IDIAP RESEARCH REPORT

IDIAP
Martigny - Valais - Suisse

An Optical Thresholding Perceptron

I. Saxena [a] P. Moerland [a] E. Fiesler [a]
A. R. Pourzand [b] N. Collings [b]

IDIAP–RR 97-16

November 97

published in
Proceedings of the Workshop on Optics and Computer Science, Geneva, Switzerland, April 1997

Dalle Molle Institute for Perceptual Artificial Intelligence • P.O.Box 592 • Martigny • Valais • Switzerland

phone +41 – 27 – 721 77 11
fax +41 – 27 – 721 77 12
e-mail secretariat@idiap.ch
internet http://www.idiap.ch

[a] IDIAP, CP 592, CH-1920 Martigny, Switzerland, E-mail: Perry.Moerland@idiap.ch
[b] IMT, University of Neuchâtel, CH-2000 Neuchâtel, Switzerland

An Optical Thresholding Perceptron

I. Saxena P. Moerland E. Fiesler A. R. Pourzand N. Collings

November 97

published in

Proceedings of the Workshop on Optics and Computer Science, Geneva, Switzerland, April 1997

Abstract. An implementation of an optical perceptron with a soft optical threshold trained with an adapted BP algorithm is described as a precursor to an optical multilayer perceptron (MLP). It has 64 inputs and ten outputs. The soft threshold is implemented by a liquid crystal light valve. Experimental results on perceptron recall are also reported. The effect of a modified grey-scale to weight mapping for weight levels implemented by LCTVs is evaluated based on the results of handwritten digit recognition.

Acknowledgements: The support of the Swiss National Science Foundation (Grant SNSF 20-43220.95) is gratefully acknowledged.

1 Introduction

An initial test of a gradient descent based learning rule such as back propagation (BP) for multi-layer perceptrons, or MLPs, is to obtain its performance on a perceptron having a soft non-linearity at the output for thresholding that is not the hard limiter such as that of Rosenblatt's perceptron. Since MLPs with a soft threshold at the hidden layer are very appealing for a large range of applications [4], these algorithms, when implemented in optics, offer the additional benefits inherent in the parallelism of the optical approach.

We implement an optical perceptron with a soft optical threshold trained with an adapted BP algorithm [6][9] as a precursor to an optical MLP. The soft threshold can be implemented by an optical device, or an optically addressed spatial light modulator (SLM) with a non-linear response based on the electro-optic or magneto-optic effect, such as a thin-film SLM or a liquid crystal light valve (LCLV) [8]. We use a commercially available LCLV having a nematic liquid crystal as the electro-optic material. Optical thresholding maintains the spatial (optical parallelism) and overcomes the serial processing bottleneck of electronic thresholding [1]. We report here experiments on perceptron recall using an optical NN system described earlier [2]. The effect of a modified grey-scale to weight mapping for a set of weight levels implemented by LCTVs is measured for handwritten digit recognition using computer learned weight matrices.

An optical thresholding perceptron is composed of two parts: matrix-vector-multiplier (MVM) and a thresholding device. The MVM is described in section 2, followed by results on the MVM in section 3. Section 4 details the results with the optical thresholding perceptron. Section 5 addresses future problems, and section 6 concludes the paper.

2 Optical MVM

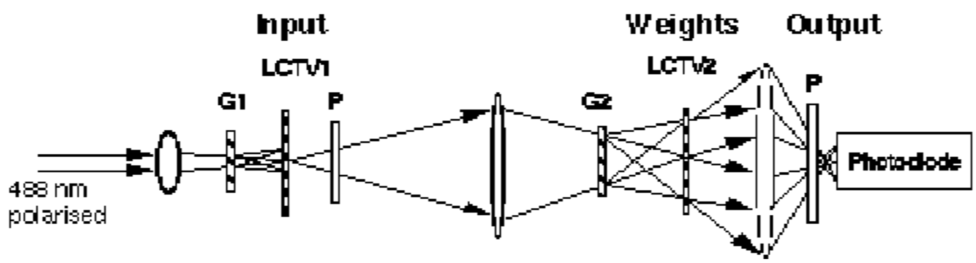

Figure 1: Optical MVM set-up.

The optical MVM is shown in Fig. 1. G1 and G2 are Damman gratings in 2-D, and each replicates the incident beams into 16×16 identical copies. Thereby 256^2 (or 65,536) optical channels are created which can be modulated, for instance, by liquid crystal television screens (LCTVs). Numerical matching between the simulation values and the transmission of the LCTV screens is ensured by choosing gray levels on the linear region of the transmittance curve of the LCTVs. The spatial matching (one to one mapping of the video frame memory and the LCTV pixels) is done through a software operation (resampling of the video frame memory before displaying it on the LCTVs). However, the capability of the 256^2 optical channels could not be implemented to its full extent, due to addressing problems with these LCTVs, and the effective number of channels had to be reduced; of inputs from 16×16 to 8×8, and of weights from 256^2 to 64^2.

Optical systems with large interconnections, favour a system with space invariant interconnects [2] using diffraction gratings. The information channels in this system (fig. 1) are coded in the intensity of the light using high resolution LCTV screens, LCTV1 and LCTV2, taken from a Seiko-Epson VPJ700 video projector.

Input pattern presentation using a combination of a laser with a grating fan-out and LCTV1 as in fig. 1 has been chosen over a more compact system resulting from using a laser diode array and collimation optics, due to the following advantages offered :

- reliability
- versatility for scalability
- misalignment tolerance

Also, in fig. 1, the input is not required to be in digital (electronic) form, and allows for a photographic film or transparency with the input pattern to replace LCTV1 for pattern presentation.

3 On the Perceptron: Inputs and Weights

Experiments were performed on the two layer ONN or optical perceptron. Training of the perceptron using the adapted, error backpropagation BP algorithm for all-positive optically thresholding ONNs with discretized weights was first simulated on (SUN) workstations. The new weights are proportional to the derivative of the thresholding function. The matrix-vector product output is also thresholded by this same function to obtain the outputs which are then compared with the desired outputs. The measured response function of an LCLV [9] is used as the soft thresholding function.

A subset of 50 patterns of hand-written digits was taken from the NIST database set [3] as the training data. Training is continued until the maximum difference over all output units and all patterns becomes less than a predefined acceptable interval. The interconnection weight matrix (IWM) defining the connections between the input and output neurons that is thus obtained [7] is used in our optical perceptron to test recall (or recognition) of presented input patterns on which it has been trained. The idea is to compare the computer simulated and the optical hardware performance.

The recall is verified in two stages: firstly, the integration is performed in software (thresholding is done by the stored LCLV response curve); in the second stage, the integration is done by a demagnifying telescopic system and thresholding by the LCLV, as discussed in the following sections.

3.1 Input Test Patterns

The NIST benchmark on handwritten digits consists of a database with 20790 digitized handwritten characters from 2100 individuals [3]. Each digit was scaled to fit into a 32×32 matrix, and each pixel is represented by an eight bit value. For our purposes the 32×32 matrix has been converted to an 8×8 matrix, by taking the average of 4×4 sub-matrices, and the input values have been scaled to the interval $[0, 1]$. The target patterns are the ten unit vectors for training of the multilayer neural network on the computer for which a subset of 500 patterns is first used. A subset of 50 digits is taken from this benchmark set, on which the computer simulated perceptron training was performed, from which an example set of ten digits is shown below in fig 2.

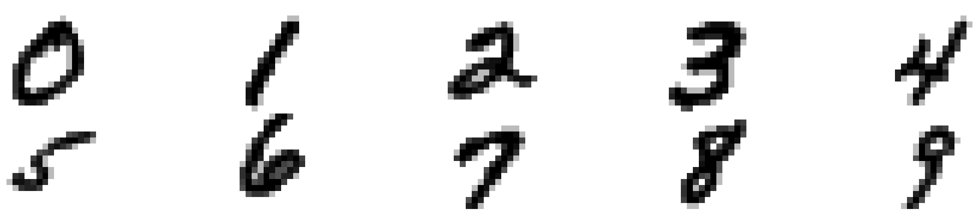

Figure 2: Test Inputs: A set of hand-written digits.

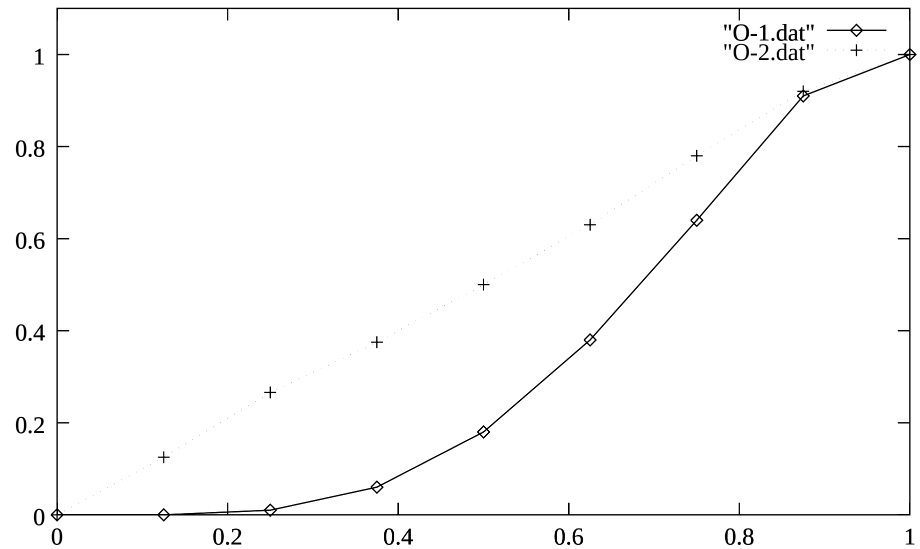

Figure 3: Linearization of LCTV2 behaviour with O-2.

3.2 Mapping TV Gray-Levels to Weights

The weights of the IWM applied to the LCTVs are encoded via the grey levels which translate into the electrical fields applied to the LCTV pixels, by which their transmission is changed. The transmission of the LCTVs varies non-linearly with the gray levels and so also with the weights (when they are linearly related to the grey levels). However, if the relationship of these eight weight levels (the number of levels required in our training algorithm) to the gray levels is made non-linear, then the weights to transmission relationship is linearized as shown in figure 3.

An initial weight-mapping, O-1, using equidistant gray levels for the eight weight levels, and 16 equidistant gray levels for the inputs, both in a range of 0-255, was first taken. A second weight-mapping, O-2, for LCTV2 was then taken, while keeping the inputs the same as in O-1. These new gray levels for the weights are 0, 119, 153, 170, 187, 204, 221, 238. The relation between the weights and the LCTV2 transmission under the two mappings, O-1, and O-2 are shown in figure 3.

3.3 Recall in the Optical MVM

Initially the perceptron configuration with software integration was tested. A performance metric, or discrimination measure, to assess the optical NN recall performance on the digit pattern set, is defined similar to the often used mean squared error or an average error on all patterns. The pre-thresholded neuron outputs, or the sum of all the inputs at each neuron are then compared on the basis of this error.

$$MSE = \frac{1}{N}\Sigma_N (C^N - O^N)^2 \tag{1}$$

$$Standard\,Error = \sqrt{MSE} \tag{2}$$

where Σ_N indicates summation over all output neurons, and C^N and O^N are the normalized outputs, defined as:

$$C^N = \frac{\Sigma C}{(\Sigma C)_{max}} \tag{3}$$

$$O^N = \frac{\Sigma O}{(\Sigma O)_{max}} \tag{4}$$

where ΣC are the computed sum of inputs to the output neuron, and ΣO are the corresponding optically obtained sum of inputs. The experimental results obtained under these three mappings are shown in table 1, show quite good agreement with the computed values (they are within 15%). With O-2 showing a slight improvement in standard error than with the uncompensated O-1 mapping.

Input digit Pattern	0	1	2	3	4	5	6	7	8	9	Std. Error
Computed*	9.85	6.38	7.74	8.75	7.07	5.26	8.04	7.61	7.34	5.49	-
Optical:O-1^N	1.0	0.83	0.94	0.86	0.81	0.56	0.74	0.72	0.72	0.48	0.09
Optical:O-2^N	0.88	0.59	0.93	1.00	0.73	0.58	0.88	0.78	0.74	0.53	0.07

Table 1: ONN Recall performance: Comparison of simulated and optical results prior to thresholding, showing the outputs at the winner neurons: * Actual Simulated Values; O-1,2^N are normalized optical output values).

3.4 Discussion of Results

In Table 1 the outputs of the computer simulated results are shown. It is observed that the sum of inputs ranges from 9.8 to 5.2. This variation seems to be primarily due to the difference in the length of the input vector (or the sum of all the pixel values in each input vector). In order to make these output values independent of the size of each digit, an input normalization procedure could be used. This may be obtained by scaling the digit length vectors to the same value [5]. Such a scaling may be more readily performed by simply scaling-up the pixel values to obtain the normalized length. However, in the optical implementation, the pixel value cannot exceed a value of unity (a maximum transmission). Hence another scheme could be to magnify the size of the digit to achieve a normalized input length.

4 Recall with Integration and LCLV Thresholding

The optical system for this is shown in figure 4. Non-ideal sigmoidal responses (such as of LCLVs) can be incorporated in the BP training algorithm as described in [7]. Based on this training that is simulated on SUN SPARC-stations, a corresponding IWM is obtained, that is used to test recall on our optical perceptron.

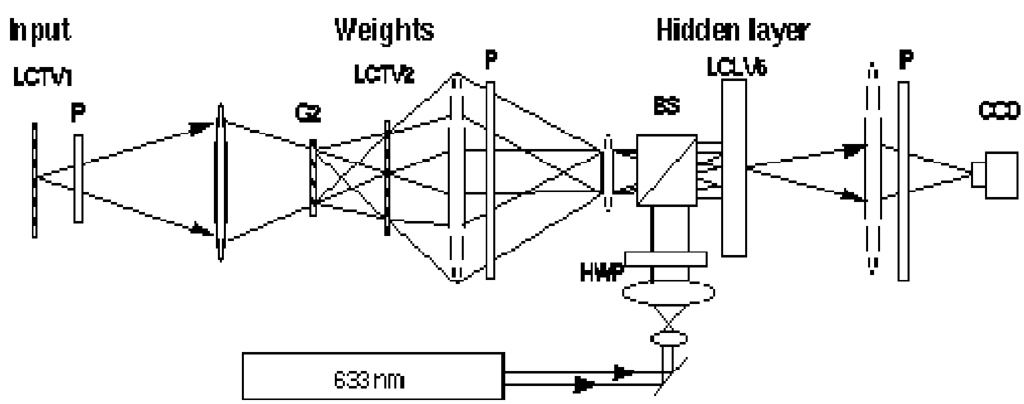

Figure 4: Complete two-layer ONN1 network with optical thresholding. P: Polariser; Gn: Gratings; BS: Beamsplitter; HWP: Half Wave Plate; CCD: camera.

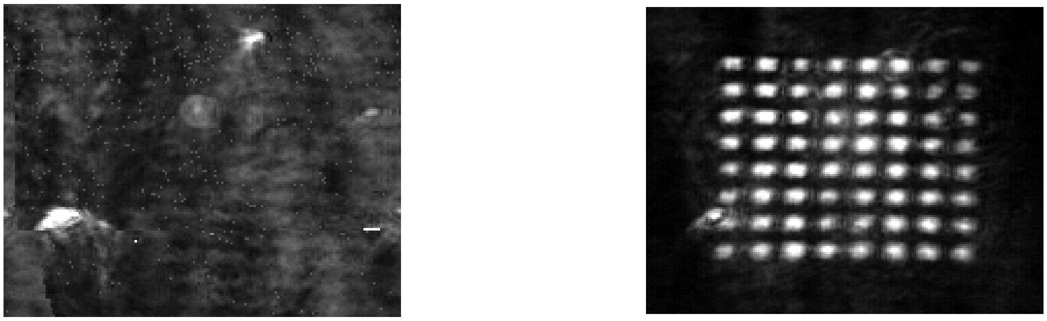

Figure 5: LCLV output images without and with write light.

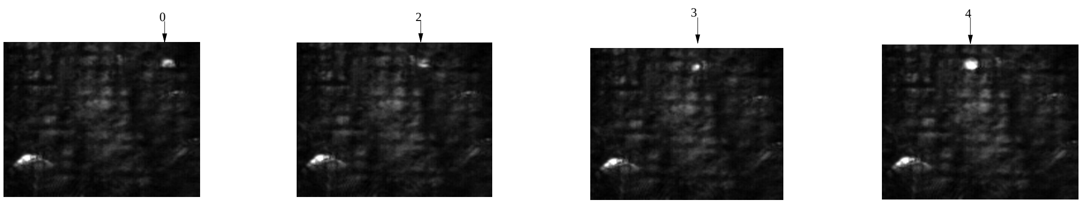

Figure 6: Recognition of a 0, 2, 3, and 4.

4.1 LCLV Spatial Characteristics

Images of the light transmitted by the LCLV at a drive voltage of +5V without the write beam, and with a uniformly equal intensity illuminated array as obtained in the optical system of figure 4 are shown in figure 5. The LCLV has some bright spots of transmitted light (as seen in the figure) even in the absence of the write beam. A differencing method is used to compensate for this. Alternatively, since only about one fourth of the LCLV area is required for the demagnified 8×8 array image, a better performance could be obtained by translating the LCLV to position the output pattern at a spot which is free from defects and non-uniformities.

4.2 Thresholded Recall

In the figures 6 and 7, the output images, showing the thresholded outputs after the LCLV on presentation of various hand-written digit patterns, are shown. The outputs are encoded such that the upper rightmost element of the 8×8 array represents a recognition of a '0', and we count from right to left in increasing numbers, until the last leftmost element in the first row represents a '7'. In the second row, the rightmost block represents an '8', and the next left element, a '9'. The other 54 output blocks do not have any significance in the present set of experiments, as only ten output classifications are required.

A few broad interference fringes running across the image, slanting to the left are also observed in these figures. These are present in the absence of the write light beam as well and are due to the shear interferometer that the sandwich of glass plates of the LCLV makes up. This causes a modification of the optically thresholded outputs as compared to those prior to optical thresholding. This can be seen by comparing the relative magnitudes of the outputs in figures 6 and 7 which are quite different from those in table 1. Before thresholding, the output signal at '5' is expected to be low as compared to at '6', but after the LCLV, due to its location close to the middle of the interference maxima, it now gives a higher output than at the location of '6' which now lies close to an interference minima region. This interference effect can be removed by making the substrate glass of the LCLV having a wedged form, with a very small wedge angle which can be as small as 1 degree.

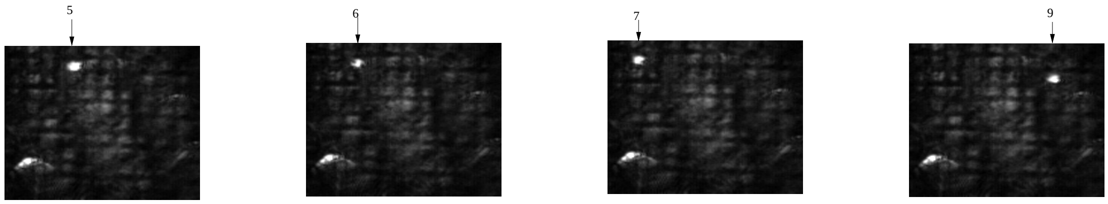

Figure 7: Recognition of a 5,6,7, and 9.

5 Requirements for On-Line Learning with LCLVs

For successful on-line or 'chip-in-the-loop' learning when the optical system includes LCLV thresholding, several aspects are important that include: alignment stability, sufficiency of light intensity for LCLV switching and the requirement of spatial uniformity of the LCLV.

The correct relative alignment of the LCTVs so that the fanned-out images of TV1 (of 64 blocks) are aligned with the electrically-addressed blocks of pixels on TV2, must remain unchanged during the learning or training process.

The power transmitted by the optical system which is incident on the LCLV is about 0.1% of that incident on the first grating G1. The operating state of the LCLV is verified by measuring its response over the range that the intensity is modulated by the LCTVs. When the incident power on G1 is 1.56 mW, then the 8×8 array intensity incident on LCLV is 200 nW, which is less than the power required to switch it fully 'on'. The training on the optical system does therefore not proceed predictably.

The on-line ONN training is expected to proceed successfully only if the behaviour of the optical thresholding element is the same as that stored in the training program. As the derivative of the non-linear response required in our training algorithm is obtained from the stored thresholding functional form. That is to say, the response should be identical to that stored, and that the LCLV should be ideally spatially uniform in its response.

The background light distribution on the CCD detector (behind the LCLV) is removed by taking the difference of the actual signal and the background. However, there remain spatial non-uniformities in the response of the LCLV which cannot be removed by this background elimination. Therefore, during training, the outputs at certain locations on the LCLV may not increase as expected and training may not converge.

6 Conclusions

Results on the functional integration of the optical matrix-vector multiplier and the hidden layer with the LCLV, of the ONN are presented. Performance of this ONN is extremely satisfactory as an optical perceptron configuration, when the integration is done in software independent of the optical LCLV thresholding.

On including the LCLV for thresholding, the reduced recall performance is due to interference effects and spatial non-uniformity of the LCLV, which are only partially compensated for by taking the difference with the existing background. Interference effects can be removed by using a wedge substrate for the LCLV, but other schemes may also need to be developed to overcome the remaining non-uniformities. On-line learning could also compensate, to some extent, for non-uniformities provided the conditions outlined in sec. 5 are satisfied. In following work, an improved optical system having a higher interconnectivity ONN and with an improved LCLV is to be implemented using new LCTV screens.

References

[1] Alain Bergeron, Henri H. Arsenault, Erik Eustache, and Denis Gingras, “Optoelectronic thresholding module for winner-take-all operations in optical neural networks,” *Applied Optics*, volume 33, number 8, pp. 1463–1468, March 1994.

[2] N. Collings, A. R. Pourzand, and R. Volkel, “Construction of a programmable multilayer analogue neural network using space invariant interconnects,” in *Proc. SPIE*, Vol. 2565, pages 40–47, 1995.

[3] M. D. Garris and R. A. Wilkinson, data from NIST Special Database 3, National Institute of Standards and Technology, Advanced System Division, Image Recognition Group, February 1992.

[4] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, volume 2, Pages 359–366, 1989.

[5] Thomas O. Jackson, “Data Preprocessing,” sec. B4.4 in the *Handbook of Neural Computation*, Editors E. Fiesler and R. Beale, OUP and IOP, 1996.

[6] P. Moerland, E. Fiesler, and I. Saxena, “The Effects of Optical Thresholding in Backpropagation Neural Networks,” *Proceedings of the International Conference on Artificial Neural Networks (ICANN'95)*, Editors F. Fogelman-Soulié and P. Gallinari, Volume 2, Pages 339–343, Paris, France, October 9–13, 1995, 1995.

[7] P. Moerland, E. Fiesler, and I. Saxena, “Incorporating LCLV Non-Linearities in Optical Multilayer Neural Networks,” *Applied Optics*, volume 35, number 26, pages 5301-5307, September 1996.

[8] I. Saxena and E. Fiesler, “Adaptive Multilayer Optical Neural Network with Optical Thresholding,” *Optical Engineering*, volume 34, number 8, pp. 2435–2440, August 1995.

[9] I. Saxena, E. Fiesler, and P. Moerland, “A Method for All-Positive Optical Multilayer Perceptrons,” in *Proceedings of the Third IEEE International Conference on Electronics, Circuits, and Systems (ICECS'96)*, vol. 1, pp. 448–451, IEEE, Piscataway, NJ, 1996.