An Adaptive Multilayer Optical Neural Network Design

I. Saxena and E. Fiesler

Institut Dalle Molle d'Intelligence Artificielle Perceptive (IDIAP)
Case postale 609, CH-1920 Martigny, Switzerland
Electronic mail (Internet): ISaxena@IDIAP.CH, EFiesler@IDIAP.CH

Abstract

An adaptive multilayer dual-wavelength optical neural network design with all-optical forward propagation, based on a large number of modifiable optical interconnections and a special weight discretization algorithm to compensate for system noise, is described. The presentation of input and interconnection weights is performed by liquid crystal television screens, and optical thresholding at the hidden layer by a liquid crystal light valve.

Keywords: optical computing, optical multilayer neural network, learning, spatial light modulator, liquid crystal light valve, optical neural network.

1 Introduction

Multilayer neural networks have proven their usefulness in an ever growing number of application domains. However, their main potential, which is their massive parallelism, is yet to be fully exploited, since they are usually simulated on electronic, single processor computers. Only true hardware implementations of neural networks, (*NN*s), will be able to exploit their massive parallelism. The most common hardware approach is electronic, using parallel computers or dedicated neural network VLSIs. However, electronic hardware is approaching physical speed limitations due to the enormous number of interconnections required by *NN*s.

Light, on the other hand, offers some very appealing characteristics: travelling at the ultimate speed, not requiring physically limiting wires, and having no cross coupling while several optical channels intersect in free space. Besides having these crucial advantages, optical *NN*s (*ONN*s) also have the potential to be scaled up in size without appreciable compromises.

One of the main problems with analog hardware implementations of *NN*s is the system noise which limits the number of different values that can be distinguished. A *NN* weight discretization technique is therefore required that restricts the number of possible weight values without performance loss.

Few *multilayer ONN*s, with limited applicability, have been suggested [1, 2]. Some opto-electronic *NN*s have been demonstrated, for example, with fixed interconnection weights [3], performing a 'recall-only' function; or the simulation of a multilayer *NN* by cycling through one perceptron many times [4]. *Adaptivity* refers to more than the updating of weights during the training process. Unlike "hard-wired" opto-electronic solutions where pre-trained weights of simulated *NN*s are fixed with masks or holograms, adaptive *NN*s are application independent.

A paucity of techniques in implementing efficiently a large number of modifiable interconnects has prevented the realisation of practical adaptive *ONN*s. The availability of liquid crystal screens with high pixellation densities, has opened up new possibilities [6] by providing simpler modifiable interconnects. These high density liquid crystal television (LCTV) screens can be used for the interconnection weights in large multilayer optical *NN*s. The description of an *ONN* design with a large number of adaptable weights using LCTV screens and optical thresholding, follows.

2 Methodology

Our three layer[1] neural network has 256 input neurons, 256 hidden layer neurons, and 16 output neurons. The input layer consists of a 16×16 data array, which is presented to the *ONN* for processing by LCTV1; (see figure 1). The hidden layer non-linearity, concerning 16×16 neurons, is implemented by a Liquid Crystal Light Valve (or LCLV). The 256×256 interconnection weights between these first two layers are represented by the analog transmission values of the pixels of a second LCTV (LCTV2). Similarly, the 256×16 interconnection weights between the hidden and output layer are represented by LCTV3. Figure 1 shows the optical system for

[1] In this report, a *layer* of a neural network is defined as a layer of neurons (cf. [7]). In the three layer neural network described in the text, the input layer is considered number one, the hidden layer number two, and the output layer as layer number three.

replicating the 16×16 data array (LCTV1) 256 times onto the transmissive LCTV2 for the vector-matrix multiplication. The design is based on positive results obtained with smaller scale, (7×7 input), *NN* implementations [5, 6].

The information flow through the *ONN* described above is as follows. After presenting the inputs to LCTV1, an initial random set of weights is presented to LCTV2. The intensity of light transmitted by LCTV2, representing the vector-matrix product of inputs and weights[2], is integrated by a micro-lens array and thresholded by the LCLV to form the outputs of the hidden layer. These outputs, read by a different light source, as explained in the following section, are similarly presented to LCTV3 and the following micro-lens array. The final outputs are detected by a 4×4 photodetector array. If thresholding is desired at the output layer, it may be performed either optically or electronically.

In this implementation of a multilayer *ONN*, training shall be performed with the aid of a personal computer. For a given input pattern on LCTV1, the personal computer presents initial weights to the LCTVs and subsequently collects (photodetector) outputs of the hidden and output layers. With these outputs, weight updates shall be calculated (by the algorithm described in section 3.4) which are again used to drive the LCTVs. This process is repeated iteratively, until the network has been trained.

3 Challenging Aspects

3.1 Three-layered structure

Both Hopfield and two-layer (perceptron like) electro-optical neural networks have been demonstrated with relative ease, as only a single vector-matrix multiplication is performed optically, while the electronics does the rest (like thresholding and the calculation of new weights).

A three-layer *ONN* is more complicated. A large fraction of the light is attenuated after having transmitted through the optical elements comprising the first two layers of the *ONN*, namely LCTV1, LCTV2, the gratings, and the LCLV. Solutions which allow enough light for further processing are therefore desired.

In our system, an Argon-ion laser (480nm) provides the light which transmits through LCTV1 and LCTV2, and writes to the LCLV. A He-Ne (633nm) light source reads the (thresholded) hidden layer output from the LCLV. This method should compensate for the loss of light in the first two layers of the neural network, and maintain sufficient optical power to drive optical processing in the following layer.

3.2 Large number of interconnections

A salient feature of this *ONN* is the large number of connections, nearing almost seventy thousand, inter-connecting the three layers of the network. The highest density of interconnections are implemented between the first two layers by LCTV2, in which 256×256 pixels shall be activated for the interconnection weights, (410×440 addressable pixels are available on each LCTV). In order that 256^2 (that is, 65,536)

[2] When light of unit intensity transmits through a material, a fraction of it is transmitted corresponding to the product of the incident intensity with the transmittance of the material.

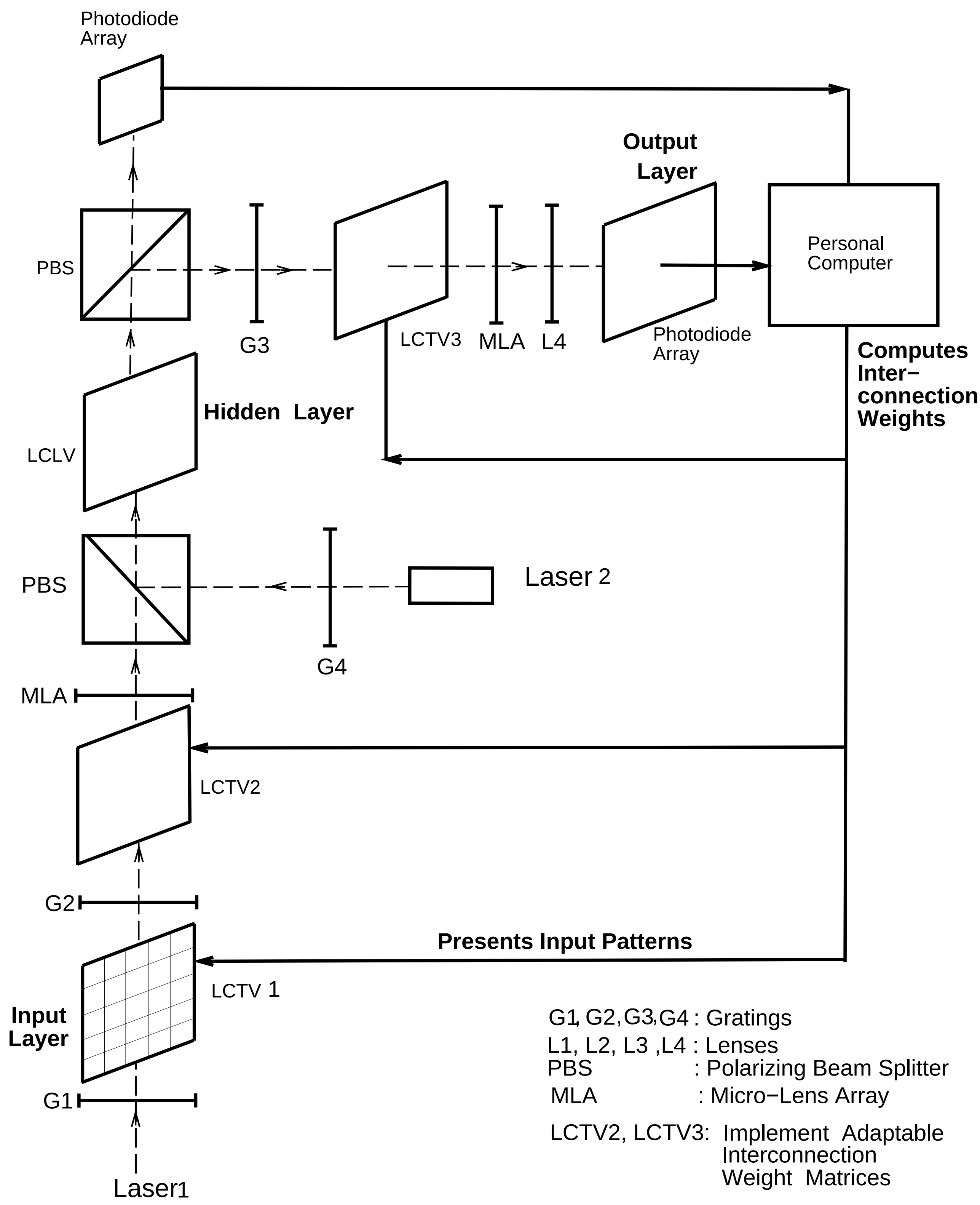

Figure 1: Schematic overview of the multilayer optical neural network.

parallel beams be generated and correctly aligned to transmit through an identical number of pixels on the LCTVs requires the optical elements to have tight performance tolerances. This necessitates building precision optical systems which are stable, robust, and modular.

3.3 LCLVs as Optical Thresholding devices

An LCLV is used as a non-linear optical thresholding device in the hidden layer. This reduces the overhead involved in conversion of optical signals into electronic, and re-conversion to opical after electronic thresholding. Such an 'all-optical' process results directly in optical inputs for subsequent layers, and a truly optical interface is established.

The incorporation of LCLV thresholding into the *ONN*, however, strongly influences the algorithm employed in the training process. Even though the transfer characteristic of the LCLV resembles that of a standard-sigmoid normally used in computer simulations of neural networks, it is truncated, non-centred, and largely assymetric in comparison. The transfer characteristic of one possible LCLV device, as compared to that of a standard sigmoid's, is shown in figure 2. The neural network learning rule needs to compensate for non-ideal thresholding functions of practical devices as this one. Preliminary simulation results with the LCLV transfer curve show that the performance of the neural network degrades as compared to that with standard-sigmoid thresholding.

3.4 Weight Discretization

Analog optical implementations incur inevitable penalties regarding the precision of computation. A principal focus of the research is the development of training methods for neural networks that compensate for this imprecision, while preserving the performance of the system. A method for discretizing weights and determining the degree of discretization with respect to performance has been developed [8]. Computer simulations based on this weight discretization method have shown very promising results in which neural networks with only a small number of discretization levels show no considerable loss in performance compared to their counterparts with continuous weights. This method of weight discretization shall be applied and customized to the *ONN* described here.

4 Summary

The design of a large, versatile adaptive multilayer optical neural network using available opto-electronic hardware, has been presented. It exhibits some of the major benefits of using optics, namely, massive parallelism and free-space interconnections with negligible beam interference. The system with close to seventy thousand interconnections is *large* compared to other SLM based systems. It is *versatile*, because it is application independent; only limited in maximum interconnection capability, which is easily upgradable by using larger SLMs. It is *adaptive* as it is application independent with modifiable weights to enable on-line learning. Above all, it is the first *multilayer NN* implementation with all-optical feed-forward propagation.

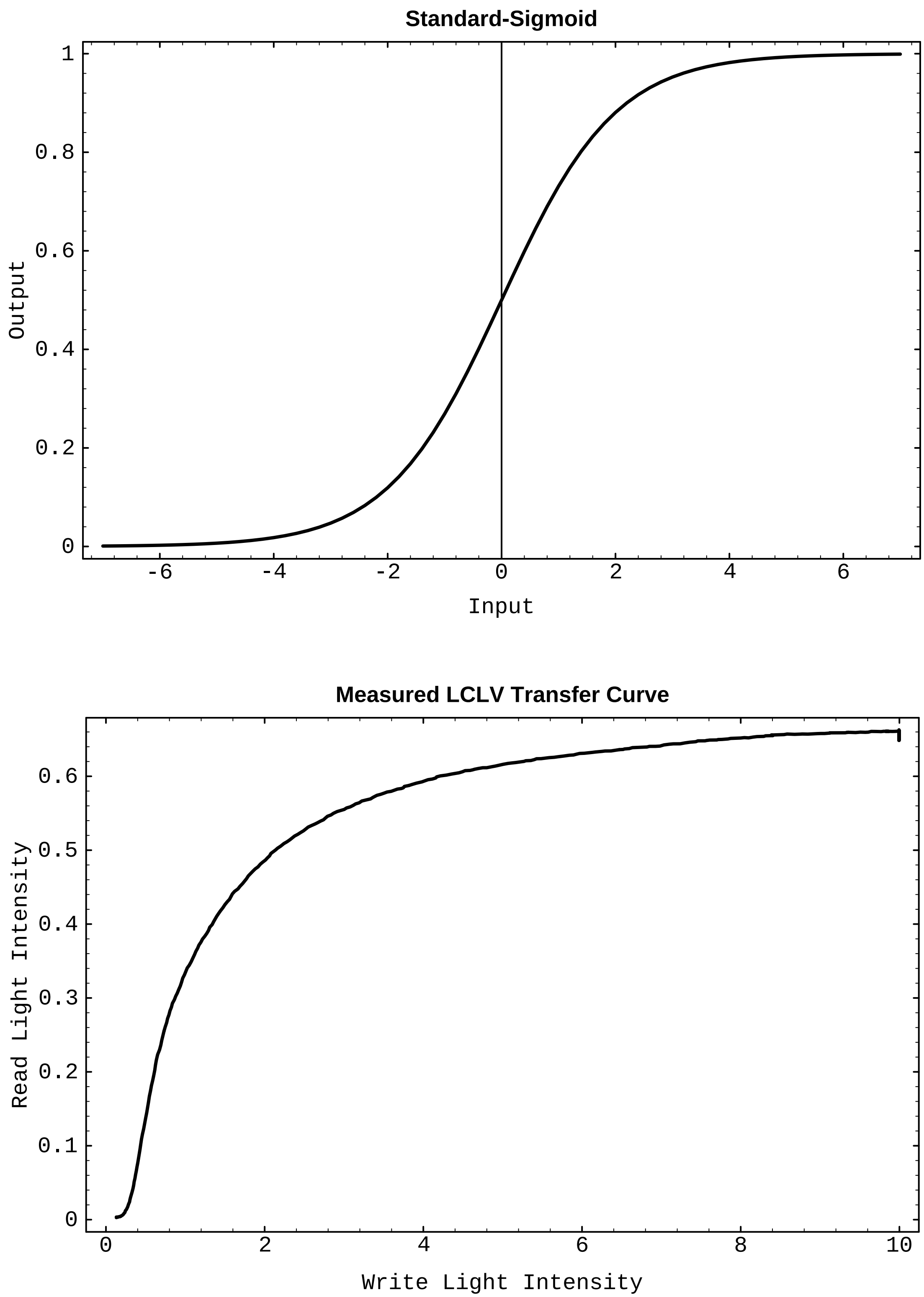

Figure 2: Transfer Curve of a Standard Sigmoid used in NN Computer Simulations and that of a typical LCLV.

The knowledge gained by the implementation of this design shall pave the way towards larger multilayer *ONN*s by device optimisation and further refinement of the learning algorithm.

5 Acknowledgements

This publication is an updated version of the original research proposal to the Swiss "Fond National". We would like to thank Neil Collings for his contribution towards the proposal, the design, and the project's progress in general. We are also grateful for the support from Daniel Osherson and René Dandliker.

This work is partially supported by the Fond National Suisse de la Recherche Scientifique, project number 21-36497.92.

References

[1] Kelvin Wagner and Demetri Psaltis, "Multilayer Optical Learning Networks", *Applied Optics*, Vol. 26, No. 23, pp. 5061–5076, December 1987.

[2] D. Psaltis and Y. Quio, "Optical Neural Networks", *Optics and Photonics News*, pp. 17–21, December 1990.

[3] J. Ohta et al., "Optical Neurochip Based on a Three-layered Feedforward Model", *Optics Letters*, Vol. 15, 1362–1364, 1990.

[4] J.-S. Jang, S.-G. Shin, S.-W. Yuk, et. al., "Dynamic optical interconnections using holographic lenslet arrays for adaptive neural networks", *Optical Engineering*, Vol. 32, No. 1, pp. 80–87, January 1993.

[5] K. J. Weible, G. Pedrini, W. Xue, and R. Thalmann, "Optical Implementation of a Neural Network Associative Memory Using Diffraction Gratings", *Japanese Journal of Applied Physics*, Vol. 29, No. 7, pp. L1301–L1303, July 1990.

[6] K. J. Weible, N. Collings, and A. Pourzand, "Initial Results of a Fully Interconnected Neural Network with Modifiable Interconnects", European Optical Society: Topical Meetings Digest series, Vol. 1, $8^{th.}$ *Workshop on Optics in Computing*, Paris, $8^{th.}$–$9^{th.}$ September 1992.

[7] E. Fiesler, "Neural Network Classification and Formalization", *Computer Standards & Interfaces*, Vol. 16, special issue on Neural Network Standards, John Fulcher, editor, North-Holland/Elsevier Science Publishers B. V., Amsterdam, The Netherlands, 1994, in press.
A preprint is available by anonymous ftp: ftp Maya.IDIAP.CH (or: ftp 192.33.221.1); directory: pub/papers/neural; file: fiesler.formalization.ps.Z (or fiesler.formalization.ps).

[8] E. Fiesler, A. Choudry, and H. J. Caulfield, "A Universal Weight Discretization Paradigm for Backward Error Propagation Neural Networks", accepted for publication by *IEEE Transactions on Systems,*

Man and Cybernetics.
See also: E. Fiesler, et al. "A Weight Discretization Paradigm for Optical Neural Networks." *Proceedings of the International Congress on Optical Science and Engineering*, volume SPIE-1281, pages 164–173. The International Society for Optical Engineering Proceedings, Bellingham, Washington, U.S.A., ISBN 0-8194-0328-8. 1990.