AN UNIPOLAR TERMINAL-ATTRACTOR BASED NEURAL ASSOCIATIVE MEMORY WITH ADAPTIVE THRESHOLD AND PERFECT CONVERGENCE

Chwan-Hwa "John" Wu
Dept. of Electrical Engineering
Auburn University
Auburn, Alabama, 36849-5201
and
Hua-Kuang Liu
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109-8099
Telephone (818)354-8935 FAX (818)354-9509

Abstract

For the first time, a unipolar terminal-attractor based neural associative memory (TABAM) system with adaptive threshold and perfect convergence is presented. By adaptively setting the threshold values for the dynamic iteration for the unipolar binary neuron states with terminal-attractors and inner-product approach, we demonstrate via computer simulation the achievement of perfect convergence and correct retrieval. The simulation is completed with a small number of stored states (M) and a small number of neurons (N) but a large M/N ratio. An experiment with exclusive-or logic operation using LCTV SLMs is used to show feasibility of the optoelectronic implementation of the models.

Key Words: Associative Memory, Terminal Attractor, Neural Network, Computer Simulation, Inner-product, Adaptive Threshold, Hopfield Model.

I. INTRODUCTION

One of the major applications of neural networks is in the area of associative memory. For example, the avalanche of intensive research interests in neural networks was initiated by the work of Hopfield[1] in which the associative memory is modeled with an neural synaptic interconnection matrix and encompasses an interesting computation scheme via nonlinear threshold and iterations. However, after further investigations, it was found[2,3] that the storage capacity of the Hopfield model is quite limited due to the number of spurious states and oscillations.

In order to alleviate the spurious states problems with the Hopfield model, the concept of terminal attractors was introduced by Zak[4]. However, the theory of the terminal attractor based associative neural net model proposed by Zak determines that a new synapse matrix totally different from the Hopfield matrix is needed. This new matrix, which is very complex and time-consuming to compute, was proven to be able to increase the speed of convergence and control the basin of attraction[4,5]. Zak's derivation shows that the Hopfield matrix only works if all the stored states in the network are orthogonal. However, since the synapse is changed, the Zak model is completely different from the Hopfield model. The impact of the terminal attractor on the Hopfield model cannot be determined based on Zak's model.

More recently, for the purpose of comparing the Hopfield model with and without the inclusion of the terminal attractor, Liu et.at. proposed a terminal attractor based associative memory(TABAM) model with binary neurons and the Hopfield matrix[6]. Several techniques for the optical implementation were introduced which include the application of the inner-product approach[7] and the exclusive-or (XOR) operation of the liquid crystal television spatial light modulator (LCTV SLM)[8].

The complexity of the optical implementation of the TABAM is discussed in Ref. 6. It is desirable to develop a unipolar neuron model which is more suitable for optical implementations. In this paper, we present a unipolar inner- product TABAM (UIT) and a across-talk reduced inner-product TABAM (CRIT). The latter provides a mechanism to put the input state vector in the correct basin of the stored vector, then, using the terminal attractor to accelerate the convergence. Besides, the cross-talk reduced model does not require any training time and therefore is readily for optoelectronic implementation. In both UIT and CRIT, a dynamical logistic function adaptive thresholding technique is developed..

Before the optical implementation is designed, it is important first to test via computer simulation the validity/practicality of the preservation of the Hopfield matrix in UIT and CRIT even though the stored states are not orthogonal. This simulation needs the determination of the threshold value that it takes to reach convergence in the associative recall. For simplicity, only a small number of neurons and a small number of stored states are used for demonstrative simulations. Secondly, the feasibility of an optoelectronic implementation of the technique is demonstrated with a small number of neurons and stored states. The application of the small-number TABAM for solving pattern classification and recognition problems is discussed.

II. ADAPTIVE THRESHOLD DETERMINATION

In the unipolar system, it is difficult to implement subtraction using optics. To avoid subtraction, an adaptive thresholding technique is used. Before the adaptive thresholding techniques is discussed the TABAM[6] model is briefly reviewed. In the model, one assumes that the i-th component of the state vector x_i at time $t+1$ may be written as a function of t as follows:

$$x_i(t+1) = \sum_{m=1}^{M} v_i^m \ \alpha^m(t) + I_i, \tag{1}$$

where

$$I_i = -\sum_{m=1}^{M} \alpha^m(t)[g(x_i(t)) - v_i^m]^{\frac{1}{3}} \delta^m(t),$$

$$\delta^m(t) = \exp\{-\beta \sum_{i=1}^{N}[g(x_i(t)) - v_i^m]^2\},$$

$$\alpha^m(t) = \sum_{j=1}^{N} v_j^m \ g(x_j(t)),$$

where $g(x_j(t))$ is a logistic function.

In Eq. (1), v_i^m denotes the ith component of the mth stored vector V^m and β is a constant, M and N are the number of stored vectors and the number of neurons respectively. Based on the property of the unipolar representation of a binary system, this Equation can be rewritten as:

$$x_i(t+1) = \sum_{m=1}^{M} \alpha^m(t)[(1+\delta^m(t))v_i^m - g(x_i(t))\delta^m(t)], \tag{2}$$

Instead of implementing the substraction in Eq. (2) optically, an adaptive threshold function is introduced following the modification of Eq. (2) as follows:

$$x_i(t+1) = \sum_{m=1}^{M} \alpha^m(t)[(1+\delta^m(t))v_i^m + g(x_i(t))\delta^m(t)] \tag{3}$$

We refer to the model of Eq. (4) as the UIT.

In Eq. (3), a logistic function, as illustrated in Fig. 1, with an adaptive threshold, $\theta(t)$, is used to perform the nonlinear transform such that the output becomes quasi-binary states. The adaptive threshold will be determined according to whether the stored states are orthogonal to one another or not. The system is crosstalk-free if all the stored states are orthogonal.

When there is no crosstalk, then the threshold $\theta(t)$ is set as

$$\theta(t) = \alpha^{m'}(t)[\frac{1}{2} + \delta^{m'}(t)], \tag{4}$$

where $\alpha^{m'}(t)$ is the inner product between $x(t)$ and the m'th stored vector, $V^{m'}$, and

$$\alpha^m(t) = 0 \qquad \text{for all } m \neq m'.$$

In order to cause the states to converge to the stored states, a threshold can be set based on Eq. (4) for maximum noise immunity as shown in Fig. 2. The threshold is set between the lower value of the desired output state 1 and the higher value of the desired output state 0.

A more practical situation is where the stored states are not orthogonal to one another. In this situation, crosstalk often occurs between similar stored vectors. For this case, $\theta(t)$ is set to be:

$$\theta(t) = \sum_{m=1}^{M} \alpha^{m}(t)[\frac{1}{2} + \delta^{m}(t)] \tag{5}$$

In order to reduce/eliminate the crosstalk problem, a new model called CRIT is proposed as follows:

$$x_i(t+1) = \sum_{m=1}^{M} \alpha^{m}(t)\delta^{m}(t)[(1+\delta^{m}(t))v_i^{m} + g(x_i(t))\delta^{m}(t)] \tag{6}$$

It can be seen from Eq.(6) that the crosstalk between the nonorthogonal stored vectors is exponentially weighted and reduced by the term $\delta^{m}(t)$. The size of the basin of terminal attractors is controlled by the value of β in $\delta^{m}(t)$. Using the model, the retrieved state vector is first placed in a correct basin, then the corresponding terminal attractor makes the state vector converge rapidly to the bottom of the basin. This model does not need time for training and is more suitable for optoelectronic implementation since the stored vectors can be directly used. The threshold can be set similar to Eq. (5) as:

$$\theta(t) = \sum_{m=1}^{M} \alpha^{m}(t)\delta^{m}(t)[\frac{1}{2} + \delta^{m}(t)]. \tag{7}$$

III. COMPUTER SIMULATION

In order to test the effectiveness of the models of UIT and CRIT using the adaptive threshold, a computer simulation code is developed. The XOR logic operation is used to detect the Hamming distance between the state vector, $x(t)$, and the stored vectors, V^{m} . An input-output relationship of XOR between the input and output is shown in Table I.

The Hamming distance between the state vector, $x(t)$, and the stored vectors can be computed by using XOR as follows. For a unipolar system,

$$[g(x_i(t)) - v_i^m]^2 \approx | g(x_i(t)) - v_i^m | \tag{8}$$

Based on the above Table, the Hamming distance is found to be

$$\begin{aligned} \sum_{i=1}^{N}[g(x_i(t)) - v_i^m]^2 &= \sum_{i=1}^{N} | g(x_i(t)) - v_i^m | \\ &= \sum_{i=1}^{N}[1 - g(x_i(t))\mathrm{XOR} v_i^m] \end{aligned} \tag{9}$$

Equation (9) provides a similarity measure between $g(x(t))$ and V^m, and is equivalent to the inner product of a bipolar system.

Based on the above algorithm, a computer simulation program was used to test the feasibility of the models. For simplicity of computation, the number of neurons and stored states selected are small. The number of stored vectors are 2 and 3 for both 3 and 4 neurons in each state. With the small N and M, all possible combinations of vectors have been stored and the retrieval was tested with all possible input vectors. In other words, an exhaustive test was done with the simulation. An IBM PC 386 was used for the simulation. An associative recall result is considered to be accurate when an input vector converges to a stored vector with the smallest Hamming distance to the input vector. The accuracy can be determined in all occasions except in those cases when the input vector has an equal Hamming distance to all of the stored vectors. In this case one cannot decide which of the stored state should the input converge to. In reality, it really does not matter because of equal Hamming distance. The simulation results of the UIT (Eq. (3) and CRIT (Eq. (6), are presented in Table II. The 95.7% convergence accuracy of $M = 3$ and $N = 4$ of the UIT is due to the cross-talk effect. The inaccuracy is removed by the CRIT as shown.

Based on the results of the exhaustive computer simulation, it can be seen that even $M = N$, perfectly accurate convergence can be accomplished. The details of the simulation of a particular case will be provided in an example below, while the optoelectronic implementation of the models is discussed.

The output images following the XOR operations are focused to become two spots on the CCD camera. The intensity ratio of the two spots should be $9 : 4$. The experimental result showing the calculated intensity ratio is displayed by an oscilloscope and shown in Fig. 4. The computer can dynamically use a logistic function to perform the computation as guided by Eqs. (3) or (6).

The result obtained for UIT before taking the threshold is

$$\begin{bmatrix} 7.7486 & 6.3743 \\ 5.4779 & 7.7486 \end{bmatrix}$$

and the threshold θ of the logistic function is 3.874309 (when one sets $\beta = 1.0$). After taking the threshold, the result is

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

which shows an accurate retrieval. The thresholded output will remain in the same state by the iterative process.

On the other hand the result obtained for CRITA before thresholding is

$$\begin{bmatrix} 2.2596 & 1.8169 \\ 1.9523 & 2.2596 \end{bmatrix}$$

After taking the threshold with $\theta = 1.1298(\beta = 1.0)$, the result is the same as that of the UIT.

The dynamical system converges to the correct state and remains in the same thresholded output state forever..

The above example offers a feasibility demonstration of using a computer to perform nonlinear thresholding and feedback. An example of an optical thresholding and feedback is illustrated in reference No. 6, but the fact that a suitable threshold device is not available limits pure optical implementaiton capability. Examples of states thresholding to zeros are also available but not shown due to space limitations.

V. CONCLUSION

In this Letter, two unipolar terminal-attractor based associative memory models UIT and CRIT, are presented with adaptive logistic functions. Computer simulations on the associative retrieval of these models with a small number of neurons and a small number of stored states demonstrate perfect recall and convergence. Corresponding experiment using XOR operation of LCTV SLMs demonstrated the feasibility of optoelectronic implementation of the models.

The perfect convergence for the case of $N = M$, even though both M and N are small, indicates that the terminal attractors are unique and effective in making tremendous improvements on the otherwise limited Hopfield model. In a pattern recognition/classification problem, the number of pixels of the input iamge is usually on the order of hundreds. In order to solve the problem, it is conceived that the large dimensions of neurons may be divided into small "cells". The data may therefore be reduced by cascading the processors following the rules of UIT and CRIT. At each stage, the accuracy of convergence is perfect. The potentials of these models such as scaling, cellular approach, and pyramedial multi-resolution image classification based on UIT and CRIT are under investigation.

ACKNOWLEDGEMENT

This paper describes research which was jointly sponsored by the Defense Advanced Projects Agency and the Information Systems Branch, Office of 'Space Science and Applications, National Aeronautics and Space Administration, and was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboraotry, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. We would like to thank Mr. Joseph Diep for experimental assistance and Annie Aroyan for typing the manuscript. One of the authors (C. Wu), would like also to thank the support from Center of Commercial Development for Space Power and Advanced Electronics/NASA.

REFERENCES

1. J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", Proc. Nat. Acad. Sci., USA, *79*, 2254 (1982).

2. R. J. McEliece, E. C. Posner, E. R. Rodemich, and S.S. Venkatesch, "The Capacity of the Hopfield Associative Memory", IEEE Transactions on Information Theory, *T-33*, 461 (1987).

3. B. L. Montgomeryand B.V.K. Vijaya Kumar, "Evaluation of the use of Hopfield Neural Network Model as a Nearest Neighbor Algorithm", Appl. Opt., *25*, 3759 (1986).

4. M. Zak, "Terminal Attractors for Addressable Memory in Neural Networks," Phys. Lett. *A-133*, 18 (1988).

5. M. Zak, "Terminal Attractors in Neural Networks," Neural Networks, *2*, 259-274 (1989).

6. H.K. Liu, J. Barhen and N.H. Farhat, "Optical Implementation of Terminal Attractor Based Associative Memory," submitted for publication.

7. S.Y. Kung and H.K. Liu, "An Optical Inner-Product Array Processor for Associative Retrieval," Proc. SPIE *613*, 214 (1986).

8. H.K. Liu and T.H. Chao, "Liquid Crystal Television Spatial Light Modulators," Appl. Opt., *28*, 4772-4780 (1989).

FIGURE CAPTIONS

Figure 1. A dynamic logistic function with a sharp transition at $x = \theta$ (threshold).

Figure 2. The adaptive threshold, θ, is set to achieve the maximum noise immunity.

Figure 3. An experimental setup for the optical IOR operations.

Figure 4. Two focused spots of optical image intensities following XOR operations. The intensity ratio between the two is approximately $9 : 4$.

NAMES OF TABLES

Table I. Exclusive OR (XOR) Relationship

Table II. Computer Simulation Results of UIT and CRIT

Table I. Exclusive OR (XOR) Relationship

Input 1	Input 2	Output
1	1	1
0	1	0
1	0	0
0	0	1

Table II. Computer Simulation Results of UIT and CRIT

Model	N	M	Convergence Accuracy
UIT	3	2	100%
UIT	3	3	100%
UIT	4	2	100%
UIT	4	3	95.7%
CRIT	3	2	100%
CRIT	3	3	100%
CRIT	4	2	100%
CRIT	4	3	100%

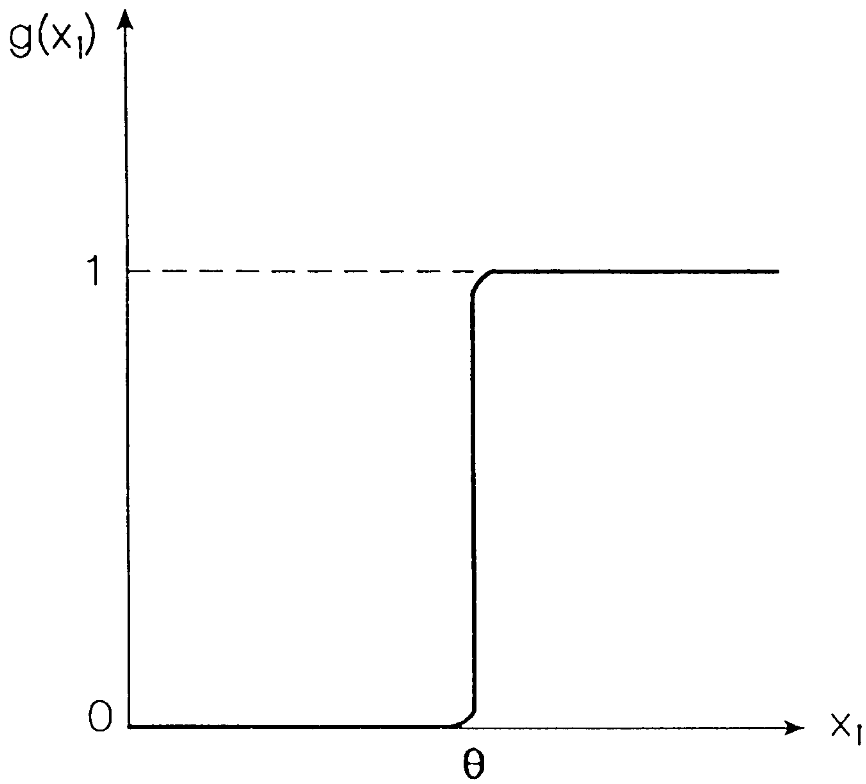

Fig. 1

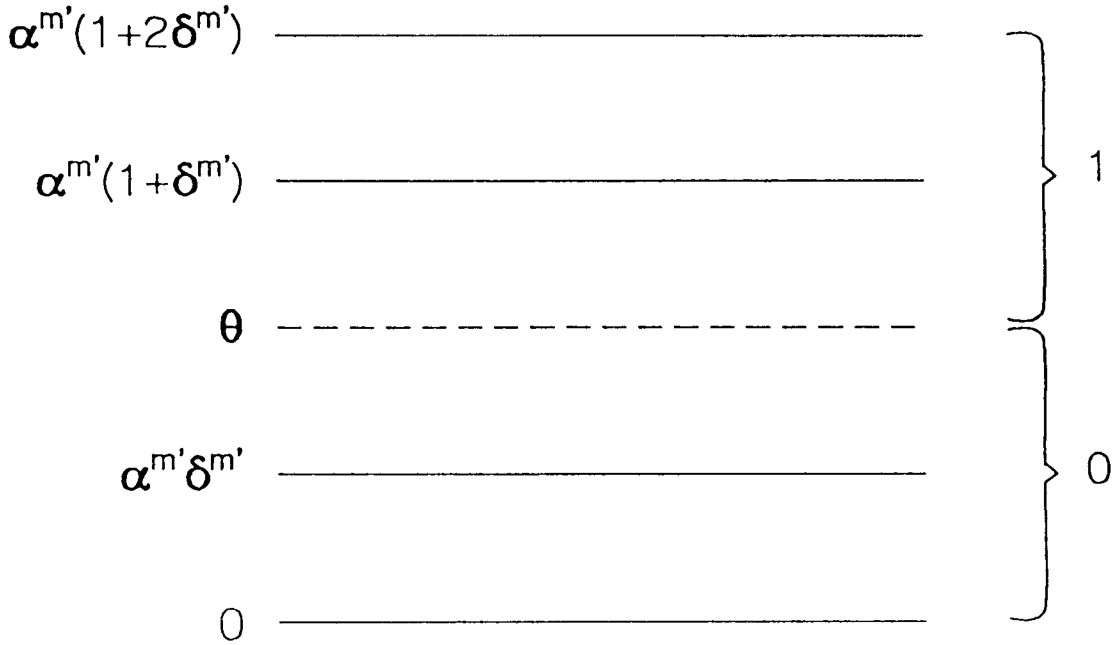

Fig. 2

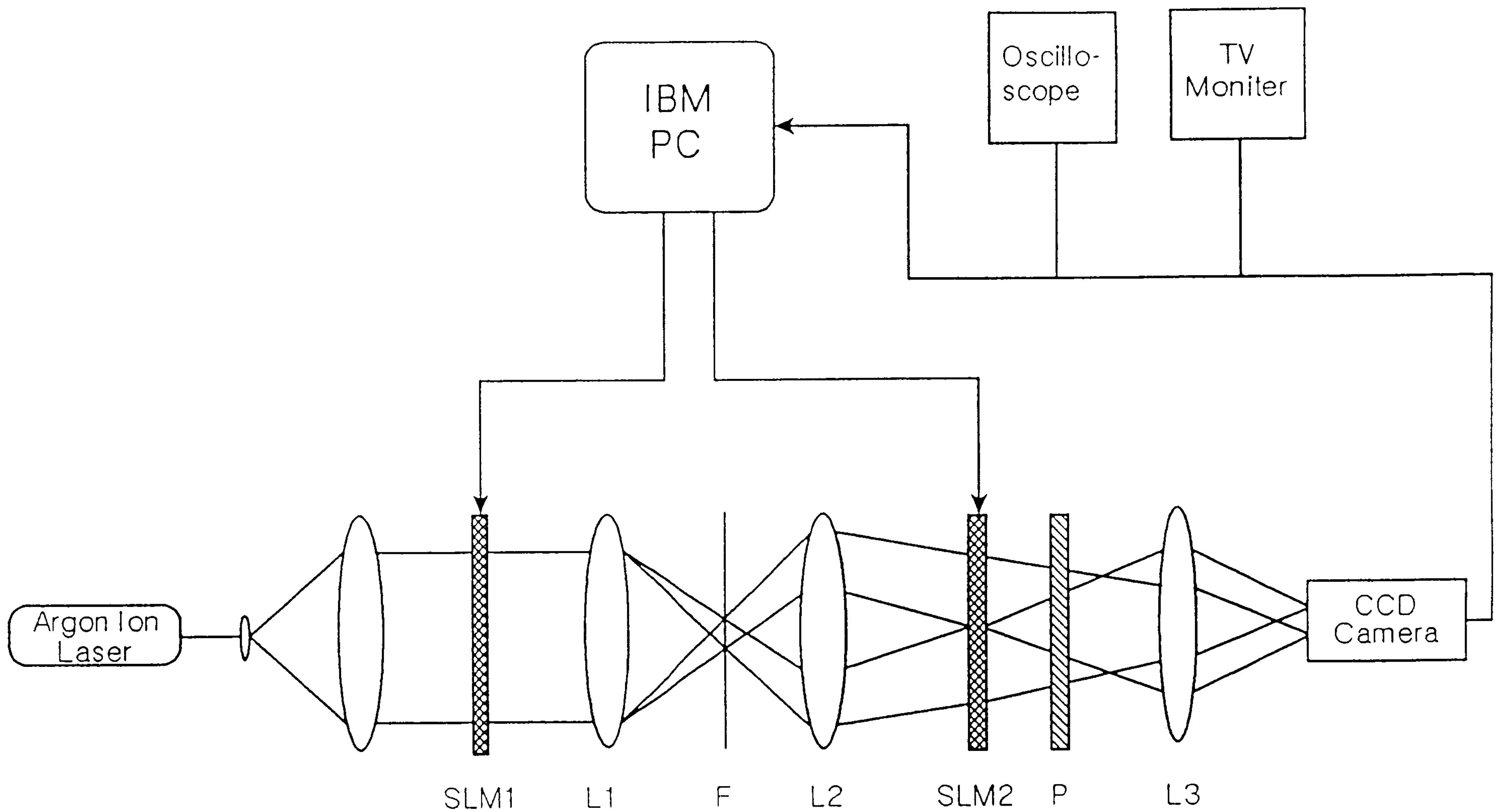
IBM
PC
Oscillo-
scope
TV
Moniter
Argon Ion
Laser
CCD
Camera
SLM1
L1
F
L2
SLM2
P
L3

Fig. 3

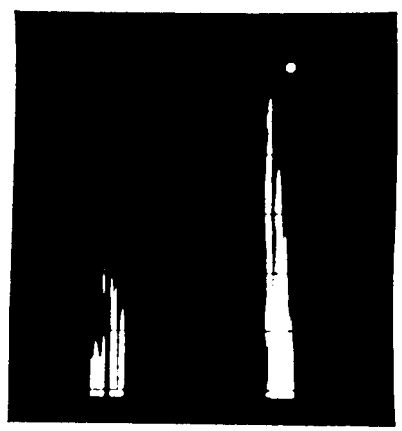

Fig. 4