

11-16-2004

Generalized Design of Diffractive Optical Elements Using Neural Networks

Ajay Pasupuleti

Rochester Institute of Technology

Anand Gopalan

Rochester Institute of Technology

Ferat Sahin

Rochester Institute of Technology

Mustafa A.G. Abushagur

Rochester Institute of Technology

Follow this and additional works at: <https://scholarworks.rit.edu/other>

Recommended Citation

Ajay Pasupuleti, Anand Gopalan, Ferat Sahin, Mustafa A. G. Abushagur, "Generalized design of diffractive optical elements using neural networks", Proc. SPIE 5579, Photonics North 2004: Photonic Applications in Telecommunications, Sensors, Software, and Lasers, (16 November 2004); doi: 10.1117/12.567191; <https://doi.org/10.1117/12.567191>

This Conference Paper is brought to you for free and open access by the Faculty & Staff Scholarship at RIT Scholar Works. It has been accepted for inclusion in Presentations and other scholarship by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.

Copyright 2004 Society of Photo-Optical Instrumentation Engineers.

This paper was published in The Journal of Electronic Imaging and is made available as an electronic reprint (preprint) with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Generalized design of Diffractive Optical Elements using Neural Networks

A. Pasupuleti, A. Gopalan, F. Sahin and M. A. G. Abushagur
Rochester Institute of Technology
Rochester, New York, USA
axp1014@rit.edu

Abstract

Diffractive optical elements (DOE) utilize diffraction to manipulate light in optical systems. These elements have a wide range of applications including optical interconnects, coherent beam addition, laser beam shaping and refractive optics aberration correction. Due to the wide range of applications, optimal design of DOE has become an important research problem. In the design of the DOEs, existing techniques utilize the Fresnel diffraction theory to compute the phase at the desired location at the output plane. This process involves solving nonlinear integral equations for which various numerical methods along with robust optimization algorithms exist in literature. However all the algorithms proposed so far assume that the size and the spacing of the elements as independent variables in the design of optimal diffractive gratings. Therefore search algorithms need to be called every time the required geometry of the elements changes, resulting in a computationally expensive design procedure for systems utilizing a large number of DOEs.

In this work we have developed a novel algorithm that uses neural networks with possibly multiple hidden layers to overcome this limitation and arrives at a general solution for the design of the DOEs for a given application. Inputs to this network are the spacing between the elements and the input/output planes. The network outputs the phase gratings that are required to obtain the desired intensity at the specified location in the output plane. The network was trained using the back-propagation technique. The training set was generated by using GS algorithm approach as described in literature. The mean square error obtained is comparable to conventional techniques but with much lower computational costs.

Keywords: DOE design, neural networks, back propagation

1. Introduction

Over the past decade with the advent of high resolution multimedia streaming, Voice over IP and other such technologies, the demand for data communication bandwidth has increased exponentially. In 2001, it was predicted that the load on the Internet backbone would be as high as 11 Tb/s [1]. Current estimates have already tripled this number, and the estimated load by 2005 is now 35 Tb/s with estimates being revised each year [2]. This is in addition to ever-increasing telephony traffic. The optic fiber has firmly established itself as the medium of choice for long haul transmission, as well as for backbone communications because of its high transmission bandwidth, low error rates and high reliability [1].

Due to this there is a considerable interest in developing technologies that enable the increase of the data handling capacity of optical networks. In lieu, of these developments there is an increasing need to develop optical switches that can switch between larger numbers of fibers. Several techniques have been used to design and manufacture optical switches ranging from technologies based on reflection like MEMS mirror technologies such as the DMD to those based on diffraction such as diffractive optical elements (DOEs) [3]. In the past few years the use of DOEs in the design of optical switches along with many other optical applications such as beam splitting has gained a lot of attention, due to the ease of manufacture using traditional semiconductor fabrication processes.

This is especially true in the field of space variant optical interconnects (SVOIs) which are used to manufacture optical switches [4]. A major drawback to the use of DOEs has been the design complexity associated with designing large scale systems with several DOEs. This is because the design of a single DOE is in itself a rather complex

optimization problem, which requires significant computation to reach an efficient solution. This problem is exacerbated when the system involves several DOEs with varying requirements.

The current work proposes a technique for DOE design that represents extremely low computational overheads in comparison to existing techniques, but yet produces a solution that is very close to the global minimum, and hence has very low losses. The technique uses a neural network to develop a generalized solution for the DOE design for a given application. This paper presents a proof of concept for this technique in solving the DOE design problem for the purpose of space variant optical interconnects. After the neural network has been trained, the only inputs to the neural network are four distances that uniquely define the geometry of the particular interconnect to be designed. Since the neural network has already been trained, it uses these inputs to simply produce an output phase matrix for the DOE to be designed with almost no computational overhead. This would drastically alleviate the design process, especially in the case of optical systems with a large number of DOEs. The current work provides evidence that the neural network can indeed arrive at a general solution for the DOE design and in some cases produce better results than those obtained using standard error reduction techniques.

This paper is organized as follows. In section 2 some of the prior techniques for the design of DOEs are described. In section 3 the proposed technique is illustrated followed by a discussion on the simulation setup. Finally results are discussed in section 4.

2. Prior Techniques for DOE Design

Several techniques have been used in the past for the purpose of DOE design in the SVOI problem. Since the resultant DOE design is a recognized optimization problem, most of these techniques are based on the utilization of some technique to optimize the phase of the grating for a maximum power throughput in the system. Although several of these techniques find applications in other classes of the DOE design problem, we will focus on the aspects that deal with the SVOI problem, since this is the application that we have chosen as a proof of concept for our method.

2.1 Linear Error Reduction using Gerchberg Saxton Algorithm

A linear error reduction technique that was initially used for DOE design in the SVOI problem is the Gerchberg Saxton Algorithm. In the Single Gerchberg-Saxton (GS), the first DOE is designed to maximize illumination of the second DOE, and then the second is designed to focus the light beam into the second fiber. [5] In another method known as the Repetitive GS algorithm, the DOEs are designed using the standard GS algorithm, and then an amplitude constraint is applied for a second pass of the algorithm to design the second DOE [5]. The most commonly observed variant of the GS algorithm seen in literature is the coupled GS algorithm [5, 6]. This is the algorithm that has been used to generate the initial training sets for the neural network described in the current work. Of all the standard linear error reduction techniques mentioned above, the coupled GS algorithm performs the best in terms of the power throughput of the optical switch system [5]. Figure 1 shows a flowchart of the application of the coupled GS algorithm to calculate the phase of two DOE gratings used in the SVOI problem.

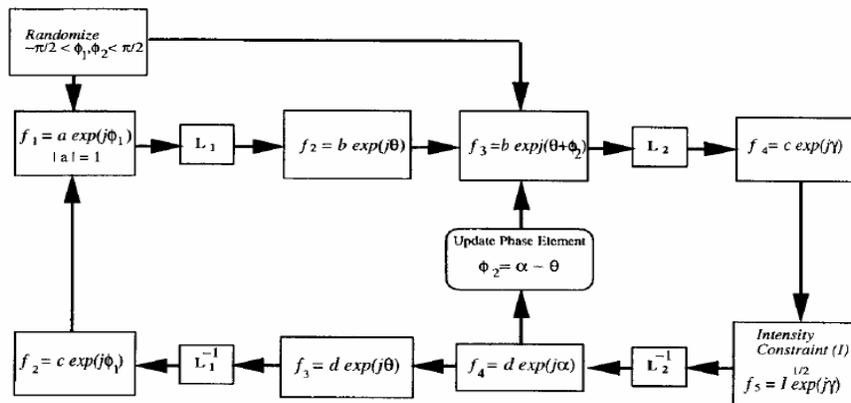


Figure 1: Coupled GS Algorithm for DOE phase calculation [5]

However all linear error reduction techniques suffer from the tendency to get stuck in a local minimum solution and not converge on the global minimum solution. The algorithms discussed above also do not provide any mechanism to move out of this local minimum. In order to overcome the local minima problem, non linear optimization techniques such as the genetic algorithm have been used.

2.2 Genetic Algorithm

The logical choice of non-linear optimization algorithm to solve the SVOI problem is the genetic algorithm [7]. Here the phase grating is quantized into 8, 16 or 32 levels. Each element of the phase grating is then treated as a gene and a standard GA procedure is applied involving optimally selected crossover techniques and mutations [7]. This technique has been found to yield extremely good results, with very low losses in the resultant optical link. However, the major drawback of the genetic algorithm is the massive amount of computation power and time required to arrive at a solution for a single problem. As the size and consequently number of elements in the phase grating increases the computational complexity increases greatly. Therefore although the genetic algorithm converges eventually to a global minimum solution, it does so with great computational cost. Therefore there has been considerable interest in the research community towards developing a method that is computationally less intensive, but can still converge close to a global minimum solution for the DOE design problem.

2.3 Micro-Genetic Algorithm

The micro-genetic algorithm is one such attempt to arrive at a global solution with relatively lower computational overheads than the GA [8]. The micro-GA algorithm advocates the use of merely five starting parents in any generation. Generally a minimum of two points are used to cross over the genes to come up with the next generation. Again, in the next generation only five total genes are retained and so on. The concepts of fitness function and mutation are similar to the corresponding GA implementation. Although this technique holds promise according to [8, 9], and converges to solutions with much lower power losses, the technique is very sensitive to the process of parent selection. If the right combination of cross over and selection criteria are not used, the technique can often get lodged in a local minimum.

3. Proposed Technique

All existing techniques for solving the DOE design problem treat it as a standard optimization problem and apply various linear and non-linear search techniques to arrive at a global minimum solution. In order to arrive at a solution close to the global minimum these techniques have to endure massive computational cost and complexity. The current technique proposes to arrive at a similar solution close to the global minimum by using the generalization property of a standard single hidden layer neural network [10].

One of the greatest advantages of using a neural network to solve a class of problems is its ability to learn a particular solution to a problem and then apply it towards finding a general solution to the entire class of problems. In the current work, the same property has been applied to train a single hidden layer neural network to arrive at a generalized solution for the DOE problem. This solution has the potential of creating DOEs with high diffraction efficiencies as will be demonstrated by the results in Section 4.

Let us consider a standard SVOI design setup as shown in Figure 2. The light from fiber I_1 needs to be focused using DOE 1 & 2 onto fiber O_2 . For a fixed fiber and DOE size as dictated by the application and the limits of fabrication, the entire setup in Figure 2 can be completely described by five distances X_1, X_2, X_3, X_4 and bias. The bias is used as an additional control parameter to ensure proper functioning of the network. Therefore, for a given set of fiber and beam characteristics, the entire DOE design is uniquely defined by these four distances. Therefore these five quantities (four inputs and bias) form the inputs to the proposed neural network. The outputs of the neural network are the phase matrices corresponding to DOE 1 and 2 such that the power entering the fiber is maximized.

Figure 3 illustrates the structure of the neural network as defined by the above considerations that has been used for the purpose. As discussed earlier the network has only five input nodes which store the real values corresponding to the five distances X_1, X_2, X_3, X_4 and bias. The hidden layer consists of 20×20 hidden nodes, while the output layer has 50×50 output nodes corresponding to the 2500 elements that define the phase grating. The size of the hidden layer is determined by the number of elements in the desired phase grating. This is due to the fact, that as the number of phase

grating increases, the amount of information that the network needs to learn also increases substantially. The weights of the network are real quantities and the output nodes store quantized phase values.

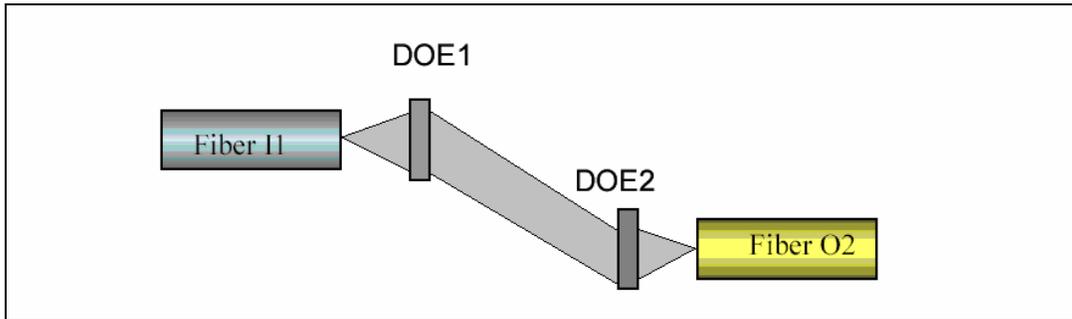


Figure: 2 Space Variant Optical Interconnect (SVOI) design setup

We have used a modified version of the standard back propagation algorithm to train the network [10]. Weights are initialized at random between -1 and +1 using a uniform distribution. The network is then presented with the training vector comprising of the five distances. The intermediate outputs of the hidden layer are obtained using the sigmoid function. The output of the neural network is the quantized phase matrix. The phase values are quantized since; this form is needed anyway in order to fabricate the actual DOE.

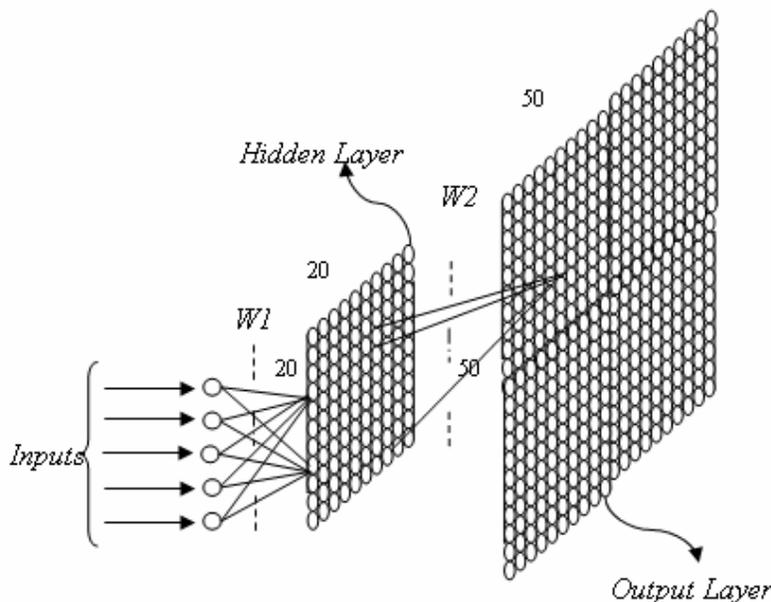


Figure 3: Schematic structure of the proposed neural network

The quantization of phase also incidentally reduces the amount of information that the neural network has to learn. This matrix is then used to compute the resultant intensity at the output fiber using standard Fresnel diffraction formulae [5]. The phase matrix is then compared with the desired phase matrix that would result in a maximum amount of power focused at the output beam. The difference between the resultant phase and the desired phase is used to generate an error signal using the formula

$$\epsilon_s = (D_p - Cpo) \times \sigma'(O) \tag{1}$$

Where ϵ_s is the error signal, $\sigma'(O)$ is the output of the neurons in the previous stage passed through the derivative of the standard sigmoid function, D_p is the desired phase and Cpo is the current phase output. This error signal is used to readjust the weights in a back propagation like manner using the formula

$$\Delta W = \alpha \times \varepsilon_s \times O \quad (2)$$

Where ΔW is the change in the weights, α is the learning rate and output is the output signal of the neurons in the previous stage. The learning rate alpha is adjusted using an exponential decay. Alpha is decayed drastically in the first 1000 iterations after which it is held at a more or less constant value. Figure 4 shows a pseudo code algorithm for the training procedure [10].

4. Results and discussion

The first step in the design of the optical switches is to calculate the spatial distances between the various elements of the switch in order to minimize losses. The loss in the system occurs due to various factors such as number of side lobes captured by the DOEs, diffraction efficiency, loss due to spatial misalignment and scattering losses [5]. The computed spatial distances can be transformed into the mathematical model by setting up an appropriate coordinate system. In this design, as illustrated in Figure 5, a symmetric coordinate system with its origin at the center of the input fiber surface was used because it was found that the algorithm converged faster in a symmetric coordinate frame as opposed to an asymmetric coordinate frame.

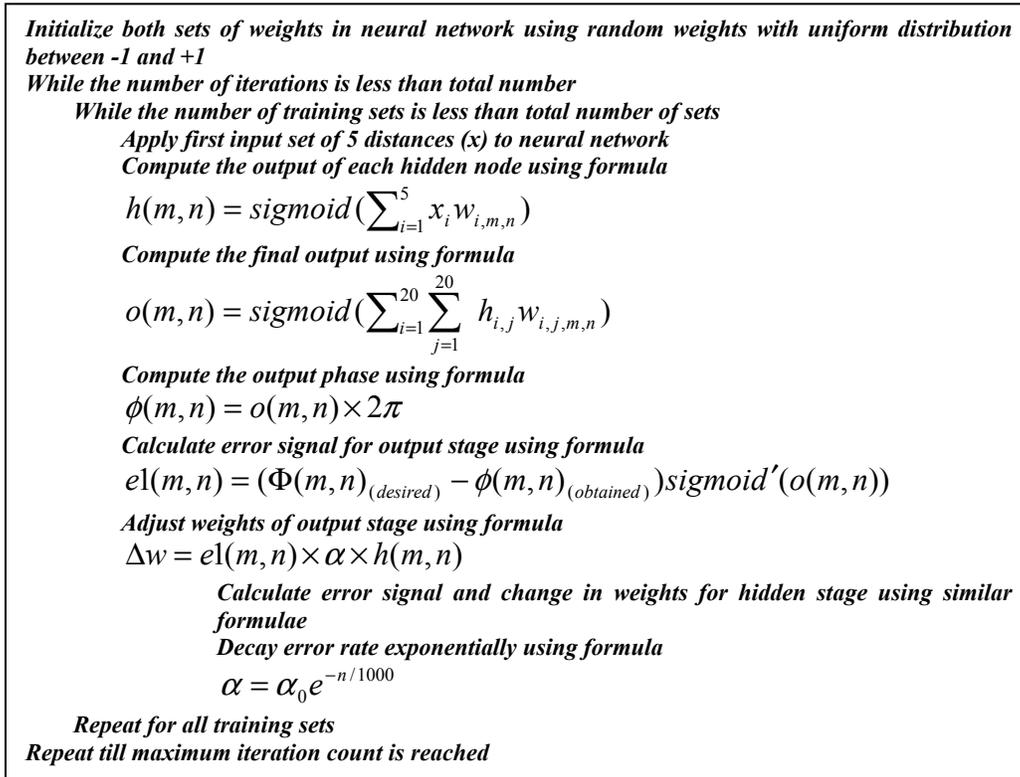


Figure 4: Algorithm for training the neural network

The above configuration of the input and the output fibers could result in four possible coupling configurations of which only one coupling configuration (input fiber I_2 to output fiber O_1) was considered for the proof of concept. Simulation results of the DOE phase as computed by the generalized neural network are presented. The proposed neural network was trained by two training vectors consisting of 12 data sets and 50 data sets. The mean convergence plots for these training vectors is illustrated in Figure 6.

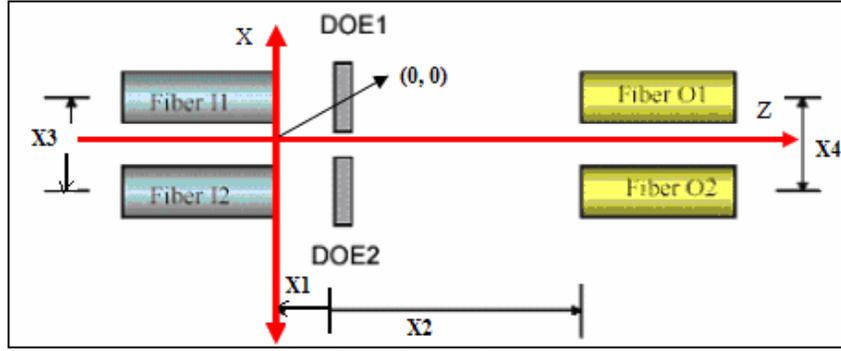


Figure 5: Simulation setup for the proposed technique

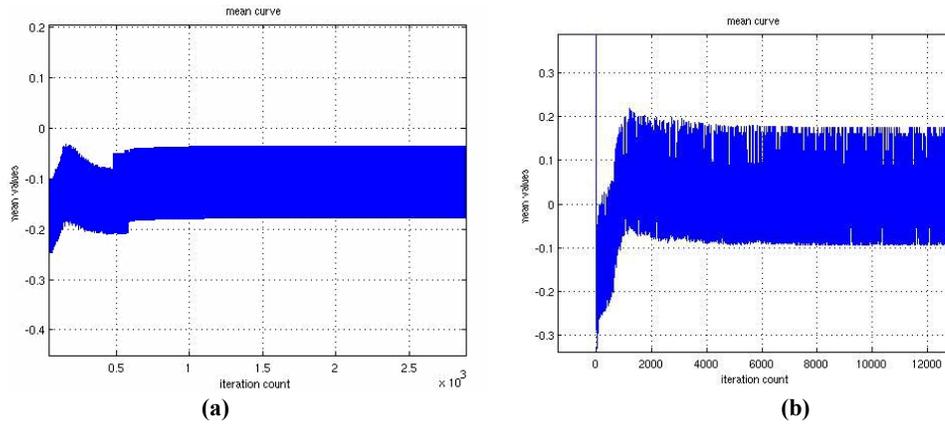


Figure 6: Mean convergence plot for training vector with (a) 12 data sets and (b) 50 data sets

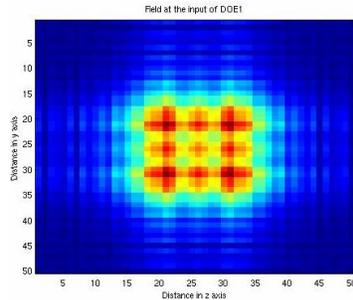
The above graphs indicate that the neural network takes only about 2000 iterations to converge for 50 data sets as opposed to 6000 iterations for 12 datasets. From this observation it can be inferred that as the size of the training vector increases the network learns in fewer iterations indicating that there is high correlation between the training sets [9]. However, the amplitude of the mean convergence curve for the training vector consisting of 12 datasets is relatively small when compared to that of the training vector consisting of 50 datasets. This indicates that the network when trained with fewer datasets is able to learn the trained set well but may suffer from a loss of generalization. On the other hand, when the network is trained with 50 datasets, it tends to generalize the DOE phase that results in a relative change in output power.

Table 1: Performance of the neural network with respect to GS Algorithm

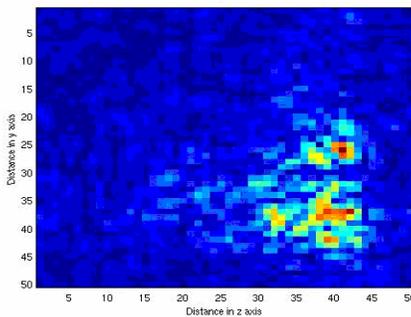
No.	Test sample	Distances (m)				MSE (%)	
		X1	X2	X3	X4	GS	NN
1	1	0.00012	0.00188	0.0005	0.0005	33.05	9.77
2	8	0.00026	0.00174	0.0005	0.0005	28.97	49.50
3	33	0.00016	0.00184	0.0004	0.0003	30.33	11.83
4	44	0.00018	0.00182	0.00025	0.00035	27.40	25.02
5	50	0.0003	0.0017	0.00025	0.00035	20.82	57.54

As described before, the neural network is trained by using the phase values computed by the GS algorithm, thus the performance of the trained network is compared to that of GS Algorithm. Table 1 illustrates the mean square error (MSE) of the output power in the region of interest for GS as well as the NN algorithms. Analyzing the results in Table 1 illustrate that the network is able to learn correlation between the various inputs and improve on the performance of some of the vectors in the training set. Figures 7(a)-7(g) illustrate the input power at the DOE along with the output power computed by GS and NN algorithms for some of the training. Figures 7(b) and 7(c) correspond to the output power as computed GS Algorithm and Neural network respectively for the first training vector. Likewise, Figures 7(d)

and 7(e) correspond to 44th training vector and 7(f) and 7(g) correspond to the 50th training vector. These cases illustrate the three possible situations which are GS outperforms NN (Figures 7(b) and 7(c)), the performance of the two algorithms is comparable (Figures 7(d) and 7(e)), and NN outperforms GS (Figures 7(f) and 7(g)).

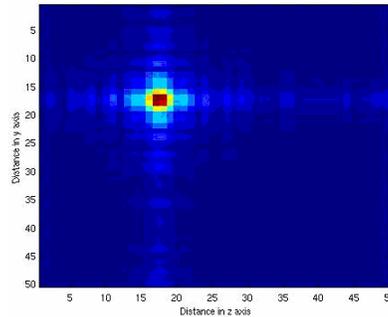


(a) Input field at the DOE



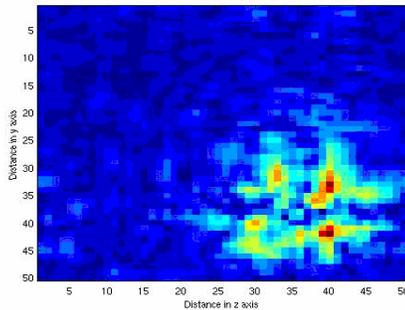
(b)

Output Power by GS algorithm for Test sample #1



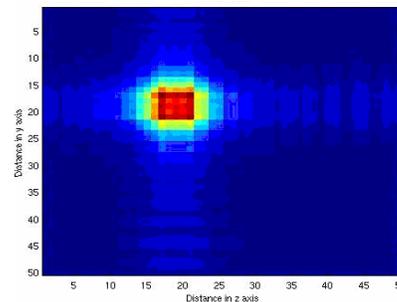
(c)

Output Power by NN algorithm for Test sample #1



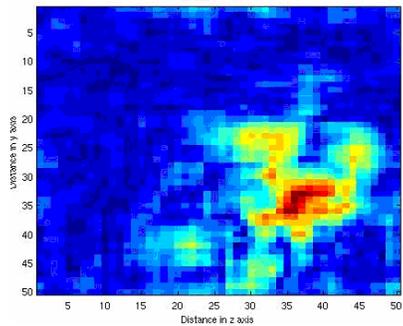
(d)

Output Power by GS algorithm for Test sample #44



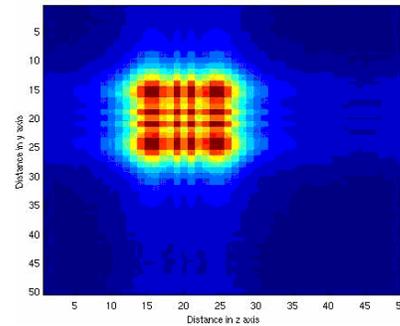
(e)

Output Power by NN algorithm for Test sample #44



(f)

Output Power by GS algorithm for Test sample #50



(g)

Output Power by NN algorithm for Test sample #50

Figure 7: Performance comparison of GS Algorithm and Neural network for some of the training vectors

5. Applications

At this time, two possible applications are envisioned for the proposed technique. This system can greatly reduce the design time for multiple input multiple output space variant optical interconnects. As discussed in section 2, existing algorithms can compute optimal phase values for the coupled DOEs for a given coupling configuration. However, all these algorithms are computationally intensive and time consuming processes that result in long design time as they need to be executed for each and every coupling configuration. The proposed technique reduces the design time by learning from a few training vectors and gaining the ability to generalize for the entire data set.

The other application for this work is in correcting the phase values for the fabricated DOEs. Due to the alignment and other fabrication issues, the distances between the fibers or the DOEs could be different from the designed values. This error could result in the drop of performance of the system. To overcome this difficulty, an external module containing the proposed algorithm could be used to compute the actual phase values for the fabricated DOEs and thus correct the phase.

6. Conclusion and future work

A neural network based algorithm for the design of diffractive optical elements for a space variant optical interconnect system was proposed and its performance was compared with the existing GS algorithm. The proposed technique is found to have extremely low computational overheads yet produce solutions that are extremely close to the global minimum, and hence has very low losses. Two possible applications have been suggested for this work. Extension of this work could include designing a neural network with more input parameters such as dimensions of DOE (as opposed to fixed value used in this work) as well as the beam collimator (present at the output of the fiber) and material properties of the DOE and the rest of the system. Including these parameters could reduce the losses, thus increasing the performance of the system.

7. References

- [1] Bourouha M A, Bataineh M, Guizani M, "Advances in optical switching and networking: past, present and future," *Proc. IEEE Southeastcon*, 2002, pp.405-413.
- [2] Probe research Inc, <http://www.proberesearch.com>
- [3] Dobbelaere P, "Advances in integrated 2D MEMS-Based Solutions for optical network applications," *Proc. IEEE optical communications*, 2003, p 16-23.
- [4] Kusko M, Cojoc D, Apostol D, Muller R, Manea E, Podaru C, "Design and fabrication of diffractive optical elements," *Proc. IEEE semiconductors conference*, 2003, p 170-174.
- [5] Zaleta D, Larsson M, Daschner W, Lee S H, "Design methods for space-variant optical interconnections to achieve optimum power throughput," *Applied Optics*, v 34, n 14, May 10, 1995, p 2436-2447.
- [6] Urquhart K S, Marchand P, Fainman Y, Lee S H, "Diffractive optics applied to free space optical interconnects," *Applied Optics*, Vol. 33, 1994, p 3670-3682.
- [7] Yoon J S, Kim N, Jeon S H, Suh H H, "Real-time optical interconnection using gratings applied by simple genetic algorithm," *CLEO/Pacific Rim '99*, 1999, p 397-398.
- [8] Johnson E G, Abushagur, M AG, "Microgenetic-algorithm optimization methods applied to dielectric gratings," *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, v 12, n 5, May, 1995, p 1152-1160.
- [9] Krishnakumar K, "Micro- genetic algorithms for stationary and non-stationary function optimization," *Intelligent Control and Adaptive Systems*, Vol. 1196, 1989, p 289- 296.
- [10] Haykin S, *Neural Networks A Comprehensive Foundation*, Prentice-Hall, 1994.