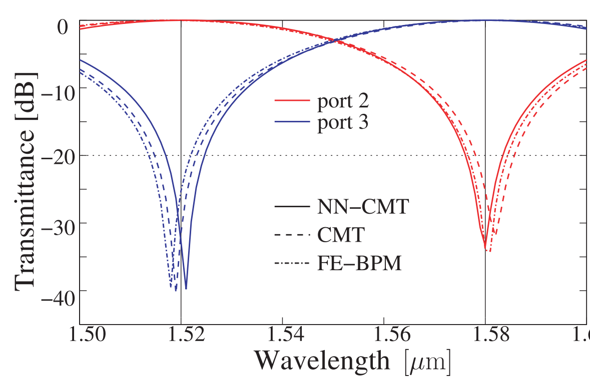


Fig. 16 Transmittance spectrum of wavelength demultiplexer shown in Fig. 15 (a).

design approach. The computational efficiency is demonstrated better than optimal design utilizing the CMT. The extension of this design approach to 3D design problems is now under consideration.

Acknowledgments

This work was supported by JSPS (Japan) KAKENHI Grant Number 18K04276.

References

[1] A. Koda, K. Morimoto, and Y. Tsuji, "A study on topology optimization of plasmonic waveguide devices using function expansion method and evolutionary approach," J. Lightw. Technol., vol.37, no.3, pp.981–988, Feb. 2019.
[2] M.H. Taherslma, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons, "Deep neural network inverse design of integrated nanophotonic devices," Sci. Reports, vol.9, #1386, Feb. 2019.
[3] T. Zhang, J. Wang, Q. Liu, J. Zhou, J. Dai, X. Han, Y. Zhou, and K. Xu, "Efficient spectrum prediction and inverse design for plasmonic waveguide sytems based on artificial neural netowork," Photonics Research, vol.7, no.3, pp.368–380, March 2019.
[4] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," ACS Photon., 5, pp.1365–1369, Feb. 2018.
[5] E. Bor, O. Alparslan, M. Turduev, Y.S. Hanay, H. Kurt, S. Arakawa, and M. Murata, "Integrated silicon photonics device design by attractor selection mechanism based on artificial neural networks: Optical coupler and asymmetric light transmitter," Opt. Exp., vol.26, no.22, pp.29032–29044, Oct. 2018.
[6] A.M. Gabr, C. Featherston, C. Zhang, C. Bonfil, Q.-J. Zhang, and T.J. Smy, "Design and optimization of optical passive elements using artificial neural networks," J. Opt. Soc. Amer. B, vol.36, no.4, pp.999–1007, April 2019.
[7] A. Yariv, "Coupled-mode theory for guided-wave optics," IEEE J. Quantum Electron., vol.9, no.9, pp.919–933, Sept. 1973.
[8] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning representations by back-propagating errors," Nature, vol.323, no.9, pp.533–536, Oct. 1986.
[9] L. Zhang, L. Liu, X.-S. Yang, and Y. Dai, "A novel hybrid firefly algorithm for global optimization," PLoS ONE, vol.11, pp.1–17, Sept. 2016.
[10] D. Guo and T. Cho, "Compact broadband silicon 3 dB coupler based on shortcuts to adiabaticity," Opt. Lett., vol.43, no.19, pp.4795–4798, Sept. 2018.
[11] A. Takagi, K. Jinguji, and M. Kawachi, "Silica-based waveguide-type wavelength-insensitive couplers (WINC's) with series-tapered coupling structure," J. Lightw. Technol., vol.10, no.12, pp.1814–1824, Dec. 1992.
[12] Y. Tsuji and M. Koshiba, "A finite element beam propagation method for strongly guiding and longitudinally varying optical waveguides," J. Lightw. Technol., vol.14, no.2, pp.217–222, Feb. 1996.

Koji Kudo received the B.S. degree in department of information and electrical engineering from Muroran Institute of technology, Muroran, Japan, in 2019. He is presently working toward the M.S, degree in Information and Electronic Engineering from Muroran Institute of Technology. Mr. Kudo is a student member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Keita Morimoto received the B.S. and M.S. degrees in information and electronic engineering from the Muroran Institute of Technology, Muroran, Japan, in 2017 and 2019, respectively, where he is currently pursuing the Ph.D. degree in information and electronic engineering. Mr. Morimoto is a Student Member of the Institute of Electronics, Information and Communication Engineers (IEICE) and IEEE.

Akito Iguchi received the B.S., M.S., and Ph.D. degrees in electronic engineering from Muroran Institute of Technology, Muroran, Japan, in 2015, 2017, and 2019. From 2019 to 2020, he was a Post-Doctoral Research Fellow of Japan Society for the Promotion of Science (JSPS). He is currently an Assistant Professor at Muroran Institute of Technology. Dr. Iguchi is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and IEEE.

Yasuhide Tsuji received the B.S., M.S., and Ph.D. degrees in electronic engineering from Hokkaido University, Sapporo, Japan, in 1991, 1993, and 1996, respectively. In 1996, he joined the Department of Applied Electronic Engineering, Hokkaido Institute of Technology, Sapporo, Japan. From 1997 to 2004, he was an Associate Professor of Electronics and Information Engineering at Hokkaido University. From 2004 to 2011, he was an Associate Professor of Electrical and Electronic Engineering at Kitami Institute of Technology, Kitami, Japan. Since 2011, he has been a Professor of Information and Electronic Engineering at Muroran Institute of Technology, Muroran, Japan. He has been interested in wave electronics. Dr. Tsuji is a Senior Member of the Institute of Electronics, Information and Communication Engineers (IEICE), and the IEEE, and a Member of the Japan Society of Applied Physics and the Optical Society of America (OSA). In 1997, 1999, and 2019, he was awarded the Best Paper Award from IEICE. In 2000, he has received the Third Millennium Medal from IEEE. In 2019, he has received the IEEE Photonics Technology Letters Outstanding Reviewer Award.