LEARNING DYNAMICS OF PHOTOREFRACTIVE NEURAL NETWORKS

Thesis by

Yong Qiao

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1994

(Defended June 29, 1993)

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Professor Demetri Psaltis, for his superb guidance, support and constant encouragement throughout my graduate studies at Caltech. His creativity, enthusiasm and technical insights have been my constant inspiration, and his kindness and understanding helped me grow. It has been a privilege to be a member of his highly motivated and talented Optical Information Processing Group.

The work described in this thesis reflects collaboration and discussion with Dr. Ken Hsu, Dr. John Hong, Dr. Claire Gu, Dr. Fai Mok, Sidney Li, and Sergei Orlov. Special thanks go to Dr. Ken Hsu, who taught me a lot of lab skills when I first joined this group. I would also like to thank Dr. Mark Neifeld, Dr. David Brady, Dr. Robert Snapp, Dr. Nabeel Riza, Dr. Cheol-Hoon Park, Dr. Alan Yamamura, Dr. Chuanyi Ji, Dr. Scott Hudson, Dr. Steve Lin, Charles Stirk, Seiji Kobayashi, Subrata Rakshit, Kevin Curtis, Annette Grot, David Marx, Robert Denkewalter, Jean-Jacques Drolet, Geoffrey Burr, Yayun Liu, Jiafu Luo, Xin An, Michael Levene, Allen Pu, and Ernest Chuang for many helpful discussions and their friendship. I am also grateful to Su McKinley and Helen Carrier for their administrative support and kindness.

I would like to thank Dr. Ratnakar Neurgaonkar and the Rockwell International Science Center, who supplied the SBN crystals used in some of my experiments.

I am very grateful to my grandparents Huishan and Baozen Feng. Without their unconditional support and sacrifice, what I have achieved today would still remain a dream. I want to thank my parents Yimou and Jinmei Qiao, and my

sister Jun Qiao for their constant encouragement. Finally, I would like to express my deepest thanks to my wife, Patty, for her love, support, and patience. The arrival of our son, Danny, has made my last year at Caltech more enjoyable.

ABSTRACT

This thesis investigates the optical implementation of neural networks utilizing dynamic photorefractive volume holography. The number of accessible degrees of freedom in a general holographic interconnection system is derived, and a cascaded-grating scheme that provides full, nondegenerate interconnections between two unsampled planes is presented. The dynamics of the formation of photorefractive volume holograms is considered. The impact of time-constant asymmetry on multiple hologram recording is evaluated. A basic framework for controlling the dynamics of photorefractive holograms is described and a number of dynamic copying methods for rejuvenating decayed holograms are identified. Experiments of linear dynamic copying using phase conjugation and nonlinear copying using an optical feedback loop are presented. The electrical fixing of photorefractive holograms in $Sr_{0.75}Ba_{0.25}Nb_2O_6$ crystals is experimentally demonstrated and the physical mechanism is discussed. A number of neural learning algorithms are investigated for optical implementation. An Anti-Hebbian local learning algorithm is proposed to simplify the optical architecture of feedforward multilayer networks. Experimental demonstrations of several optical neural networks are presented. An optical perceptron is trained for face classification, and the use of dynamic copying for improving its performance is demonstrated. A two-layer network based on Kanerva's sparse, distributed memory model is implemented and trained for real-time handwritten character recognition. Finally an optical two-layer network for real-time face recognition, with moderate tolerance to shift, rotation, scale, and facial expression, is presented.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Volume Holography for Information Storage

Volume holography, first invented as an interference method of recording the light waves originated from a subject, represents an effective method of controlling and probing the 3-D optical properties of media. The ability of volume holograms to store information in 3-D with high density and parallel access has made them attractive for various applications including optical information processing. It is based on this technology that the optical learning systems presented in this thesis are built.

Holography was first invented by Dennis Gabor [1] to improve the resolution in images obtained with an electron microscope. Although unable to demonstrate the validity of his principle with electron waves, he was able to do so with visible light, which proved to be a step in the right direction. The inspiration for this two-step, wave-reconstruction microscopy came from the work done by Sir Lawrence Bragg, who invented the "x-ray microscope" [2] for the investigation of atomic arrangement in crystals based on optical Fourier synthesis. After a decade or so, Leith and Upatnieks [3-5] demonstrated off-axis holography that eliminates the twin image effect in Gabor's original in-line holograms. They gave convincing proof that holography was practical and that the use of coherent laser light was an important factor.

Volume holography was invented by Denisyuk [6] to improve the quality of images stored in a hologram. Van Heerden studied volume holograms for information storage, and estimated their storage capacity [7]. The analysis of volume

holograms is different from that of planar holograms in that the wave-coupling effect must be considered. The coupled wave theory was first developed by Kogelnik [8]. Since then, many refinements have been added to Kogelnik's theory, including the treatment of anisotropic effects and composite holograms.

The control of the vast number of degrees of freedom available in the volume has always been a challenge. In fact, the capacity of a holographic memory or interconnection system is usually limited by the recording scheme rather than the capacity of the crystal. Information storage in a volume is achieved by various multiplexing techniques. Spatial and angular multiplexing are the two methods that have been most widely used [9-13]. Recently, with the advent of spatial light modulators (SLMs) that are capable of phase modulation (such as liquid crystal SLMs), phase-code multiplexing has been proposed [14,15] that eliminates problems associated with the physical realization of angular multiplexing. Wavelength multiplexing, although an old concept, is made possible with the advent of wavelength tunable lasers, especially solid-state lasers [16-18].

In this thesis, we are primarily concerned with volume holograms in photorefractive media, which are electro-optic materials in which phase holograms can be recorded through the migration and trapping of photo-generated charges. Typical photorefractive materials used in this work include lithium niobate ($LiNbO_3$), barium titanate ($BaTiO_3$), and strontium barium niobate (SBN). A number of theories have been developed to described the formation of photorefractive holograms [19]. The most commonly used theory is the band transport model developed by Kukhtarev [20]. This theory is reviewed in chapter 2, with emphasis on the dynamical behavior of the formation of photorefractive volume holograms.

The general control of volume holograms was discussed by Brady and Psaltis [21]. This thesis deals with the dynamic control of photorefractive holograms. Several methods are derived and experimentally demonstrated that allow effective use of the whole dynamic range of the crystals. Optical learning systems that are capable of performing a very large number of adaptations can be implemented based on these techniques.

1.2 Artificial Neural Networks and Optical Implementations

Neural networks are massively parallel computers in which a large number of simple processing elements (neurons) are densely interconnected and operate in parallel [22]. Typically, there are several hundred to several thousand connections for each neuron. Therefore, the most difficult part in the practical implementation of neural networks is the realization of the interconnections. If the implementation is a simulation on a digital computer, then the implementation of a large network becomes very time consuming because each connection needs to be realized by a digital multiplication. Moreover, for a large network, there is a serious problem with the storage and retrieval of the weights of interconnections from a mass memory. Hardware implementations can solve these problems through parallelism and the use of a distributed memory in which the weights are stored adjacent to the neurons that they connect. In this way, the logical function of the hardware replicates directly the neural network and the network can be simulated extremely fast. Electronics and optics are the two main approaches for neural network implementation. The advantages of the electronic implementation derive from the fact that it is based on a very mature technology and hence it is possible to easily fabricate

chips that reliably simulate complex neural network functions. The problems with the electronic implementation arise when the size of the network exceeds what can be accommodated on a single chip. For large networks requiring a large number of chips, the problems associated with interconnecting the chips and sequencing the operations properly make the electronic implementation difficult. The optical implementation, on the other hand, is not based on a well established technology which makes the fabriction of optical systems relatively difficult and expensive, and it is more difficult to control their characteristics. The great advantage of the optical implementation derives from the ability to implement interconnections in three dimensions. This allows us to have an architecture consisting of planes of neurons separated by optical systems that implement the connections between the neurons in the planes. The neural planes consist of non-linear optoelectronic processing elements whereas the interconnecting system typically consists of holograms and/or spatial light modulators. The ability to store and process information in three dimensions makes it relatively easy to build very large optical networks, within a relatively small volume and with small power requirements. Typically, a large electronic chip can accommodate 10^4 to 10^5 weighted connections. We will see in this thesis that optical networks with 10^8 to 10^9 connections are readily realizable. Therefore, these optical systems can be functionally equivalent to several thousand chips without the difficulties involved in connecting and synchronizing these chips.

The most important feature of a neural network is learning. Learning is achieved by exposing the network to the problem, and modifying the weights of interconnections and/or the characteristics of the neuron response based on the error it made. In most neural network models the functionality of the network

is determined by the values of the weights of the interconnections. The weights are set during a learning phase in which examples of the function that we wish to implement are presented to the network and its performance is monitored. The weights are then modified to improve the performance. There is a wide variety of weight modification algorithms and almost all of them can be thought of as a variation of the Hebbian law [23]

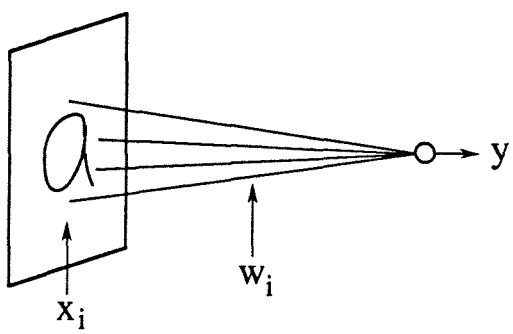
$$\Delta w_{ij} \propto o_i o_j, \tag{1.1}$$

where w_{ij} is the weight connecting the ith and jth neurons and o_i and o_j are the activation functions (typically, nonlinear thresholding functions) of the two neurons. This simple rule not only forms the basis of most neural network learning algorithms, but also has a very direct analogy to holography. Two optical beams, interfering to form a hologram, reinforce the recorded hologram in proportion to the product of their amplitudes. If the hologram implements the interconnection between the ith and jth neurons, then the Hebbian law can be realized simply by allowing light emanating from the two neurons to interfere and modify the interconnecting hologram.

A similar history of development is another common feature of optical and neural computers. Both topics became popular in the late 1950s and early 1960s and both fell out of favor soon thereafter, until recently. The difficulty that both approaches ran into was that they could not provide competitive, practical solutions to interesting problems, despite their promise for superior capabilities. This remains true today.

Nevertheless, there is renewed optimism and the level of interest in these two fields is more intense now than ever before. Whatever the reasons for this

resurgence in interest may be, the remarkable fact is that the interest has been increasing for the last ten years or so. The underlying reason is the transition, in both fields, from the single-layer machines of the early 1960s to the multi-layer machines of today. Most of the work in neural networks in the 1960s focused on a single neuron (see Fig. 1.1), modelled as a simple threshold gate. Each input x_i to the neuron is multiplied by a weight w_i and the weighted signals are summed. If this sum exceeds a threshold, the neuron turns on; otherwise it stays off.

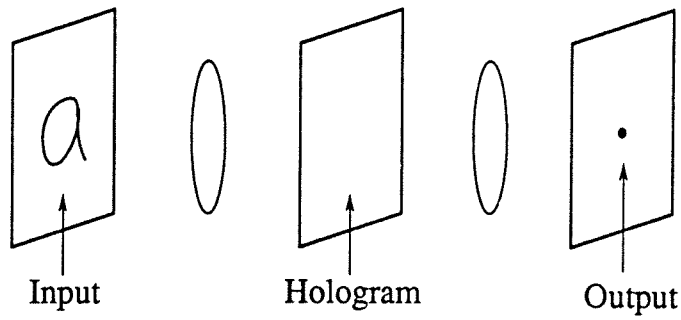


$$y = \mathrm{sgn}\{\sum w_i x_i + \theta\}$$

Figure 1.1. A single neuron modelled as a threshold unit.

What created excitement about this sort of thing was the discovery of algorithms for training such a system. If we are given a set of inputs along with the desired output for each of the inputs, then procedures such as the perceptron [24] and the adaline [25] can iteratively develop a set of neuron weights to map each of the given inputs correctly. Of course, these procedures can produce a set of weights to solve the problem at hand only if such a solution exists. The realization that there are no problems of practical interest that can be solved efficiently by a single neuron essentially brought to an end the early efforts in neural networks.

The most intensively studied optical computing system that originated in that era is the optical correlator [26] (Fig. 1.2). This system correlates an input image with a reference that is stored holographically. The value of the correlation function at the center of the output plane is the weighted sum of the input pixels, the weights being the corresponding pixel values of the reference image. At other positions we obtain the weighted sum of a shifted version of the input image. Therefore, the optical correlator is functionally equivalent to an individual neuron with the additional feature of shift invariance. Optical correlators have not found significant practical applications for the same reason that single neuron systems have not.



$$y_j = \mathrm{sgn}\{\sum w_i x_{i-j} + \theta\}$$

Figure 1.2. Optical correlator.

Research in neural networks in recent years has been focused on the study of large sets of neurons. The discovery of training algorithms for multi-layered feed-forward networks [27,28] is the development that has had the strongest impact in the field. In a two-layer network (Fig. 1.3), the input pattern is first processed by a

set of "hidden" neurons, and the final output is calculated by an addtional neuron which accepts as its inputs the outputs from the hidden units. If we are given enough hidden units, any function can be approximated with arbitrarily good accuracy [29]. Therefore, the question for multi-layer networks is not whether there are any interesting problems they can solve, but whether they can solve some of them more efficiently than other approaches.

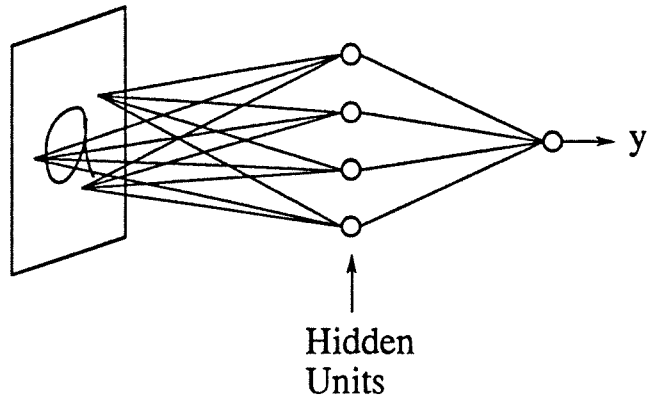


Figure 1.3. Multilayer neural network.

While progress in neural networks occurred primarily through algorithmic developments, in optical computing the most significant development in the 1980s was the maturing of several key optoelectronics technologies. Spatial Light Modulators (SLMs) have always been considered the key component for optical computing systems. An optically addressed SLM is a 2-D array of devices, each device detecting light incident on it, electronically processing the detected signal (typically in a very simple way), and then regenerating an optically modulated signal through a light modulator or a source.

The most significant advances in SLM technology in the past decade came from the development of liquid crystal devices [30,31] and semiconductor optoelectronic devices [32,33]. These advances have made it possible to construct multilayer optical systems in which the output of one layer (or SLM) is optically interconnected to another. The linear portion of the system (the interconnections) is the strength of the optical approach, but the nonlinear, optoelectronic portion provides the system with computational power and diversity. Optical neural networks, digital optical computing [34], and optical interconnections for electronic systems [35] are all examples of this new breed of multilayer optical machines.

The recent activities in the area of optical neural networks began with the optical implementation of a Hopfield network [36,37] in which every neuron is connected to all the rest. Since then, there have been many research efforts in this area, investigating associative memories [38-47], higher-order networks [48-50], methods of learning [51-56], perceptron networks [57-59], self-organizing systems [60], and feedforward multilayer networks [61-63].

The general architecture for optical neural networks used throughout this thesis is shown in Fig. 1.4, which is similar to the optical correlator shown in Fig. 1.2. A plane of sources at the input is connected to a plane of detectors at the output via a volume hologram, and the light sources on a training plane are used together with the input sources to control the hologram. This basic module can be cascaded to form a multilayer network, and feedback can be added as desired. The hologram in this system implements a linear transformation between the input and output plane. The principal theme of this thesis is to develop methods for the dynamic control of this hologram, and use these methods to implement real-time

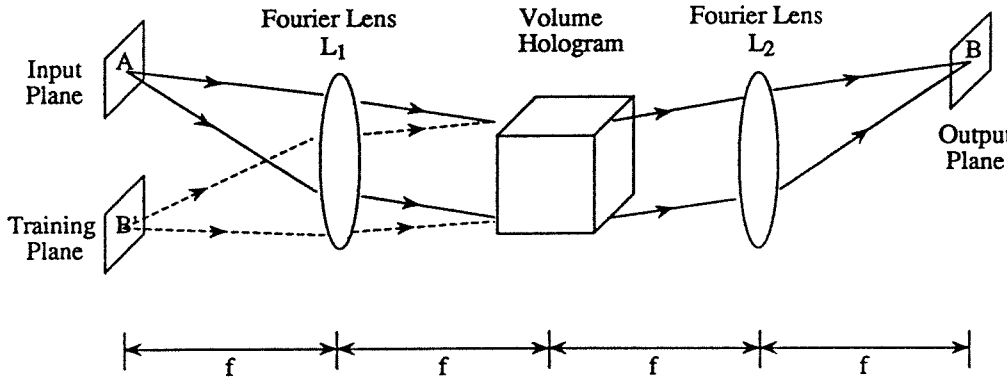


Figure 1.4. General architecture of optical neural networks.

learning in optical neural networks.

1.3 Thesis Overview

The primary goal of this thesis is to develop methods for dynamic control of photorefractive volume holograms and apply them to applications such as optical neural networks and optical memories.

The basic mechanism of 3-D information storage in volume holograms is discussed in Chapter 2. The number of accessable gratings in the optical processing architecture shown in Fig. 1.4 is derived, and a two-layer scheme is proposed that provides full, nondegenerate interconnections between the input and output plane without the need of fractal sampling at these two planes. The physical mechanism of the photorefractive effect is reviewed in the second part of this chapter, and the dynamic equation governing the formation of photorefractive volume holograms is derived. The time-constant asymmetry involved in multiple hologram recording

is analyzed. The effective use of such asymmtry improves the utilization of the available dynamic range of the photorefractive crystal, and therefore enhances the diffraction efficiency of the recorded holograms.

Based on the theory derived in Chapter 2, Chapter 3 describes a general framework for controlling the dynamic behavior of photorefractive holograms. Several types of dynamic copying methods are derived to refresh holograms that have decayed due to read-out or multiple exposures. These methods are analyzed in Section 3.1, and the experimental demonstrations of linear and nonlinear dynamic copying are presented in Sections 3.2 and 3.3, respectively.

While the dynamic copying methods provide a solution to nondestructive read-out of photorefractive holograms, the other approach is permanent fixing. Chapter 4 presents the experimental demonstration of electrical fixing of photorefractive holograms recorded in SBN:75. Electrical fixing of photorefractive holograms is favored in practice because of its simplicity. The experimental results are presented in this chapter, and the physical mechanism involved in this process is discussed. Theoretical predictions based on charge compensation in the band transport model are shown to match closely with the experimental data.

The optical implementation of neural network models depends largely the complexity of the learning algorithms. Chapter 5 investigates different types of learning algorithms that can be implemented optically. The perceptron network is described and its optical implementation, as a basic building block of multilayer optical networks, is presented. The backward error propagation (BEP) algorithm [27] as an example of the algorithms for training distributed-representation multilayer networks is described along with its optical implementation. An Anti-

Hebbian Local Learning (ALL) algorithm is presented that ellimates some of the difficulties encountered in the implementation of BEP-based multilayer networks. As an example of the local-representation networks, the Radial Basis Function (RBF) network and its optical implementation is reviewed. Finally, we study learning algorithms that do not require a fixed network structure but rather determine it as a part of learning. One example is the tiling algorithm, and its optical implementation is discussed.

Chapter 6 is devoted to the experimental demonstration of optical learning networks. An optical perceptron implemented using the dynamic copying technique is presented. This perceptron was trained in real time to classify two classes of human faces, and it demonstrates improved performance over optical perceptron without dynamic copying. A two-layer optical network based on Kanerva's sparse, distributed memory (SDM) model is demonstrated. This system was trained to recognize handwritten character alphabet in real time. The last experimental system is a two-layer network that is capable of recognizing at standard video rate the identity of faces it has been trained for. The faces are presented under a wide variety of conditions to the system, which demonstrates moderate tolerance to shift, rotation, scale, tilt, and facial expressions.

2. OPTICAL INTERCONNECTIONS USING PHOTOREFRACTIVE VOLUME HOLOGRAMS

2.1 Introduction

The basic architecture for each stage of a multilayer optical network is shown in Fig. 2.1. The neurons are arranged in planes, with the $(n-1)$th and nth neural planes being the input and output layers of the nth stage, respectively. Neurons in the input plane are connected to the neurons in the output plane via holographic gratings recorded in a photorefractive crystal. As shown in Fig. 2.1, the light from the ith neuron at the input, with its field amplitude $o_i^{(n-1)}$ representing the response of that neuron, is collimated by a Fourier lens and then diffracted by a holographic grating. The diffracted light is focused by a second Fourier lens onto the jth neuron in the output plane.

An interconnection between the ith neuron in the input plane and the jth neuron in the output plane is formed by interfering $o_i^{(n-1)}$, the light emanating from the ith input neuron, with $t_j^{(n-1)}$, the light emanating from the jth neuron in the training plane. The image of the jth training neuron coincides with the jth neuron in the output plane. The interference of the input signal and the training signal creates a refractive-index modulation grating, with the grating vector $\mathbf{k}_{ji}$ equal to $\mathbf{k}_j - \mathbf{k}_i$, where $\mathbf{k}_i$ ($\mathbf{k}_j$) is the wave vector of the light that is emitted by the i-th (j-th) neuron and collimated by a Fourier lens. The strength (i.e., the weight value) of the interconnection is determined by $A_{ji}^{(n)}$, the amplitude of the refractive index modulation. Later in this chapter it will be shown that $A_{ji}^{(n)}$ is

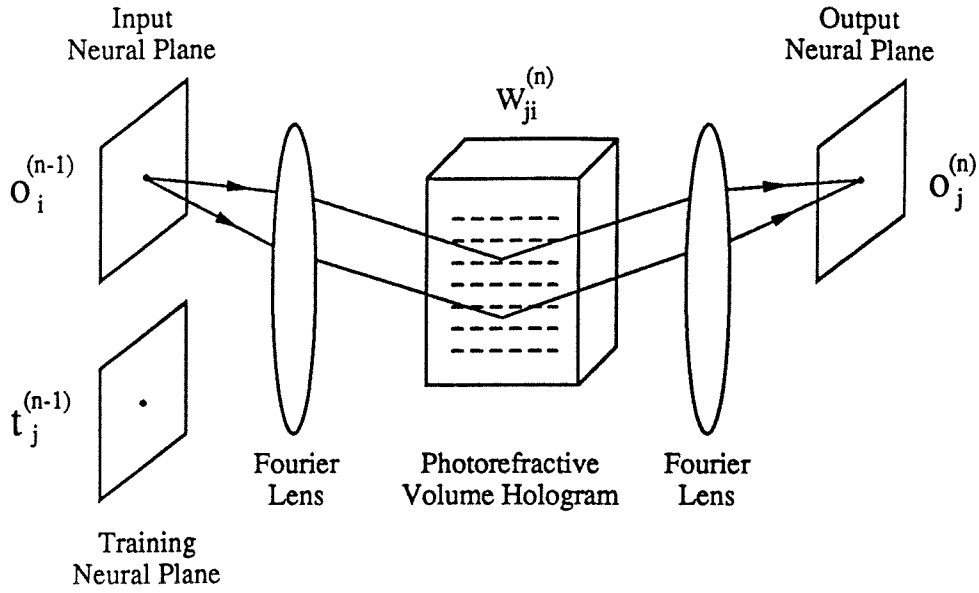


Figure 2.1. Basic architecture for an optical multilayer neural network.

proportional to the modulation depth of the interference pattern:

$$A_{ji}^{(n)} \propto \frac{o_i^{(n-1)} t_j^{(n-1)}}{I_o}, \tag{2.1}$$

where I_o is the total illuminating intensity. This grating diffracts an input beam with wave vector $\mathbf{k}_\alpha$ into an output beam with wave vector $\mathbf{k}_\beta$ if these two beams satisfy the Bragg condition

$$\mathbf{k}_\beta - \mathbf{k}_\alpha = \mathbf{k}_{ji}. \tag{2.2}$$

In section 2.1, we will derive the number of interconnections that can be supported by the system of Fig. 2.1, and apply the result to the design of fully connected optical systems. Section 2.2 describes the physical model for the formation of photorefractive holograms, with emphasis on the first-order dynamic properties.

2.2 Volume Holographic Interconnections

2.2.1 Degenerate Interconnections

We will analyze the capacity of the holographic interconnection architecture of Fig. 2.1 using the same global and local coordinate systems as Ref. 64 (see Fig. 2.2). The light emanating from a pixel at the input (training) plane is converted to a plane wave propagating in the air by the Fourier transforming lens L1 (L2), as shown in Fig. 2.2. This plane wave is refracted at the surface of the crystal, resulting in a plane wave propagating inside the crystal. It is assumed that the angle between the optical axis of lens L1 and the z-direction of the global system is the same as that between the optical axis of the lens L2 and the z-direction, and this same angle is denoted by θ. In addition, all the lenses are assumed to have the same focal length f. Finally, the crystal is assumed to be uniaxial and it is oriented such that its c-axis coincides with the y-direction.

We assume that the area of the input plane is small enough so that the paraxial approximation is valid. Consider a point source at the input plane, with local coordinates (x_i', y_i'). The components of the wave vector $\mathbf{p}_i$ of the corresponding plane wave propagating in the air can be found, in the global coordinate system (x, y, z), as

$$\begin{aligned} p_{ix} &\approx -\frac{k_o}{f} x_i', \\ p_{iy} &\approx \cos\theta(-\frac{k_o}{f} y_i') - \sin\theta k_o, \\ p_{iz} &\approx \sin\theta(-\frac{k_o}{f} y_i') + \cos\theta k_o, \end{aligned} \tag{2.3}$$

where $k_o = 2\pi/\lambda$, with λ being the wavelength in vacuum.

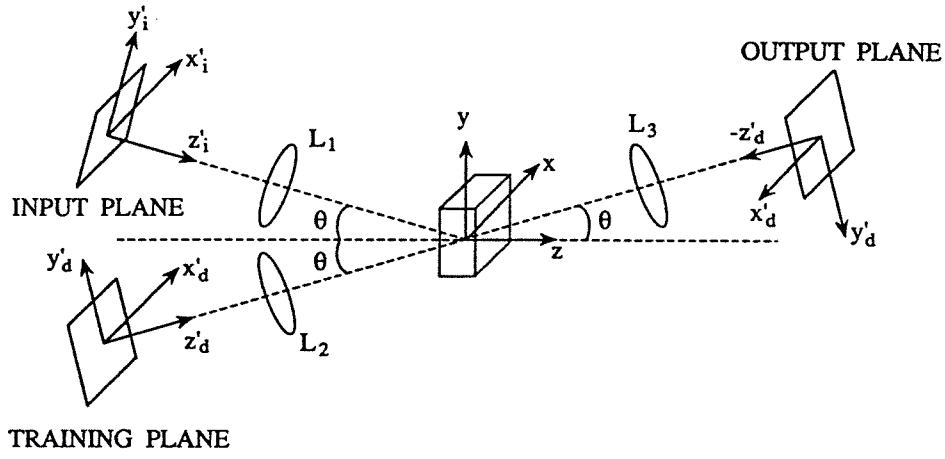


Figure 2.2. The global and local coordinate system.

Similarly, the wave vector of the plane wave originated from a point (x'_d, y'_d) at the training plane is given by

$$\begin{aligned} p_{dx} &\approx -\frac{k_o}{f} x'_d, \\ p_{dy} &\approx \cos\theta(-\frac{k_o}{f} y'_d) + \sin\theta k_o, \\ p_{dz} &\approx -\sin\theta(-\frac{k_o}{f} y'_d) + \cos\theta k_o. \end{aligned} \tag{2.4}$$

The wave vector of the refracted plane wave inside the crystal can be found through the boundary conditions and the normal surface equation of the crystal. We will consider only the extraordinary wave case, since the normal surface associated with the ordinary wave is a sphere, which can be treated as a special case of the ellipsoidal normal surface associated with the extraordinary wave. Let the

wave vectors of the input and training plane waves inside the crystal be denoted by $\mathbf{k}_i$ and $\mathbf{k}_d$, respectively. For $\mathbf{k}_i$, the boundary condition requires that $k_{ix} = p_{ix}$ and $k_{iy} = p_{iy}$. The wave vector $\mathbf{k}_i$ is confined to the normal surface given by [65]

$$\frac{k_{ix}^2}{n_e^2} + \frac{k_{iy}^2}{n_o^2} + \frac{k_{iz}^2}{n_e^2} = k_o^2. \tag{2.5}$$

In the first-order approximation, the components of the input plane wave inside the crystal can be calculated:

$$\begin{aligned} k_{ix} &\approx -\frac{k_o}{f} x_i', \\ k_{iy} &\approx \cos\theta(-\frac{k_o}{f} y_i') - \sin\theta k_o, \\ k_{iz} &\approx n_e k_o \sqrt{1 - \frac{\sin^2\theta}{n_o^2}} - \frac{\sin\theta\cos\theta k_o (n_e/n_o^2)(y_i'/f)}{\sqrt{1 - \sin^2\theta/n_o^2}}. \end{aligned} \tag{2.6}$$

Similar expressions can be found for the training wave

$$\begin{aligned} k_{dx} &\approx -\frac{k_o}{f} x_d', \\ k_{dy} &\approx \cos\theta(-\frac{k_o}{f} y_d') + \sin\theta k_o, \\ k_{dz} &\approx n_e k_o \sqrt{1 - \frac{\sin^2\theta}{n_o^2}} + \frac{\sin\theta\cos\theta k_o (n_e/n_o^2)(y_d'/f)}{\sqrt{1 - \sin^2\theta/n_o^2}}. \end{aligned} \tag{2.7}$$

The grating vector of a grating written by an input point located at (x_i', y_i') and a training point located at (x_d', y_d') is given by the Bragg condition

$$\mathbf{K} = \mathbf{k}_d - \mathbf{k}_i, \tag{2.8}$$

and its components can be found using Eqs.(2.6) and (2.7):

$$K_x \approx k_o(x_i' - x_d')/f,$$

$$K_y \approx \cos\theta k_o(y_i' - y_d')/f + 2k_o\sin\theta, \tag{2.9}$$
$$K_z \approx \frac{\sin\theta\cos\theta k_o(n_e/n_o^2)}{\sqrt{1-\sin^2\theta/n_o^2}}(y_i' + y_d')/f.$$

The mapping between a grating vector and a pair of input and output wave vectors is not one to one. A given grating vector can be mapped to different pair of input and reference wave vectors as shown in Fig. 2.3. This effect, called Bragg degeneracy, indicates that in the system of Fig. 2.2, different pairs of input and output points can be connected by the same grating. Clearly, the characteristics of Bragg degeneracy have to be fully understood before any optical processing systems can be effectively constructed based on the setup of Fig. 2.2.

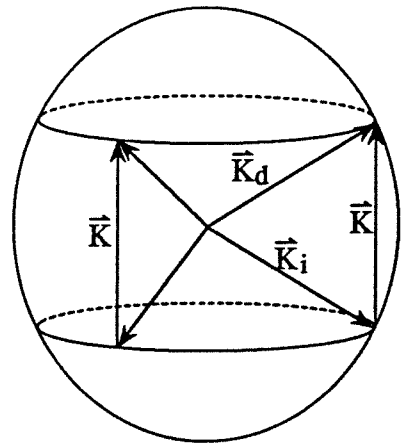


Figure 2.3. Illustration of the Bragg degeneracy.

All pairs of degenerate input-output wave vectors matched by the same grating vector are located on two degeneracy ellipses on the normal surface of the crys-

tal. Specifically, a pair of input-reference wave vectors $(\mathbf{k}_i, \mathbf{k}_d)$ that are matched by a given grating vector $\mathbf{K}$ must satisfy Eq. (2.8) and the following conditions:

$$\frac{k_{ix}^2}{n_e^2} + \frac{k_{iy}^2}{n_o^2} + \frac{k_{iz}^2}{n_e^2} = k_o^2, \tag{2.10}$$

$$\frac{k_{dx}^2}{n_e^2} + \frac{k_{dy}^2}{n_o^2} + \frac{k_{dz}^2}{n_e^2} = k_o^2. \tag{2.11}$$

By subtracting Eq. (2.11) from Eq. (2.10) and substituting $\mathbf{k}_d$ by $\mathbf{K} + \mathbf{k}_i$, we obtain the following equations that consist of only the components of $\mathbf{k}_i$:

$$\begin{gathered} \frac{k_{ix}^2}{n_e^2} + \frac{k_{iy}^2}{n_o^2} + \frac{k_{iz}^2}{n_e^2} = k_o^2, \\ \frac{(2k_{ix} + K_x)K_x}{n_e^2} + \frac{(2k_{iy} + K_y)K_y}{n_o^2} + \frac{(2k_{iz} + K_z)K_z}{n_e^2} = 0, \end{gathered} \tag{2.12}$$

which represent a degeneracy ellipse obtained by cutting the normal surface with a plane. Similarly, the degeneracy ellipse for $\mathbf{k}_d$ is given by:

$$\begin{gathered} \frac{k_{dx}^2}{n_e^2} + \frac{k_{dy}^2}{n_o^2} + \frac{k_{dz}^2}{n_e^2} = k_o^2, \\ \frac{(2k_{dx} - K_x)K_x}{n_e^2} + \frac{(2k_{dy} - K_y)K_y}{n_o^2} + \frac{(2k_{dz} - K_z)K_z}{n_e^2} = 0. \end{gathered} \tag{2.13}$$

The degeneracy ellipses given by Eqs. (2.12) and (2.13) can be mapped to the input and training planes by substituting Eqs. (2.6) and (2.7) into Eqs. (2.12)

and (2.13), respectively. The result is a straight line at the input plane (in terms of local coordinates)

$$(\frac{2K_x k_o}{fn_e^2})x_i' + (\frac{2\cos\theta K_y k_o}{fn_o^2} + \frac{2\sin\theta\cos\theta K_z k_o}{fn_o^2 n_e\sqrt{1-\sin^2\theta/n_o^2}})y_i'$$
$$+ (-\frac{K_x^2}{n_e^2} - \frac{K_y^2}{n_o^2} - \frac{K_z^2}{n_e^2} + \frac{2\sin\theta K_y k_o}{n_o^2} - \frac{2K_z k_o\sqrt{1-\sin^2\theta/n_o^2}}{n_e}) = 0, \quad (2.14)$$

and a straight line at the reference plane (in terms of local coordinates)

$$(\frac{2K_x k_o}{fn_e^2})x_d' + (\frac{2\cos\theta K_y k_o}{fn_o^2} - \frac{2\sin\theta\cos\theta K_z k_o}{fn_o^2 n_e\sqrt{1-\sin^2\theta/n_o^2}})y_d'$$
$$+ (\frac{K_x^2}{n_e^2} + \frac{K_y^2}{n_o^2} + \frac{K_z^2}{n_e^2} - \frac{2\sin\theta K_y k_o}{n_o^2} - \frac{2K_z k_o\sqrt{1-\sin^2\theta/n_o^2}}{n_e}) = 0. \quad (2.15)$$

These two degeneracy lines are almost parallel, since the difference between their slopes is very small when $K_z << K_y$ which is true when the separation distance between the input and reference planes is much bigger than the separation between two input or reference pixels.

Once the degeneracy lines are found for a grating, all the degenerate interconnections corresponding to the same grating can be found. Any pair of input and reference pixels that are picked from the degeneracy lines, with a separation distance of $K_x f/k_o$ in the x-direction, is connected by the grating specifying the two lines. A direct result from this discussion is that any two pairs of input and output pixels that are connected by a same grating should form a parallelogram.

2.2.2 Number of Accessible Gratings

The Bragg degeneracy inherent in volume holograms indicates that in the system of Fig. 2.2, it is impossible to obtain independent connections between the 2-D input plane and the 2-D output plane. The number of independent connections that can be specified or accessed by the system can be found by determining the accessible space of grating vector K. Let us assume that both the input and reference planes are confined within a square, defined as $-a/2 \leq x_i' \leq a/2$, $-a/2 \leq y_i' \leq a/2$, and $-b/2 \leq x_d' \leq b/2$, $-b/2 \leq y_d' \leq b/2$. Without loss of generality, it is assumed that $b \geq a$.

According to Eq. (2.9), the maximum and minimum values of K_x, K_y, and K_z are

$$\begin{aligned}
K_{xmin} &= -k_o \frac{a+b}{2f}, \\
K_{xmax} &= k_o \frac{a+b}{2f}, \\
K_{ymin} &= 2k_o \sin\theta - \cos\theta k_o \frac{a+b}{2f}, \\
K_{ymax} &= 2k_o \sin\theta + \cos\theta k_o \frac{a+b}{2f}, \\
K_{zmin} &= -\frac{\sin\theta \cos\theta k_o (n_e/n_o^2)}{\sqrt{1-\sin^2\theta/n_o^2}} \frac{a+b}{2f}, \\
K_{zmax} &= \frac{\sin\theta \cos\theta k_o (n_e/n_o^2)}{\sqrt{1-\sin^2\theta/n_o^2}} \frac{a+b}{2f}.
\end{aligned} \tag{2.16}$$

According to Eq. (2.16), the value of K_x does not depend on the values of either K_y or K_z. Therefore the cross section of the accessible grating space cut by any plane perpendicular to the K_x-direction is the same. The cross section is

a parallelogram shown in Fig. 2.4, where

$$\begin{aligned} K_a &= \frac{\sin\theta\cos\theta k_o(n_e/n_o^2)}{\sqrt{1-\sin^2\theta/n_o^2}}\frac{b-a}{2f}, \\ K_b &= 2k_o\sin\theta + \cos\theta k_o\frac{a-b}{2f}, \\ K_c &= 2k_o\sin\theta + \cos\theta k_o\frac{b-a}{2f}. \end{aligned} \tag{2.17}$$

The area of the parallelogram is given by

$$\begin{aligned} S =&(K_{zmax}-K_{zmin})(K_{ymax}-K_{ymin}) - (K_{zmax}-K_a)(K_b-K_{ymin}) \\ &- (K_a-K_{zmin})(K_c-K_{ymin}). \end{aligned} \tag{2.18}$$

The shape of the accessible grating space is a parallelepiped bounded by the edge points given in Eq. (2.16), and the volume of this parallelepiped is

$$\begin{aligned} V_a &= S(K_{xmax}-K_{xmin}) \\ &= k_o^3\frac{ab(a+b)}{f^3}\frac{n_e}{n_o^2}\frac{\sin(2\theta)\cos\theta}{\sqrt{1-\sin^2\theta/n_o^2}}, \end{aligned} \tag{2.19}$$

where Eq. (2.16) has been used. In the grating space, the uncertainty volume of a grating is $(2\pi)^3/V_{xtal}$, where $V_{xtal} = L_xL_yL_z$ is the volume of the crystal. The number of accessible gratings of the system of Fig. 2.2 is therefore

$$\begin{aligned} N_a &= \frac{V_a}{(2\pi)^3/V_{xtal}} \\ &= \frac{V_{xtal}}{\lambda^3}\frac{ab(a+b)}{f^3}\frac{n_e}{n_o^2}\frac{\sin(2\theta)\cos\theta}{\sqrt{1-\sin^2\theta/n_o^2}}. \end{aligned} \tag{2.20}$$

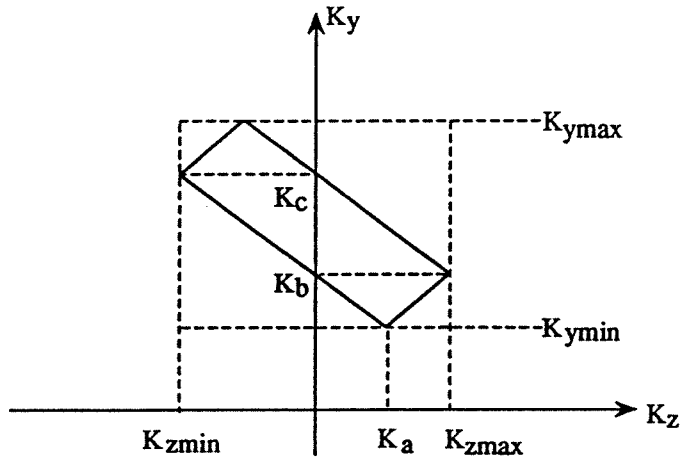


Figure 2.4. The cross section of the accessible grating space cut by a plane perpendicular to the K_x-direction.

Let the numbers of resolvable pixels at the input and reference planes be $N_1 \times N_1$ and $N_2 \times N_2$, respectively. According to Ref. 64, under certain conditions,

$$N_1 = \frac{a}{f\lambda/L_x}, \tag{2.21}$$

and

$$N_2 = \frac{b}{f\lambda/L_x}, \tag{2.22}$$

Equations (2.20), (2.21), and (2.22) yield the following relation:

$$N_1 N_2 (N_1 + N_2) = N_a \frac{L_x^2 \sqrt{1 - \sin^2\theta/n_o^2}}{L_y L_z \sin(2\theta) \frac{n_e}{n_o^2} \cos\theta}. \tag{2.23}$$

This relation shows that N_a, the total number of accessible gratings, is on the order of magnitude of $N_1 N_2 (N_1 + N_2)$. When the input and reference planes

have the same size (i.e., $a = b$), $N_1 = N_2 \equiv N$ and N_a is on the order of magnitude of N^3.

2.2.3 Cascaded-Grating Interconnections

The fact that the number of accessible gratings is on the order of N^3 implies that it is impossible to realize fully independent connections between the $N \times N$ input array to the $N \times N$ output array, since that would require N^4 independent connections. One approach for alleviating this discrepancy is to decrease the density with which the input and output planes are populated. A systematic way of sampling the input and output planes was developed so that the product of the numbers of active pixels at the input and output planes is equal to N^3 [66,67]. These sampled pixels can then be fully connected without Bragg degeneracy.

The drawback of the sampling approach is that the capability of the input and output planes, usually implemented with spatial light modulators, are not fully utilized. The obvious solution to this problem is to increase the number of accessible gratings between the input and output planes. One such method is the cascaded-grating approach proposed and demonstrated by Owechko [68]. The basic idea of the cascaded-grating approach is to use a set of angularly and spatially multiplexed gratings, rather than a single grating, to store each interconnection. By forcing a light beam to match the Bragg condition at each of a cascaded series of spatially and angularly distributed gratings (Fig. 2.5), Bragg degeneracy induced crosstalk is greatly reduced. The **k**-space illustration of this approach is shown in Fig. 2.6. It is clear that two gratings in series connect only a single input/output pair of beams through an intermediary diffracted beam. All other

pairs of input/output beams are not able to match the Bragg conditions at both gratings because the intermediary diffracted beam does not lie on the degeneracy ellipse of the second grating. An undesired beam on the degeneracy ellipse of the first grating is therefore rejected by the second grating. In Owechko's experiment, the cascaded gratings were generated through photorefractive fanning.

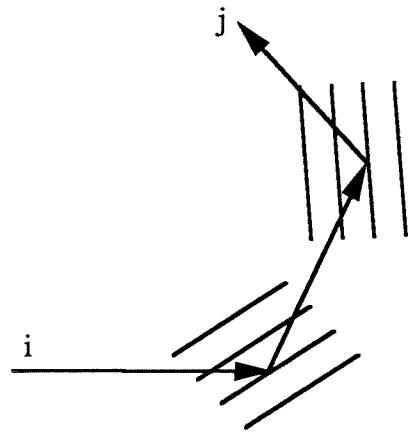


Figure 2.5. Cascaded gratings for eliminating Bragg degeneracy.

In Fig. 2.5, we can imagine that the input beam is diffracted by the first (fanning) grating, and the diffracted beam arrives at an intermediate plane. The light coming out of the intermediate plane is then diffracted by the second grating to generate the final output beam. For a given both the input and output planes have a given dimension of $N \times N$. It is desired to find out what the dimension of the intermediate plane, assumed to be $M \times M$, should be in order to achieve full connections between the input and output plane, which requires N^4 independent

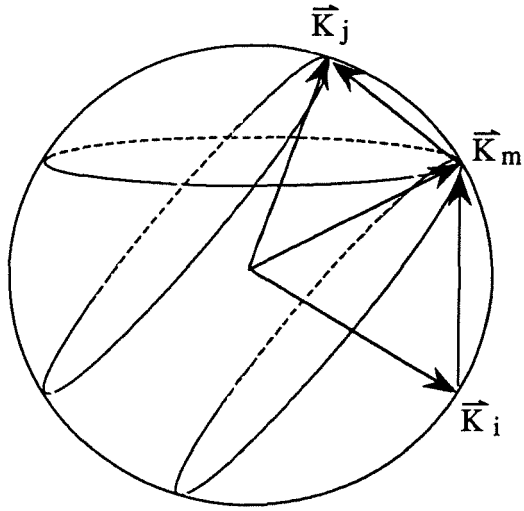


Figure 2.6. **k**-space illustration of the cascaded-grating approach.

connections. Based on the previous analysis of the number of accessible gratings, the number of gratings that can be accessed by the input and intermediate planes (called the first layer connections) is $MN(M+N)$. The number of gratings that can be accessed by the intermediate and output planes (called the second layer connections) is also $MN(M+N)$. Therefore the total number of accessible gratings is on the order of $MN(M+N)$. To have fully connected input and output planes, it is required that

$$MN(M+N) \geq N^4, \tag{2.24}$$

which yields

$$M \sim N^{3/2}. \tag{2.25}$$

The increase in the number of accessible gratings is achieved by using a larger intermediate plane, which enables the access of more gratings. It should

be noted that all the previous analyses are based on paraxial approximation, and there is a limit to which the size of the intermediate plane can be increased.

The fact that the number of accessible gratings between the input and output planes is greater than the number of required connections between these two planes does not necessarily guarantee that the two planes can be fully connected. Fortunately, it turns out that when the condition given by Eq. (2.25) is satisfied, there exists a scheme that realizes full connections between the input and output planes. The intermediate plane will be sampled as shown in Fig. 2.7a, where only one out of every N columns will be utilized. There are a total of $N^{1/2}$ $(=N^{3/2}/N)$ such columns and therefore the total number of sampled pixels in the intermediate plane is $N^{3/2} \cdot N^{1/2} = N^2$. These N^2 pixels can be fully connected to the N^2 pixels at the output plane without Bragg degeneracy, since there are no two sampled pixels at the intermediate plane and two pixels at the output plane that can form a parallelogram.

The task of the first-layer interconnections from the input plane to the intermediate plane is simply to perform a fixed mapping from the N^2 input pixels to the N^2 intermediate pixels. The only requirement is that the interconnection matrix must be nonsingular. In fact degenerate interconnections are allowed because the modification of the input-output connections can be achieved in the second layer. The simplest case is a one-to-one mapping, which can be achieved by using a planar hologram which consists of $N^{1/2}$ multiplexed gratings. As shown in Fig. 2.7b, the first grating will connect the first pixel of the first column of the input plane to the first pixel of the first column of the intermediate plane. Because of shift invariance of the planar hologram, the first column of the input plane is

connected to the first N pixels of the first column of the intermediate plane. In a similar fashion, the second grating will connect the second column of the input layer to the $(N+1)$-th pixel through the $(2N)$-th pixel of the first column of the intermediate plane. This construction continues until the first $N^{1/2}$ columns of the input plane are connected to the first column of the intermediate plane. Then next $N^{1/2}$ columns of the input plane are connected to the second column of the intermediate plane by a set of $N^{1/2}$ gratings. This process is repeated until all the input pixels are connected, in a one-to-one fashion, to the sampled pixels at the intermediate plane. Here we described an approach using a multiplexed planar hologram for achieving the first layer connection. It can be envisioned that a specially designed holographic optical element (for example, using binary optics) can do the job more efficiently.

An advantage of using the above described cascaded-grating scheme over Owechko's fanning based method is that the crosstalk due to Bragg degeneracy can be reduced to a greater extent. In the fanning based approach, the input, intermediary, and output beams tend to lie in the same plane. As the result, the degeneracy ellipses of these two gratings tend to be tangential with each other. If the input beam changes slightly along its degeneracy ellipse, the intermediary diffracted beam will not move a lot away from the degeneracy ellipse of the second grating so that it will not be completely rejected by the second grating. This results in crosstalk. In the cascaded-grating scheme described here, the input, intermediary, and output beams are deliberately designed so that they do not lie in the same plane. Therefore, as shown in Fig. 2.8, the two degeneracy ellipses can be made perpendicular with each other at their intersection. This maximizes the

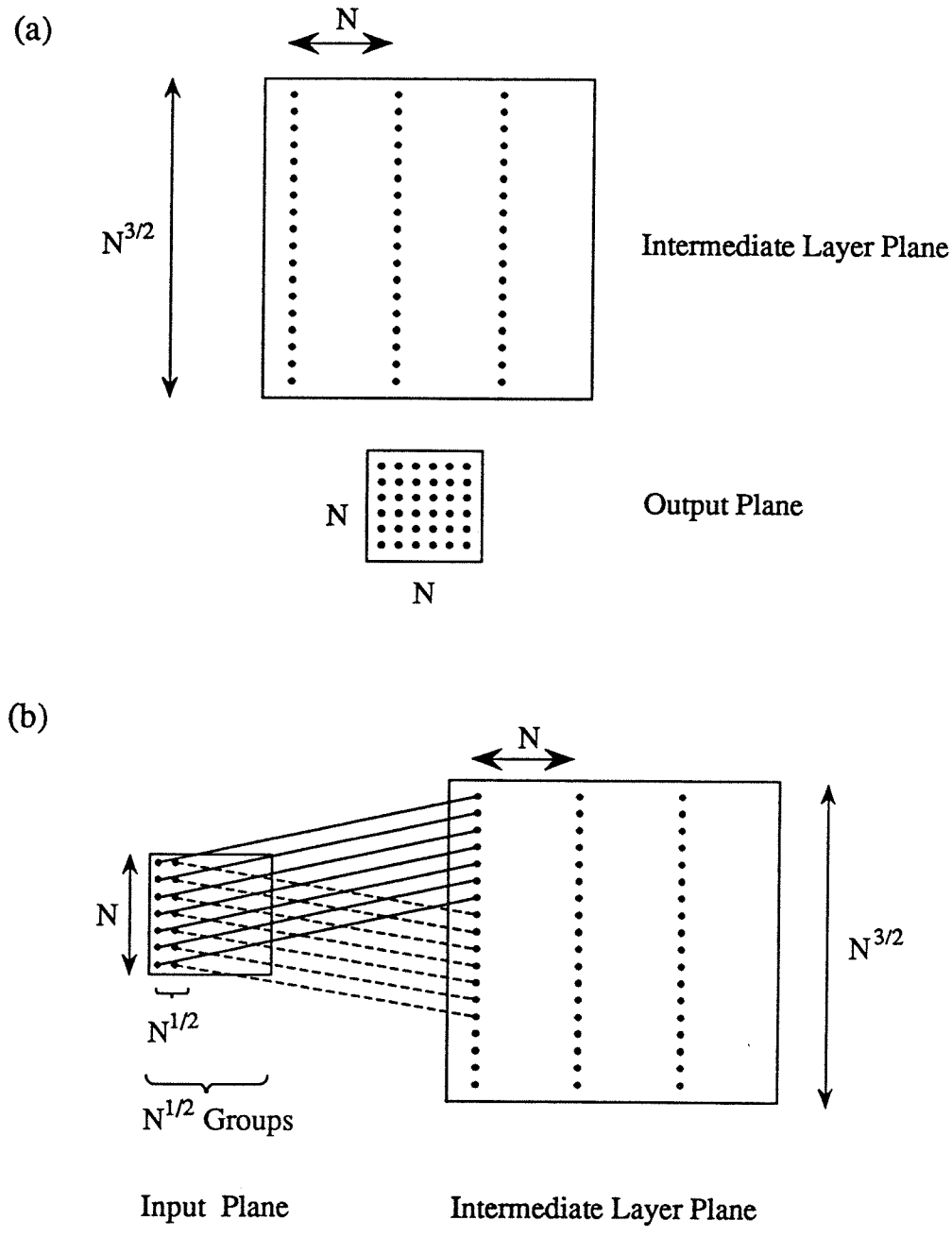


Figure 2.7. Two-layer holographic interconnection scheme that provide full, non-degenerate connections between the input and output planes. (a) The connections between the intermediate and output planes; (b) The connections between the input and intermediate planes.

sensitivity with respect to the intermediary diffracted beam, and thus the crosstalk due to Bragg degeneracy is minimized.

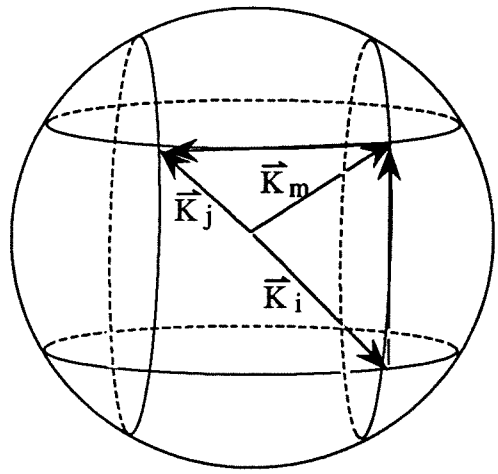


Figure 2.8. **k**-space illustration of the two-layer interconnection scheme.

2.3 Hologram Formation in Photorefractive Media

2.3.1 The Band Transport Model

The photorefractive effect describes the light intensity dependent refractive index observed in photoconductive electro-optic materials. In 1979, Kukhtarev *et al.* [20] published the band transport model. The predictions of this model have been verified experimentally by groups all over the world working on this subject, making this one of the most important theories so far.

The band transport model assumes that photo-excited electrons (or holes) are rejected from filled donor (or acceptor) sites to the conduction (or valence) band. Then they migrate to dark regions in the crystal before combining into empty donors (acceptors). The charge separation results in a space-charge field which modulates the refractive index via the linear electro-optic effect. As the result, a volume hologram, i.e., the three-dimensional pattern of refractive index modulation, is created in the photorefractive crystal.

In the simplest case, the hologram formation in a photorefractive crystal can be modeled as the migration of a single species of charge carriers (which are assumed to be electrons in the following discussion) from neural local trapping sites to ionized sites. In this case, the band transport model is described by the continuity equation for the mobile conduction band electron

$$\frac{\partial n}{\partial t} = \frac{\partial N_D^+}{\partial t} + \frac{1}{e}\vec{\nabla} \cdot \vec{J}, \tag{2.26}$$

the continuity equation for the immobile ionized donors

$$\frac{\partial N_D^+}{\partial t} = (\beta + sI)(N_D - N_D^+) - \gamma n N_D^+, \tag{2.27}$$

the current equation

$$\vec{J} = e\mu n\vec{E} + \mu k_B T\vec{\nabla} n + \vec{J}_{ph} \tag{2.28}$$

and the Poisson equation

$$\vec{\nabla} \cdot \epsilon\vec{E} = e(N_D^+ - N_A^- - n), \tag{2.29}$$

where

n is the free electron number density

N_D is the total density of donors

N_D^+ is the ionized donor number density

$\vec{J}$ is the current density

e is the electron charge

β is the the donor thermal ionization rate

s is the photoexcitation constant

I is the illuminating intensity

γ is the recombination constant

$\vec{E}$ is the electric field

μ is the electron mobility

k_B is the Boltzman's constant

T is the temperature

$\vec{J}_{ph}$ is the photovoltaic current

ϵ is the static permittivity

N_A^- is the number density of ionized acceptors.

The immobile acceptors are assumed to be completely ionized and not to take part in the photoexcitation process. Therefore, N_A^- is a constant independent of time and space.

We will consider hologram formation in the system of Fig. 2.2. We will ignore cross gratings recorded between signals generated on the same SLM, since they are weak due to large grating spacings [20]. The recording intensity is

$$I_{\vec{r}} = I_o(1 + \sum_{ij} m_{ij} cos(\vec{K}_{ij} \cdot \vec{r} + \phi_{ij})). \tag{2.30}$$

m_{ij} and ϕ_{ij} are the modulation depth and phase difference between the ith input beam and the jth training beam. I_o is the average recording intensity. The wave-mixing effect is ignored here, since they can be suppressed by using ordinary polarized recording beams. Holographic recording with coupled beams was treated in [69].

The variables n, N_D^+, $\vec{E}$, and $\vec{J}$ can be expanded in Fourier series in harmonics of the wave vectors $\vec{K}_{ij}$. Assuming $m_{ij} << 1$ and using the quasi-steady approximation [70], the first-order dynamic equation is obtained [71]:

$$\tau \frac{dE_{ij,1}}{dt} = -E_{ij,1} + m_{ij} e^{j\phi_{ij}} E_s, \tag{2.31}$$

where $E_{ij,1}$ is the amplitude of the first-order Fourier component of the space-charge field; τ is the photorefractive time constant given by

$$\tau = \frac{N_A^-}{sI_o N_D} \frac{E_{o||} + j(E_{d,ij} + E_m)}{E_{o||} - \frac{N_A^-}{N_D} E_{ph} + j(E_{d,ij} + E_q)}, \tag{2.32}$$

and

$$E_s = -\frac{1}{2} \frac{E_{o||} + E_{ph} + jE_{d,ij}}{1 - j(E_{o||} + \frac{N_A^-}{N_D} E_{ph} + jE_{d,ij})/E_q}. \tag{2.33}$$

The parameters in Eqs. (2.32) and (2.33) are defined as follows

$$E_{o||} = \frac{\vec{E}_o \cdot \vec{K}_{ij}}{|\vec{K}_i j|}$$

$$E_{ph} = \frac{\vec{\kappa} \cdot \vec{K}_{ij} \alpha_o \gamma N_A^-}{|\vec{K}_{ij}| e \mu s N_D}$$

$$E_m = \frac{\gamma N_A^-}{\mu |\vec{K}_{ij}|}$$

$$E_q = \frac{eN_A^-}{\epsilon|\vec{K}_{ij}|}$$

$$E_{d,ij} = \frac{k_B T|\vec{K}_{ij}|}{e}$$

where $\vec{E}_o$ is the externally applied electric field, $\vec{\kappa}$ is the Glass constant [72], and α_o is the mean absorption constant.

The space-charge field modulates the index of refraction of a photorefractive crystal via the linear electro-optic effect. Specifically, the corresponding perturbation to the permittivity tensor is

$$\Delta\epsilon = -\epsilon \mathbf{r} \vec{E} \epsilon, \tag{2.34}$$

where $\mathbf{r}$ is the electro-optic coefficient tensor. This also implies that the amplitude of refractive index modulation is proportional to $E_{ij,1}$.

2.3.2 Multiple Hologram Recording

When photorefractive holograms are used in neural networks, each weight update in a network requires that an exposure is made to the crystal. At the end of the training, the resulting hologram has contributions from all of the previous exposures. We now consider the effect of multiple exposures to the strength of the hologram. Let A_m be the amplitude of the mth hologram recorded. After a total of M exposures,

$$A_m = A_o[1 - exp(-\frac{t_m}{\tau})]exp(-\sum_{m'=m+1}^{M} \frac{t_{m'}}{\tau}), \tag{2.35}$$

where A_o is the saturation amplitude of the index modulation recorded in the photorefractive crystal, which is proportional to E_s; t_m is the time of the mth exposure. If we require that the final hologram is a linear sum of the recording signals over all the exposures,which means $A_m = A_{m+1}$ for all m, we obtain

$$[1 - exp(-\frac{t_m}{\tau})]exp(-\frac{t_{m+1}}{\tau}) = [1 - exp(-\frac{t_{m+1}}{\tau})]. \tag{2.36}$$

Here the dependence of τ on the grating vector K_{ij} is ignored; this dependence will be explored later in this chapter.

If we assume that $E_{o||} = E_{ph} = 0$, τ becomes real and the solution to Eq. (2.36), for maximum diffraction efficiencies of the recorded holograms, is given by $t_1 \gg \tau$ and [53,73]

$$t_m = \tau ln(\frac{m}{m-1}), \qquad m > 1. \tag{2.37}$$

With this exposure schedule, the final amplitude of the holograms can beobtained using Eq. (2.35):

$$A_m = A_M = A_o/M, \qquad m = 1, 2, \ldots, M. \tag{2.38}$$

Therefore, the diffraction efficiency of each hologram, which is proportional to the square of the recorded amplitude, decays as M^{-2}. For recording of M holograms, the total exposure time is given by

$$T = \sum_{m=1}^{M} t_m = t_1 + \tau lnM. \tag{2.39}$$

In general, the recording time constant τ_r is not necessarily equal to the erasure time constant τ_e. This asymmetry is created, for example, when the

externally applied field $E_{o||}$ or the photovoltaic field E_{ph} is not equal to zero. In this case, if we still assume that τ_r and τ_e are real, the exposure schedule is found to be [53]

$$t_m = \tau_e / m \tag{2.40}$$

for sufficiently large m. The strength of the holograms becomes

$$\begin{aligned} A_m = A_M &= A_o[1 - \exp(\frac{-\tau_e}{M\tau_r})] \\ &\approx \frac{\tau_e}{\tau_r}\frac{A_o}{M} \end{aligned} \tag{2.41}$$

for all m. For example, if $\tau_e/\tau_r = 5$, there will a gain of 25 in the diffraction efficiency of the multiplexed holograms.

In addition to the writing-erasing asymmetry associated with each hologram, there is another type of asymmetry which results from different behaviors for different holograms. For example, in angular multiplexing, the angle between the signal and reference beams changes. Therefore each hologram has a distinct spatial frequency which, according Eqs. (2.32) and (2.33), gives rise to different time constant and saturation amplitude for different holograms. Figure 2.9 shows the dependence of time constant on the angle between the writing beams, for $LiNbO_3$, $BaTiO_3$, and SBN, which shows that the time constant could vary by a factor of 10. A proper utilization of this effect can improve the final diffraction efficiency. In what follows, the effect of varying time constant in angular multiplexing is analyzed.

Assume we want to record M angularly-multiplexed holograms. Let A_m^s, A_m^f, and τ_m be the saturation amplitude, final amplitude, and time constant

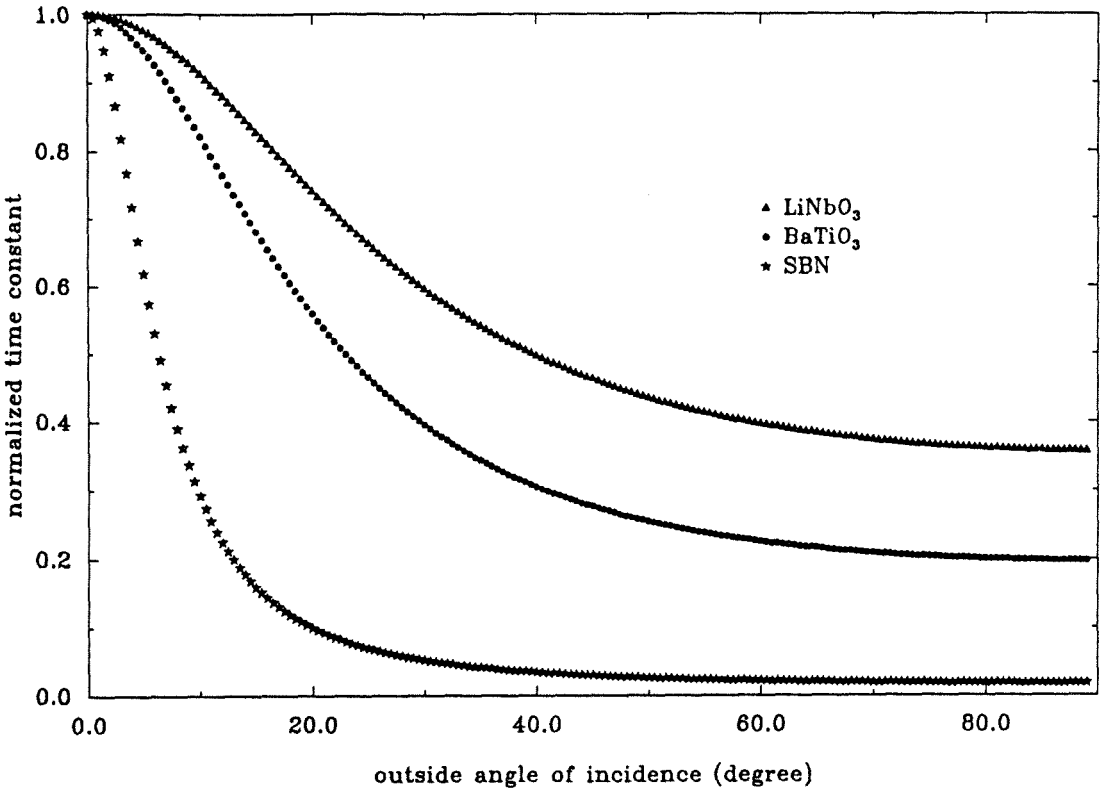


Figure 2.9. Dependence of time constant on the recording angle.

for the mth hologram, respectively. The exposure time for the mth hologram is denoted as t_m, and we define

$$T_m \equiv \sum_{i=2}^{m} t_i = t_2 + t_3 + \cdots + t_m, \quad m = 2, \cdots, M. \tag{2.42}$$

For convenience, we define

$$T_1 \equiv 0. \tag{2.43}$$

Now the question is how to find an exposure schedule $t_2, t_3, \cdots, t_M$ such that

$$A_m^f = A_1^f, \qquad m = 2, 3, \cdots, M. \tag{2.44}$$

The best strategy for the first hologram is, as usual, to make $t_1 \gg \tau_1$ so that the hologram reaches the saturation amplitude A_1^s. Equation (2.44) can then be written as

$$A_m^s(1 - e^{-t_m/\tau_m})e^{-(T_M - T_m)/\tau_m} = A_1^s e^{-T_M/\tau_1}, \tag{2.45}$$

which yields

$$T_m = \tau_m \; ln\Big[\frac{A_1^s}{A_m^s} e^{T(\frac{1}{\tau_m} - \frac{1}{\tau_1})} + e^{T_{m-1}\frac{1}{\tau_m}}\Big], \qquad m = 2, 3, \cdots, M. \tag{2.46}$$

where we have used the relationship

$$t_m = T_m - T_{m-1}. \tag{2.47}$$

The $M - 1$ unknown variables $T_2, T_3, \cdots, T_M$, can be found by solving Eq. (2.46), which consists of a total of M-1 iterative equations. They are solved numerically since closed-form solutions are difficult to find. Once T_m's are found,

the exposure schedule can be derived from Eq. (2.47). At the end of the scheduled recording, the holograms have the same amplitude, given by

$$A_m = A_1 = A_1^s e^{-T_M/\tau_1}. \tag{2.48}$$

If we define

$$\tau'_m = \tau_m/\tau_1, \tag{2.49}$$

$$T'_m = T_m/\tau_1, \tag{2.50}$$

then Eq. (2.46) becomes

$$T'_m = \tau'_m \; ln\Big[\frac{A_1^s}{A_m^s} e^{T'(\frac{1}{\tau'_m}-1)} + e^{T'_{m-1}\frac{1}{\tau'_m}}\Big], \tag{2.51}$$

The final amplitude of the holograms is given by

$$A_m = A_1^s e^{-T'_M}. \tag{2.52}$$

Next we will consider a special case where the time constant has a linear dependence on the recording angle. This corresponds to the linear regions in the curves of Fig. 2.9. For fixed τ_1 and τ_M, the linear dependence implies that τ_m has to satisfy

$$\tau_m = \tau_1 - \frac{m-1}{M-1}(\tau_1 - \tau_M). \tag{2.53}$$

Then, using the definition given by Eq. (2.49), we obtain

$$\tau'_m = \frac{c(M-m)+(m-1)}{c(M-1)}, \tag{2.54}$$

where

$$c = \tau_1/\tau_M. \tag{2.55}$$

For simplicity, we also assume that the saturation amplitude of these holograms is a constant independent of the recording angle. The saturation amplitudes can then be normalized as

$$A_m^s = 1. \qquad m = 1, 2, \cdots, M. \tag{2.56}$$

Under this assumption, Eq. (2.15) is solved numerically and the result is shown in Fig. 2.10. Figure 2.10 is a log-log plot of the final diffraction efficiency η_M (defined as $\eta_M = A_M^2$) versus M, the number of recorded holograms, for three different c values 1, 3, and 10. For these selected values of c and for the plotted range of M, the diffraction efficiency is of the form

$$\eta_M = \beta/M^2, \tag{2.57}$$

where the values of β are equal to 1, 1.6, and 2.1 for c being 1, 3, and 10, respectively. Therefore the diffraction efficiency is enhanced by a factor of c. To achieve the same final diffraction efficiency of 10^{-6}, the numbers of holograms that can be recorded for these three cases are 1000, 1260, and 1450 for c being 1, 3, and 10, respectively. If this type of asymmetry is combined with the writing-erasing asymmetry discussed earlier, the gain in the number of holograms is even greater.

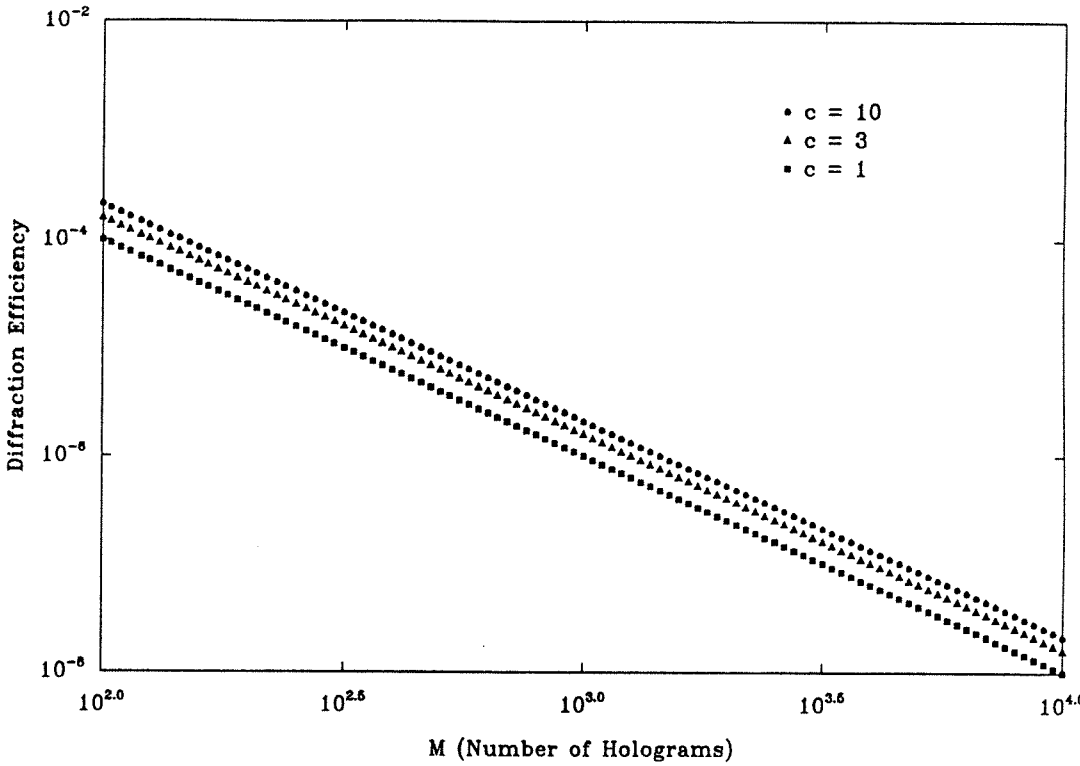


Figure 2.10. Log-log plot of diffraction efficiency versus the number of multiplexed holograms.

3. DYNAMIC CONTROL OF PHOTOREFRACTIVE VOLUME HOLOGRAMS

3.1 Introduction

Volume holograms recorded in photorefractive crystals are attractive for optical computing applications owing to their high storage density and fast parallel access [7]. The dynamics of photorefractive holograms have been extensively studied and are relatively well understood, which enables intelligent control of the hologram formation process and real-time modification of the stored holograms. In addition, the continuing improvement in photorefractive materials makes it possible to record a large number of superimposed holograms with sufficient strength. These advancements have had a great impact in areas such as optical neural networks and optical memories.

In an optical neural network, volume holograms provide an effective way of realizing massively parallel interconnections between the processing elements (neurons) [74]. One of the most interesting features of a neural network is its ability to learn. Learning is achieved by modifying the weights of the interconnections in a supervised or unsupervised fashion. In optical learning networks, this requires that the holographic interconnections can be changed in real time. Photorefractive crystals are the most promising materials for realizing modifiable holographic interconnections [53,57,75]. During each learning iteration, the required weight change is calculated according to some algorithm, and the change is made to the photorefractive hologram by exposing it to light.

For large optical networks, a large number of iterations is required to learn

a desired mapping. This means that many exposures must be made to form the hologram. When a new exposure is made to a photorefractive crystal, the previously recorded hologram decays. If we want all the exposures to make an equal contribution to the final hologram, the exposure schedule described in Chapter 2 has to be followed in which the time for more recent exposures has to be shorter than that for the earlier ones. The consequence is that the diffraction efficiency of the hologram becomes inversely proportional to the number of exposures used to form the hologram, if these exposures are statistically independent. In other words, the diffraction efficiency will rapidly decrease with an increase in the number of learning iterations. Therefore the number of learning iterations that a photorefractive hologram can provide is eventually limited by the background noise in the system. On the other hand, if we use the same time for all the exposures rather than following the exposure schedule, then the contribution of early exposures will be negligible in the determination of the final weights and this may slow down the learning or cause the system not to converge. This dilemma has to be solved before a practical optical neural network can be constructed.

Another important application of photorefractive holograms is optical memories. The ability to store and have parallel access to information in 3-D is the basic driving force for volume holographic storage. To record multiple holograms with the same strength in the same volume, the exposure schedule should again be used. Read-out is generally a destructive process in photorefractive memories because the stored holograms are partially erased by the reading beam, although this decay can be slowed down by using a relatively low-intensity reading beam. One way to achieve nondestructive read-out is by using a reading wavelength to which

the crystal is not sensitive [76]. However, the Bragg selectivity of volume holograms prevents us from reconstructing the recorded holograms with full fidelity using this approach. Another approach is to fix the recorded holograms either thermally or electrically, which will be discussed in detail in the following chapter. While it accomplishes nondestructive read-out, the hologram becomes a write-once memory. In many applications where modification of the stored information is desired, the fixing techniques are not applicable. For a real-time programmable holographic memory, it is usually required that we can selectively erase or rewrite an arbitrary hologram without affecting other stored ones. However, the fact is that during the modification of a hologram, the other holograms sharing the same volume of the crystal will always be partially erased. To summarize, the challenge we are facing is to realize a modifiable photorefractive holographic memory that is nonvolatile to read-out and selective erasure/rewriting.

In what follows, we will describe a general framework [77] for controlling the dynamic behavior of photorefractive holograms. Two types of dynamic copying methods are derived to refresh holograms that have decayed due to read-out or multiple exposures. These two methods are analyzed and their experimental demonstrations are presented.

3.2 General Approach

The problems arising from holographic neural networks and modifiable holographic memories have raised a fundamental question: is it possible to control photorefractive holograms multiplexed within the same volume such that they can be driven to any desired state? We will consider the general holographic setup shown

in Fig. 2.2, and assume that the input plane is an $N \times N$ pixel array and the training plane is a 1-D array of N pixels. Under this assumption, no degenerate interconnections will occur. There can be other arrangements for the dimensions of input and training planes (with proper sampling grids to avoid degenerate interconnections [66,67]), but the discussion presented below will apply in a similar way. By using a set of N mutually orthogonal training (or "reference") beams, holograms can be multiplexed within the volume of the crystal. These reference beams can be plane waves with different angles of incidence (which means only one pixel in the training plane is on for each reference beam), or mutually orthogonal phase-coded waves [14,15].

We first consider the control of a single hologram that is recorded with a input signal and a single reference beam. We assume that the reference beam is fixed and its intensity is much higher than the signal beam. Therefore the total illuminating intensity I_o is approximately the intensity of the reference beam, a constant. Let the amplitude of the signal beam be denoted by a vector $\mathbf{s}$, with its ith element s_i representing the amplitude of the light emanating from the ith pixel in the input plane. Let the complex amplitude of the hologram be denoted by a vector $\mathbf{w}$, with its ith element w_i representing the amplitude of the hologram component recorded as the result of the interference between s_i and the reference beam. Then the first-order dynamic equation describing the formation and decay of the photorefractive hologram is, according to Chapter 2, given by:

$$\frac{d\mathbf{w}}{dt} = \frac{1}{\tau}(-\mathbf{w} + h\mathbf{s}), \tag{3.1}$$

where τ is the photorefractive time constant and h is a constant depending upon the crystal properties, recording geometry, and the amplitude of the reference

beam. We ignore the two-wave mixing effect for the analysis. s_i is the driving signal used to control the hologram component w_i, and there is an upper bound b for the magnitude of s_i so that the modulation depth of the recording field is small enough for the first-order equation (2.1) to be valid [19].

In the terms of control theory [78], this system is controllable meaning that if the initial condition of the hologram is known, there exists a control signal $\mathbf{s}(t)$ such that it will take $\mathbf{w}$ to any desired state at least in the range $|w_i| < |hb|$ (for all i) within a finite time. This tells us that it is possible to make the diffraction efficiency of a hologram independent of the number of exposures that has been made to the hologram. The method to achieve this comes naturally out of the concept of observability of a control system [78]. For our system, the question is how to determine the initial condition $\mathbf{w}^o$, both the magnitude and the phase. It is obvious that the approach is to reconstruct the hologram $\mathbf{w}^o$ and use it to determine the control signal $\mathbf{s}(t)$ which will take $\mathbf{w}$ to the desired state.

For a constant $\mathbf{s}$, the solution of Eq.(3.1) is

$$\mathbf{w}(t) = \mathbf{w}^o exp(-t/\tau) + \mathbf{s}h[1 - exp(-t/\tau)]. \tag{3.2}$$

To drive the hologram to a desired state $\mathbf{w}^*$ at time t_o, the required control signal is

$$\mathbf{s} = \frac{\mathbf{w}^* - \mathbf{w}^o exp(-t_o/\tau)}{h[1 - exp(-t_o/\tau)]}. \tag{3.3}$$

If the goal is merely to enhance the stored hologram, then $\mathbf{w}^* = a\mathbf{w}^o$, where a is an amplification factor. Eq. (3.3) yields

$$\mathbf{s} = \mathbf{w}^{o} \frac{a - exp(-t_o/\tau)}{h[1 - exp(-t_o/\tau)]}. \tag{3.4}$$

Equation (3.4) indicates that $\mathbf{s}$ should be simply equal to the initial state $\mathbf{w}^{o}$ times a constant, suggesting that we reconstruct the original hologram and use the diffracted pattern as the control signal to enhance the hologram. This forms the basic principle of the various "dynamic copying" methods [79-83] that have been extensively investigated for refreshing the stored hologram after its strength is reduced due to read-out or multiple exposures. Dynamic copying in general is to transfer the primary hologram, whenever it becomes weak, to a another medium (i.e., to determine the initial state $\mathbf{w}^{o}$), and then copy the hologram back to the original crystal using the amplified reconstructed light as the writing signal (i.e., to use a control signal $\mathbf{s}$ that is proportional to the initial state $\mathbf{w}^{o}$).

As a first attempt to dynamic copying, Brady and co-workers successfully demonstrated a system in which the decrease in the diffraction efficiency of a multiply exposed hologram is recovered by periodic copying between a photorefractive crystal and a thermoplastic holographic medium [79]. When the hologram recorded in the crystal becomes unacceptably weak after multiple exposures, it is copied to the thermoplastic medium where it is then fixed. By reconstructing the fixed thermoplastic hologram with a strong beam, the hologram is copied back to the photorefractive crystal to rejuvenate the original. This system only recovers the magnitude of the hologram while the phase of the hologram is not maintained. This can cause problems in situations like selective hologram enhancement or erasure where the hologram phases must be well controlled. The phase-tracking problem has been solved by subsequently developed copying methods which will

be described below.

It should be noted that for an optical learning system, we do not have to draw a distinction between the learning phase and the hologram copying phase. Let us assume that during learning, the update of the interconnection weights is calculated to be $\Delta\mathbf{w}$. If the control signal $\mathbf{s}$ is simply equal to $\Delta\mathbf{w}$, then it is impossible for the weights to reach the desired state $\mathbf{w}^* = \Delta\mathbf{w} + \mathbf{w}^o$ because of the hologram decay problem. In this case, dynamic copying has to be performed after the weight update to recover the decreased hologram strength. If, on the other hand, the control signal is determined from both the desired state $\mathbf{w}^*$ and the initial state $\mathbf{w}^o$ using Eq. (3.3), then the weight update will happen properly, and it automatically incorporates the copying since we make use of $\mathbf{w}^o$ to calculate the proper control signal.

In order to specify the N^3 pixels in a 3-D crystal, with on the order of N^2 spatial degrees of freedom for each monochromatic exposure, the recording field must have temporal variations. In the simplest case, we can use N discrete exposures to alleviate this discrepancy. During each of such exposures, only one of the N mutually orthogonal reference beams is used. To enhance N holograms associated with N reference beams, we can simply cycle through the references, incrementally refreshing each hologram in turn by copying, and repeat this cycle many times. We assume that T, the exposure time for each refreshing increment, is the same for all the holograms.

The incremental refreshing of multiple photorefractive holograms can be described by a first-order dynamic equation similar to Eq. (3.1):

$$\frac{d\mathbf{w}_k}{dt} = \frac{1}{\tau}(-\mathbf{w}_k + h\mathbf{s}_k), \tag{3.5}$$

where $\mathbf{w}_k$ denotes the amplitude of the kth hologram (associated with the kth reference), and $\mathbf{s}_k$ denotes the control signal for the kth hologram. $\mathbf{s}_k$ is on only during each refreshing increment for the kth hologram and remains off for the rest of the time. When $\mathbf{s}_k$ is a constant during each refreshing increment, an iterative relationship can be found from Eq. (3.5):

$$\mathbf{w}_k^m = \left[\mathbf{w}_k^{m-1} exp(-kT/\tau) + \mathbf{s}_k^m h[1 - exp(-T/\tau)]\right] exp[-(N-k)T/\tau], \tag{3.6}$$

where $\mathbf{w}_k^m$ is the value of $\mathbf{w}_k$ at the end of mth refreshing cycle, and $\mathbf{s}_k^m$ is the value of $\mathbf{s}_k$ during the mth cycle.

We first consider the case where $\mathbf{s}_k^m$ is a constant throughout the whole copying process and is proportional to $\mathbf{w}_k^o$, the amplitude of the kth hologram before the copying starts. Specifically, we let

$$\mathbf{s}_k^m = a\mathbf{w}_k^o/h, \tag{3.7}$$

where a is again an amplification factor. In this case, the system characterized by Eq. (3.6) will finally reach a steady state $\mathbf{w}_k^s$:

$$\mathbf{w}_k^s = a\mathbf{w}_k^o exp[-(N-k)T/\tau]\frac{1 - exp(-T/\tau)}{1 - exp(-NT/\tau)}. \tag{3.8}$$

If $T << \tau$, Eq. (3.8) becomes

$$\mathbf{w}_k^s \approx a\mathbf{w}_k^o/N, \tag{3.9}$$

which indicates that this incremental refreshing method enhances the multiplexed holograms if $a > N$ and the strengths of the final holograms are still proportional to their initial strengths. Comparing to the single hologram case, the final hologram amplitude is now reduced by a factor of N, the number of holograms. This particular incremental refreshing method is similar to the incremental recording procedure described in Ref. 84. The difference is that the control signal is determined by the initial condition of the multiplexed holograms in the former case and is provided externally in the latter case.

While the copying scheme described by Eq. (3.7) produces a nice result, the physical realization of this method requires a volume holographic medium where all the initial holograms can be stored, fixed, and then selectively reconstructed. This may be possible in some cases, but in general it is difficult to achieve. In the following sections, we will describe some more practical ways of choosing the control signal for incremental refreshing.

3.3 Linear Dynamic Copying – Theory

A more practical way of choosing the control signal $\mathbf{s}_k^m$ for incremental refreshing is to make it proportional to the strength of the kth hologram at the beginning of its turn for refreshing in the mth cycle, i.e.,

$$\mathbf{s}_k^m = a\mathbf{w}_k^{m-1} exp[-(k-1)T/\tau]/h. \tag{3.10}$$

This means that only a planar medium is needed to record, fix, and reconstruct *one* hologram during each refreshing increment. Substituting Eq. (3.10) into Eq. (3.6) and assuming again that $T << \tau$, we obtain

$$\mathbf{w}_k^m \approx (1 - \frac{kT}{\tau} + \frac{aT}{\tau})[1 - \frac{(N-k)T}{\tau}]\mathbf{w}_k^{m-1}$$

$$\approx [1 + (a - N)T/\tau]\mathbf{w}_k^{m-1}. \tag{3.11}$$

Using Eq. (3.11), $\mathbf{w}_k^m$ can be expressed in terms of the initial condition $\mathbf{w}_k^o$:

$$\mathbf{w}_k^m \approx [1 + (a - N)T/\tau]^m \mathbf{w}_k^o. \tag{3.12}$$

Two properties of this copying method can be observed from Eq. (3.12). First, the strengths of all the N holograms remain proportional to their initial condition during the copying and they rise or decay at the same rate. Therefore the copying is linear. Second, in order for the holograms to get enhanced, the gain factor a for the control signal has to be greater than a threshold value N so that the factor $[1 + (a - N)T/\tau]$ in Eq. (3.12) is greater than 1. Otherwise the holograms will eventually decay to zero. In practice, of course, the holograms will not grow forever when a is above the threshold N. Either the copying process will be terminated at a certain time or the hologram strengths will saturate due to the finite dynamic range of the material.

For the linear copying method given by Eq. (3.10), copying can be done using either a 2-D holographic medium in which the recorded hologram can be fixed (for example, a thermoplastic plate) or a spatial light modulator with memory. An even simpler approach is to make the control signal $\mathbf{s}_k$ proportional to the instantaneous hologram amplitude $\mathbf{w}_k$ during the copying, i.e.,

$$\mathbf{s}_k = a\mathbf{w}_k/h. \tag{3.13}$$

In this case, Eq. (3.5) can be rewritten as

$$\frac{d\mathbf{w}_k}{dt} = \frac{1}{\tau}(a-1)\mathbf{w}_k, \tag{3.14}$$

which indicates an exponential growth of each hologram during its refreshing. An expression for $\mathbf{w}_k^m$, the amplitude of the kth hologram at the end of mth refreshing cycle, can be easily derived:

$$\mathbf{w}_k^m = exp[(a-N)mT/\tau]\mathbf{w}_k^o. \tag{3.15}$$

Obviously this copying scheme is also linear and the multiplexed holograms are enhanced as long as $a > N$. The advantage of this method is that there is no need for a copying medium where the reconstructed signal from the primary hologram can be stored for a certain amount of time. This results in a simpler optical system.

3.4 Linear Dynamic Copying – Experiment

3.4.1 Experimental Setup

In this section, the experimental system that realizes the linear dynamic copying method is described. This method allows the composite photorefractive hologram to reach a steady state with overall efficiency independent of M, the total number of exposures made to the crystal. Furthermore, the phases of the recorded holograms remain locked during the copying. The system diagram is shown in Fig. 3.1. The primary hologram is complemented by two phase-conjugating mirrors (PCMs), which are photorefractive crystals in the four-wave mixing configuration. They must share the same pair of pump beams so that the phase-conjugate beams

retain the same relative phase. The basic idea of this system is to record a primary hologram with external beams, read out this primary hologram with the reference beam o_j, and finally copy the hologram that is read out back to the same crystal using the two PCMs. For photorefractive holograms produced by diffusion only, there is a phase shift of $\pi/2$ between the interference pattern and the corresponding hologram. When the reference beam o_j is on, and if the crystal axis is oriented properly, the interference pattern formed by the reference beam o_j and the diffracted beam t_i will create a hologram that is exactly in phase with the original hologram [85]. When these two beams are phase conjugated (to produce the beams o_j' and t_i'), the hologram that the phase conjugate beams create is exactly in phase with the original hologram, and therefore the latter becomes enhanced and sustained.

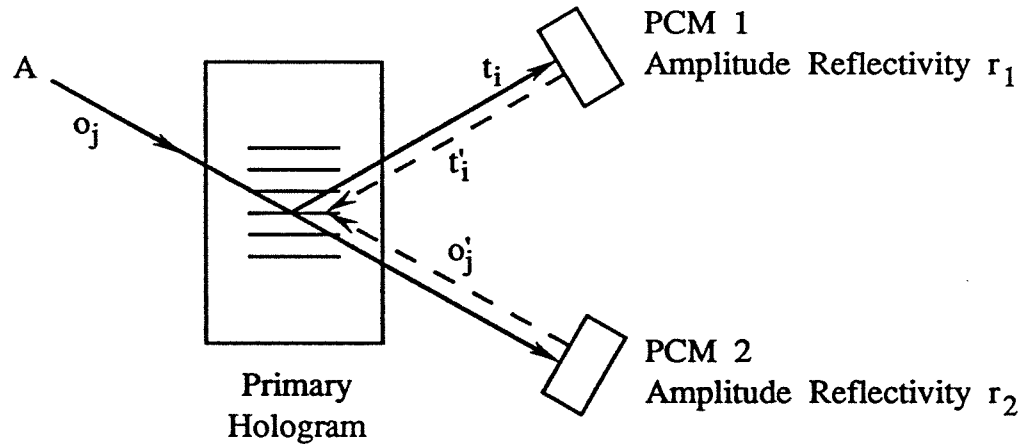


Figure 3.1 Schematic diagram for the linear dynamic copying system with a single reference beam.

3.4.2 System Analysis

The first-order dynamic equation describing the formation and decay of photorefractive holograms is given by Eq. (2.31), which is again written here:

$$\tau \frac{dE_{ij,1}}{dt} = -E_{ij,1} + m_{ij} e^{j\varphi_{ij}} E_s, \tag{3.16}$$

where τ is the photorefractive time constant, and φ_{ij} is the phase difference between the signal and reference beams. m_{ij} is the modulation depth of the interference pattern, given by

$$m_{ij} = 2t_i' o_j' / I_o, \tag{3.17}$$

where I_o is the total illuminating intensity. In this analysis, the two-wave mixing effect in the primary hologram is ignored. For holographic formation by diffusion only, the parameters in Eq. (3.16) are given by

$$\tau = \tau' / I_o, \tag{3.18}$$

$$E_s = -j|E_s|, \tag{3.19}$$

where τ' and $|E_s|$ are real parameters that depend upon the crystal properties and the recording geometry. The fact that these parameters are **real** implies that the phase of the recorded grating will not change if the phase difference between the recording beams remain constant. Since, according to the previous discussion, the phase of the copied hologram is locked , we can ignore the complex nature of Eq. (3.16) and work with the magnitude of $E_{ij,1}$, which is described by

$$\tau \frac{d|E_{ij,1}|}{dt} = -|E_{ij,1}| + m_{ij}|E_s|. \tag{3.20}$$

The amplitude diffraction efficiency of the (ij)th grating is denoted by w_{ij} and it is related to the space-charge field by [71]

$$w_{ij} = sin(\beta|E_{ij,1}|), \tag{3.21}$$

where β depends on the effective electro-optic coefficient of the crystal, the hologram thickness and the recording wavelength.

If we define

$$y_{ij} = \beta|E_{ij,1}|, \tag{3.22}$$

and

$$c = 2\beta|E_s|, \tag{3.23}$$

then a set of simplified equations is obtained:

$$\frac{dy_{ij}}{dt} = \frac{1}{\tau'}(-I_o y_{ij} + ct'_i o'_j), \tag{3.24}$$

$$w_{ij} = sin(y_{ij}), \tag{3.25}$$

where we have used Eqs. (3.17) and (3.18).

We first consider the case of single reference beam with N gratings recorded in the crystal. With the reference beam on (see Fig. 3.1 for illustration), the dynamics of this system are described by Eqs. (3.24) and (3.25), with

$$t'_i = Ar_1 w_{ij}, \tag{3.26}$$

$$o_j' = Ar_2\sqrt{1-\sum_{k=1}^{N} w_{kj}^2}, \tag{3.27}$$

$$I_o = A^2 + A^2 r_1^2 \sum_{k=1}^{N} w_{kj}^2 + A^2(1-\sum_{k=1}^{N} w_{kj}^2)r_2^2. \tag{3.28}$$

A is the real amplitude of the reference beam, and r_1 and r_2 are the amplitude reflectivities of the two PCMs. Substituting Eqs. (3.26)-(3.28) into Eq. (3.24), we obtain

$$\frac{dy_{ij}}{dt} = \frac{r_2^2 A^2}{\tau'}\{-[a+(\rho^2-1)\sum_{k=1}^{N} w_{kj}^2]y_{ij} + c\rho w_{ij}\sqrt{1-\sum_{k=1}^{N} w_{kj}^2}\}, \tag{3.29}$$

where $a = 1 + 1/r_2^2$ and $\rho = r_1/r_2$. In deriving Eq. (3.29), we have assumed that c, τ', and r_1 are all independent of the grating index i. This assumption is valid if the spatial bandwidth of the signal beams is small.

The steady state of the system is obtained by setting $dy_{ij}/dt = 0$ in Eq. (3.29):

$$y_{ij} = \frac{c\rho w_{ij}\sqrt{1-\sum_{k=1}^{N} w_{kj}^2}}{a+(\rho^2-1)\sum_{k=1}^{N} w_{kj}^2}. \tag{3.30}$$

The steady-state diffraction efficiency $w_{ij}^{(s)^2}$ can be solved from the Eqs. (3.25) and (3.30). The existence of a threshold level for the gain factor a of the control signal in the previous discussion is reflected in this experimental system by the fact that ρ must exceed a threshold level in order for the primary hologram to be sustained. Assuming low diffraction efficiencies, a sufficient condition for nonzero steady state is

$$\rho > Max\{a/c, 1\}. \tag{3.31}$$

It can be shown, using straightforward perturbation analysis, that the steady state is stable under this condition. For typical photorefractive crystals, $c = 0.2 \sim 10$. In the case of small c, the steady-state overall diffraction efficiency satisfies the condition $\sum_{k=1}^{N} w_{kj}^2 \ll 1$, which implies that $w_{ij} \approx y_{ij}$ and $\sqrt{1 - \sum_{k=1}^{N} w_{kj}^2} \approx 1$. The latter is actually the undepleted reference approximation. With these approximations, Eq. (3.30) can be solved explicitly, and it yields

$$\sum_{k=1}^{N} w_{kj}^{(s)^2} = \frac{c\rho - a}{\rho^2 - 1} \equiv \eta_1. \tag{3.32}$$

With the approximation $w_{ij} \approx y_{ij}$, Eq. (3.29) also shows that all the gratings rise or decay with the same time constant, which implies that $w_{ij}^{(s)} : w_{kj}^{(s)} = w_{ij}^{(o)} : w_{kj}^{(o)}$ for any i, k, $1 \le i, k \le N$. Here $w_{ij}^{(o)}$ represents the initial value of w_{ij}. So Eq. (3.32) can be rewritten as

$$w_{ij}^{(s)^2} = \eta_1 \frac{w_{ij}^{(o)^2}}{\sum_{k=1}^{N} w_{kj}^{(o)^2}}. \tag{3.33}$$

This property of grating strength normalization is very useful in many applications including neural network implementation since it effectively prevents interconnection weights from either decaying or saturating. If the primary hologram is formed through a sequence of M exposures using the exposure schedule described in chapter 2, then $w_{ij}^{(o)^2} \sim M^{-2}$ for all (ij) pairs. Therefore we can see from Eq. (3.33) that the steady-state diffraction efficiency $w_{ij}^{(s)^2}$ is independent of M. For large values of c, the preceding approximations do not hold and we have to solve Eqs.

(3.25) and (3.30) for the exact steady states. Table 3.1 shows some of the typical parameters and the corresponding steady-state values calculated numerically from Eq. (3.30). For small values of c, $\sum_{k=1}^{N} w_{kj}^{(s)^2} \ll 1$ is justified.

Table 3.1 Typical values for system parameters and steady-state diffraction efficiency in the case of a single reference.

c	r_1	r_2	η_1
0.2	20	2.0	0.76%
0.5	8	2.1	5%
1.0	3	1.9	13%

The approximation $w_{ij} \approx y_{ij}$, used to derive the steady-state solution given by Eq. (3.33), takes into account only the first term in the expansion of the sine function in Eq. (3.25). This approximation, however, is insufficient when the overall diffraction efficiency starts approaching its steady-state value η_1. When that happens, $dy_{ij}/dt \approx 0$ and the higher order terms of the sine expansion cannot be ignored in the dynamic equation (3.29). These higher order terms, according to our model, have an equalizing effect which will lead the system to a final steady state where all the holographic gratings reach the same diffraction efficiency. This same steady-state diffraction efficiency can be found by solving Eq. (3.25) and Eq. (3.30). For large ρ and low diffraction efficiencies, this equalizing process

occurs much slower than the grating normalization process we discussed earlier, so that in practice we usually observe the latter case as a quasi-steady state. Shown in Fig. 3.2 is a numerical simulation of Eq. (3.29) in which the primary hologram consists of two gratings with different initial amplitudes. Initially the ratio of the strengths of the two gratings remains constant until a quasi-steady state is reached. Afterwards, the strengths of the two gratings slowly converge to a common final steady-state value. The above simulation was performed with the following parameters: $\rho = 8, a = 2, c = 0.37, y_{1j}(0) = 0.08, y_{2j}(0) = 0.05$, and $(r_2^2 A^2/\tau') = 1$.

3.4.3 Experimental Demonstration

The experimental system consists of an SBN crystal (1 mm thickness) as the primary hologram and a $BaTiO_3$ crystal for the PCMs (Fig. 3.3). The $BaTiO_3$ crystal, with the c axis oriented 45° from its face, provides phase conjugation for both the reference and diffracted beams. This is done by directing the reference and diffracted beams to two separate regions of the crystal illuminated by the same pair of counter-propagating pump beams, so that the crystal acts effectively as two separate phase-locked PCMs.

The first experiment examines the dynamics of a single grating recorded in this system, where both the signal and reference beams are plane waves. The phase-conjugate reflectivity of the $BaTiO_3$ PCM for the reference beam was set to be 1. Therefore $r_2 = 1$ and $a = 2$. The c coefficient for the SBN was measured to be 0.37. Figure 3.4 shows three experimental curves measuring the changes in diffraction efficiency with time. When the phase-conjugate reflectivity for the

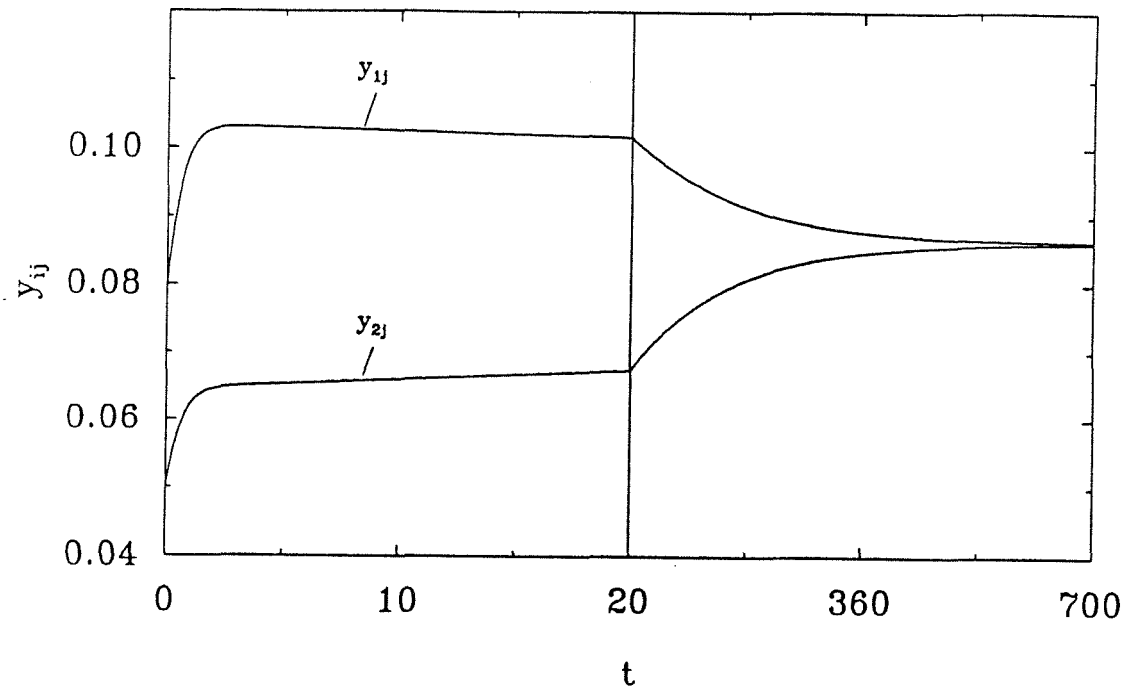


Figure 3.2 Numerical simulation of Eq. (3.29) in which the primary hologram consists of two gratings with different initial amplitudes. The simulation was performed with the following parameters: $\rho = 8, a = 2, c = 0.37, y_{1j}(0) = 0.08, y_{2j}(0) = 0.05$, and $(r_2^2 A^2/\tau') = 1$.

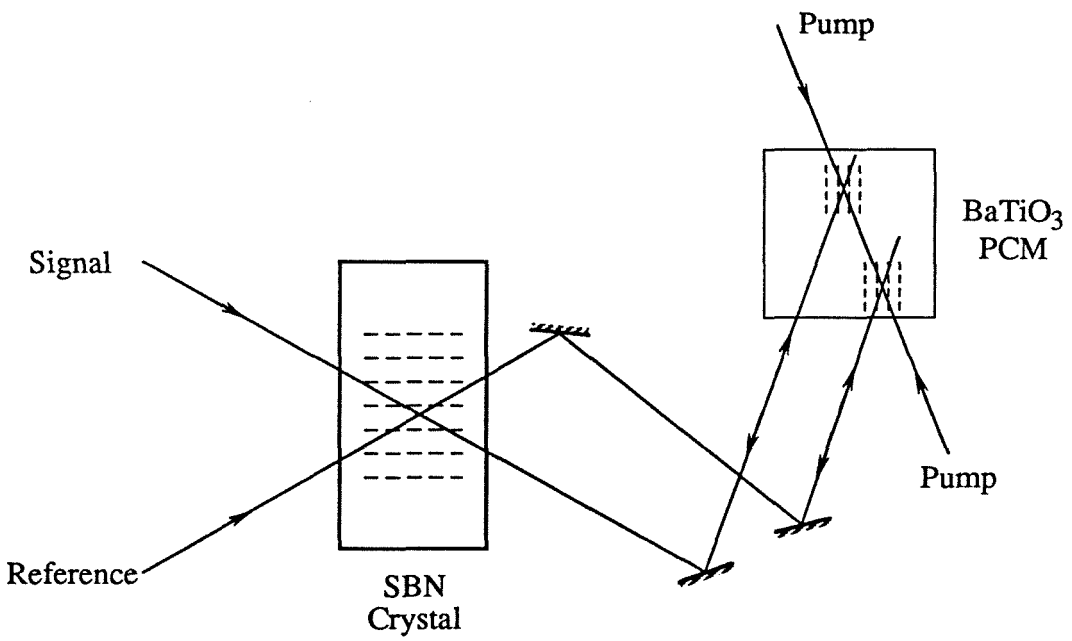


Figure 3.3 Experimental phase-locked linear dynamic copying system.

diffracted signal was 44 (therefore $\rho = r_1 = 6.63$), condition (3.31) was satisfied and the system reached an overall steady-state diffraction efficiency of about 0.845%, independent of the initial condition. For comparison, the theoretical value for the steady-state diffraction efficiency is $\eta_1 = 1.06\%$ from Eq. (3.32). The discrepancy between the experimental and theoretical results may be due to the wave-mixing effect in the SBN and the dependence of phase-conjugate reflectivity on the probe intensity. When ρ was reduced to 2.35, however, the system did not have a nonzero steady state and thus the grating decayed to zero as predicted.

The second experiment investigates the steady-state behavior of multiple gratings recorded in the system. This was done by recording the Fourier transform hologram of an image, which consists of multiple gratings resulting from different spatial frequency components of the Fourier transform. Figure 3.5(a)

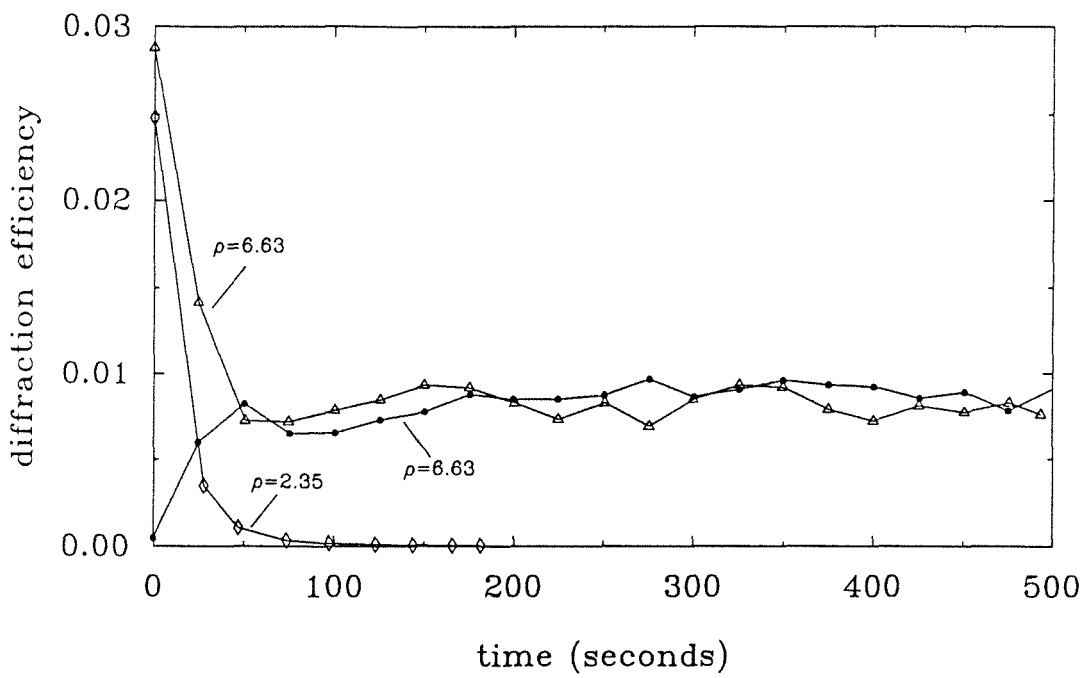


Figure 3.4 Experimental result for the linear dynamic copying system. $r_2 = 1$ in all the experiments. For $\rho = 6.63$, the same steady-state diffraction efficiency is reached when we start with either low ($\bullet$) or high ($\triangle$) diffraction efficiency. For $\rho = 2.35$, the diffraction efficiency decays to zero ($\diamond$).

shows the reconstruction of the image from the SBN when it was first recorded and Fig. 3.5(b) shows the steady-state hologram resulting from the dynamic copying. Although there is some distortion in the steady-state hologram, it can be seen that the grating normalization effect is dominant since all the spatial frequency components are roughly proportional to their initial conditions.

3.4.4 Multiple Reference Beams

To store information in a volume hologram, multiple reference beams are required. For the multiple reference beam case, assuming that there are N plane-wave components in the signal beam and R reference beams, there are two possible ways of sustaining them. One way is to bring in the reference beams cyclically (see Fig. 3.6), and the other is to use mutually incoherent reference beams and have them on simultaneously. Both of these schemes lead to the same steady state. Specifically, assuming small steady-state diffraction efficiency, both of them are described by the same dynamic equation

$$\frac{dw_{ij}}{dt} = \frac{r_2^2 A^2}{\tau'}\{c\rho - Ra - (\rho^2 - 1)\sum_{l=1}^{R}\sum_{k=1}^{N} w_{kl}^2\}w_{ij}, \tag{3.34}$$

with the steady state given by

$$\sum_{l=1}^{R}\sum_{k=1}^{N} w_{kl}^{(s)^2} = \frac{c\rho - Ra}{\rho^2 - 1} \equiv \eta_R. \tag{3.35}$$

Similar to Eq. (3.33), the grating strength normalization relationship also can be found for the multiple reference case:

(a)

(b)

Figure 3.5 (a) The reconstruction of the Fourier transform hologram of an image initially recorded in the SBN crystal. (b) The steady-state response of the hologram stored in the SBN with the initial condition being a hologram of the image shown in (a).

$$w_{ij}^{(s)^2} = \eta_R \frac{w_{ij}^{(o)^2}}{\sum_{l=1}^{R} \sum_{k=1}^{N} w_{kl}^{(o)^2}}. \tag{3.36}$$

From Eq. (3.35), the number of reference beams that can be supported is bounded by

$$R < c\rho/a = \frac{cr_1 r_2}{1 + r_2^2}. \tag{3.37}$$

The right-hand side of Eq. (3.37) reaches its maximum when $r_2 = 1$, in which case

$$R < cr_1/2. \tag{3.38}$$

Some of the typical system parameters and the corresponding number of reference beams that can be supported are shown in Table 3.2.

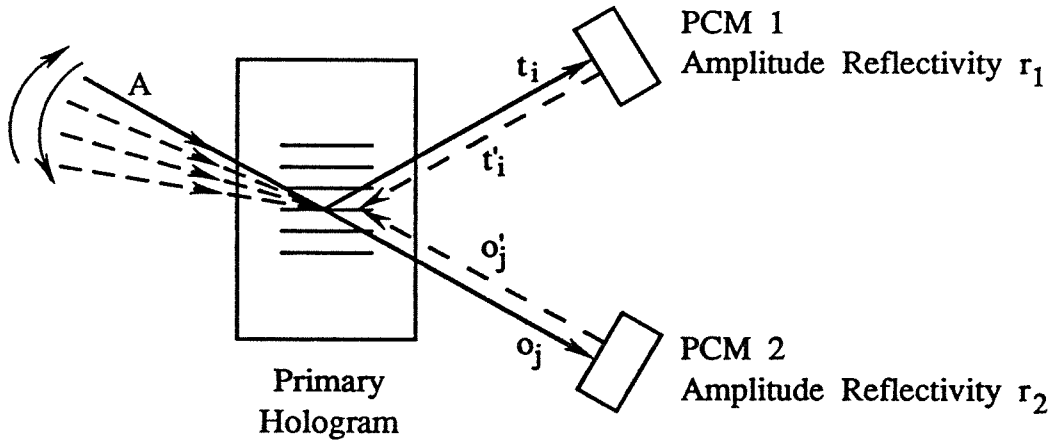


Figure 3.6 Schematic diagram for the linear dynamic copying system with multiple reference beams.

Table 3.2 Typical values for system parameters and number of reference beams that can be supported (r_2=1).

c	r_1	R
0.5	20	5
10	20	100
10	100	500

Another experimental method that falls into the linear copying category was demonstrated by Sasaki *et al.* based on dynamic copying between two functionally identical photorefractive crystals [81]. An optical amplifier was used to provide gain to the reconstructed signal from one crystal which serves as the writing signal for the other crystal.

3.5 Nonlinear Dynamic Copying – Theory

If the primary holograms can be binary images, then there is another method of choosing the control signal which will lead to a simple optical implementation. We consider the case where $\mathbf{w}_k$ can be treated as the (real) magnitude of the hologram. This is justified by the fact that the optical system will maintain the phase of the holograms during the copying (see the following section). In this case the control signal $\mathbf{s}_k$ for the kth hologram is chosen as

$$\mathbf{s}_k^m = G[\mathbf{w}_k^o]/h, \tag{3.39}$$

where $G[\mathbf{x}]$ is a nonlinear thresholding function operating on the individual elements x_i of its vector argument $\mathbf{x}$:

$$G[x_i] = \begin{cases} 0, & x_i < x_o; \\ y_{sat}, & x_i > x_o. \end{cases} \tag{3.40}$$

In Eq. (3.40), x_o is the threshold level and y_{sat}/h is the magnitude of the control signal whenever the corresponding component of the initial hologram exceeds x_o. Following a similar derivation that led to Eq. (3.9), we can find that the system will also reach a steady state where

$$\mathbf{w}_k^s \approx G[\mathbf{w}_k^o]/N. \tag{3.41}$$

Therefore the final holograms are the binarized version of the initial ones. Again the steady-state amplitude of the individual holograms is independent of the number of exposures made to the crystal before copying and is inversely proportional to N, the total number of holograms.

Let the ith component of the hologram $\mathbf{w}_k$ be denoted by $w_{k,i}$. After incremental refreshing, those components with an initial magnitude above the threshold level x_o will reach the same final magnitude of y_{sat}/N. Using Eqs. (3.40) and (3.41), we can write the steady state of those components as

$$w_{k,i}^s \approx a_{k,i} w_{k,i}^o/N, \tag{3.42}$$

where $a_{k,i}$ is an amplification factor determined by both y_{sat} and the initial condition $w_{k,i}^o$:

$$a_{k,i} = y_{sat}/w^{o}_{k,i}. \tag{3.43}$$

If the initial hologram is not binary, Eq. (3.43) indicates that this copying method changes the hologram in a nonuniform fashion. Each of those hologram components is either enhanced or reduced depending upon whether its initial magnitude is less or greater than y_{sat}/N. For those enhanced components, the ones with smaller initial magnitude are amplified more. If the initial hologram is itself binary, then $a_{k,i}$ becomes a constant independent of the subscript i and the hologram is therefore uniformly enhanced as long as the initial magnitude of those components in the on state is less than y_{sat}/N but more than x_o.

3.6 Nonlinear Dynamic Copying – Experiment

3.6.1 Experimental Setup

The experimental system for implementing nonlinear (binary) dynamic copying is shown in Fig. 3.7, which consists of a single photorefractive crystal for the storage of Fourier transform holograms, a spatial light modulator (SLM) to display input images, and a feedback loop with a liquid crystal light valve (LCLV). The LCLV performs image amplification and thresholding. The basic idea is to reconstruct the stored weak hologram, amplify and binarize the reconstructed image with the feedback loop, and rejuvenate the stored hologram by rewriting it with the amplified reconstructed image as the signal beam. Both the input and the signal reconstructed from the hologram are imaged onto the writing side of the LCLV and read out from the other side.

A basic problem with such a feedback loop [86] is that it is inherently unsta-

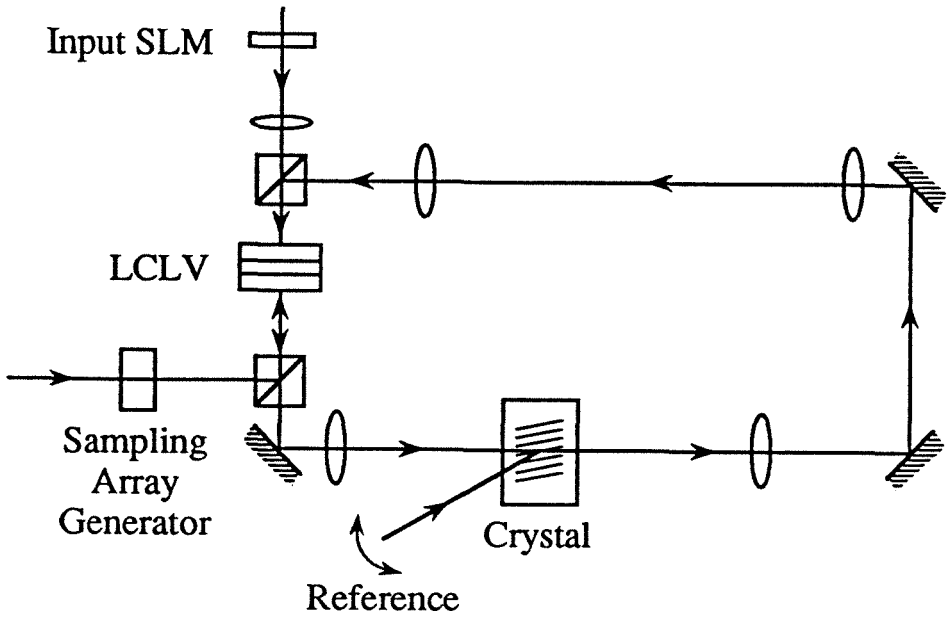


Figure 3.7 Schematic diagram of the experimental nonlinear dynamic copying system with a sampled optical feedback loop.

ble. Any slight misalignment or perturbation in the optical setup will be amplified by the feedback loop, which results in a smeared image and a smeared hologram. A novel feature of this new memory system is that it is stabilized by sampling the image on the reading side of the LCLV with an array illuminator. We used an 128×128 element micro-lenslet array and a 4-f system to generate the sampling light spot array. When the reconstructed image is fed back to the writing side of the LCLV, each pixel is magnified to a larger circle by placing the LCLV slightly before the image plane. Therefore, any misalignment between the input and feedback images will be corrected as long as the readout sampling spot does not miss the magnified pixel incident on the writing side of the LCLV, as shown in Fig.

3.8. There is an obvious trade-off in this system between image resolution and tolerance to misalignment errors.

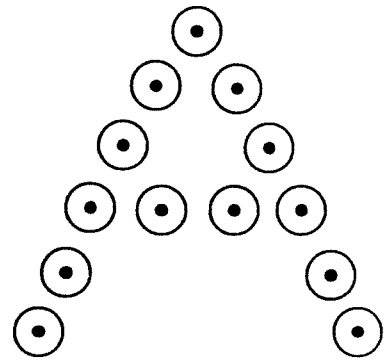


Figure 3.8 Illustration of the spatial sampling technique for stabilizing the feedback loop, where the small dots represent the sampling points on the reading side of the LCLV and the large circles represent the magnified pixels of the feedback image.

An important property of this dynamic memory is that the relative phases of the stored holograms are maintained during copying. This is true because any of the holograms is recorded and copied with the same reference and signal beams that have a fixed phase relationship. In the case of hologram formation by diffusion only (such for the $BaTiO_3$ crystal used in our experiment), the consequence is that the hologram will always remain phase locked. This property is crucial for the realization of selective memory erasure and holography-based optical neural networks, where the phases of the holograms must be accurately controlled.

Multiple holograms are recorded in the system of Fig. 3.7 by changing

the angle of the reference beam with a mirror mounted on a motorized rotary stage. These holograms are dynamically enhanced by scanning the angle of the reference beam, sequentially reading out and rejuvenating each stored hologram, and repeating this cycle many times.

3.6.2 System Analysis

A simplified diagram of the dynamic memory loop is shown in Fig. 3.9, where x_i and y_i represent the field magnitudes of the ith image pixel at the input and output of the LCLV, respectively. A is the field magnitude of the reference beam, and α_1 and α_2 account for losses in the feedback loop. The crystal is oriented so that the optical signal reconstructed by the reference is in phase with the writing signal beam. Such a geometry plus the phase-locking property of this system allow us to use only the magnitudes of all the optical signals for the analysis, similar to the linear copying case. Let w_{ij} be the amplitude diffraction efficiency of the grating corresponding to the ith pixel on the LCLV and the jth plane-wave reference. The response time of the LCLV (approximately 20 milliseconds) is much faster than that of the $BaTiO_3$ crystal (approximately 3 seconds with the light levels used in the experiment). Therefore, the signal in the loop (i.e., the output of the LCLV) will first quickly reach its steady state before there is an appreciable change in the hologram recorded in the crystal. The steady state that the signal in the loop will arrive at depends on the initial diffraction efficiency $w_{ij}(0)$, as shown in the following.

Let the nonlinear input-output characteristic of the LCLV be described by $y_i = G[x_i]$, which can be approximated by a piecewise linear function

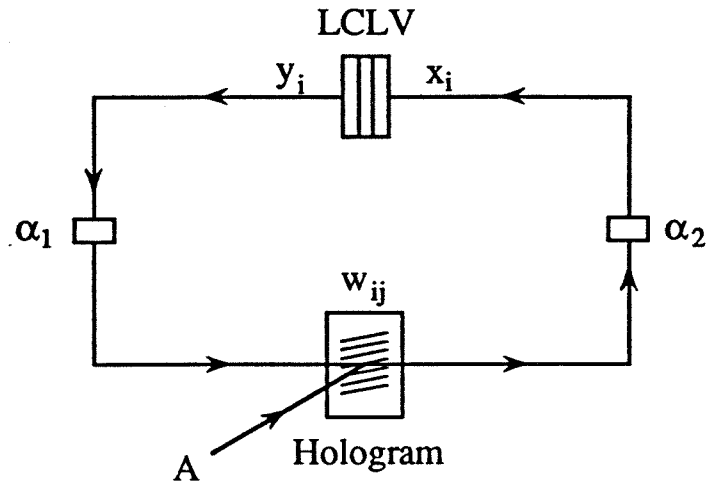


Figure 3.9. A simplified system diagram.

$$G[u] \approx \begin{cases} 0, & u \leq u_0; \\ (u - u_0)g, & u_0 \leq u \leq u_1; \\ y_{sat} \equiv (u_1 - u_0)g, & u \geq u_1, \end{cases} \tag{3.44}$$

where g is the gain in the linear region and $u \geq 0$. The magnitude of the optical field incident on the LCLV is

$$x_i = \alpha_2(\alpha_1 y_i + A w_{ij}(0)). \tag{3.45}$$

Equations (3.44) and (3.45) are illustrated in Fig. 3.10, where we have assumed that $\alpha_1\alpha_2 g > 1$. Eq. (3.45) is plotted for two cases: (a) $\alpha_2 A w_{ij}(0) < u_0$ (dashed line), and (b) $\alpha_2 A w_{ij}(0) > u_0$ (solid line). For case (a) there are three steady states, but only two of them, the on state $y_i = y_{sat}$ and the off state $y_i = 0$, are stable. In this case, the ith pixel will remain in the off state, since the initial

condition for dynamic copying is $(x_i(0), y_i(0)) = (\alpha_2 A w_{ij}(0), 0)$ which is itself the off state. For case (b), there is only one steady state, i.e., the on state $y_i = y_{sat}$. Thus the ith pixel will always go to the on state. To summarize, the pixels on the LCLV has only two stable states: the zero (off) state and the saturation (on) state. This implements the thresholding function $G[\mathbf{x}]$ defined in Eq. (3.40). The condition for switching on the ith pixel on the LCLV is

$$\alpha_2 A w_{ij}(0) > u_0, \tag{3.46}$$

which simply means that the light amplitude incident on the writing side of the LCLV, which results from diffraction off the corresponding grating, must exceed the threshold u_0 of the LCLV. Because of the thresholding effect, the control signal for any particular hologram remains unchanged throughout the whole copying process — it depends only upon the initial condition of the hologram. This realizes the copying method given by Eq. (3.39).

In this system, the cyclical readout and enhancement of the stored holograms lead to a self-driven incremental refreshing. A very useful feature of the self-driven incremental refreshing is that there is no need to provide repeatedly external writing signals during each exposure. Rather, the memory loop itself generates the writing signals by tuning to the appropriate angles of the reference beam, which is a great simplification for practical applications. The final diffraction efficiency of each hologram scales as $1/R^2$ (where R is the number of reference beams), and it does not depend on the number of exposures used to form the initial holograms (before copying).

The dynamics of photorefractive holograms in this system can be described

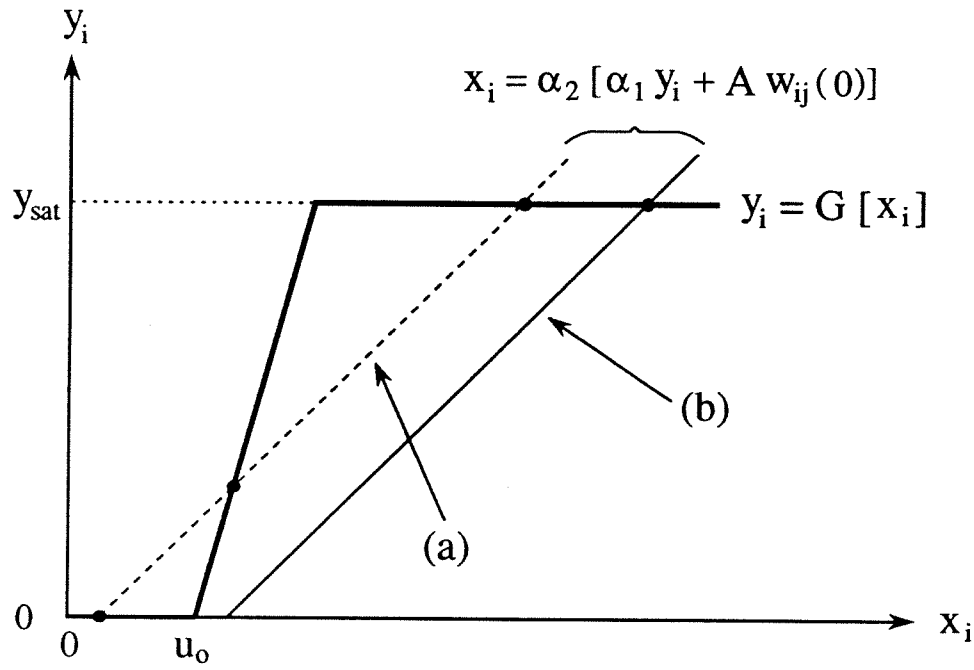


Figure 3.10. Steady state behavior of the dynamic memory.

by the following equation (derived from Eq. (3.20)):

$$\tau dw_{ij}/dt = -w_{ij} + c\alpha_1 y_{sat}/A, \tag{3.47}$$

where c is a unitless parameter determined by the crystal properties and the recording condition. Strictly speaking, Eq. (3.47) is valid only for weak gratings. However w_{ij} scales as $1/R$ and it becomes small when R is very large. In this context, Eq. (3.47) is almost exact. From Eq. (3.47), we find that the steady-state amplitude diffraction efficiency for R recorded holograms is $c\alpha_1 y_{sat}/(AR)$. Using this value for $w_{ij}(0)$ in the switching condition (3.46), we obtain the upper bound for the number of angularly multiplexed holograms:

$$R < c\alpha_1\alpha_2 y_{sat}/u_0. \tag{3.48}$$

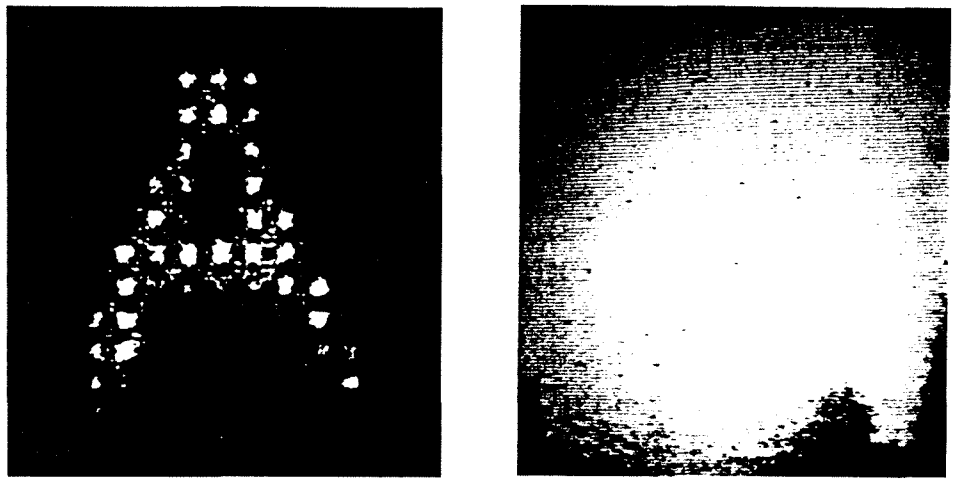
For example, assuming $c = 1, \alpha_1 = \alpha_2 = 0.9$ and $y_{sat}/u_0 = 10^3$, the maximum value of R is 810.

In deriving Eq. (3.48) , we have implicitly assumed that the amplitude of the reference beam for copying is the same as that for recording. This does not have to be so. In fact, we can use a strong reference pulse (with an amplitude A_p that is larger than A) during the reconstruction phase of the copying, so that a smaller $w_{ij}(0)$ would still satisfy the switching condition (3.46). The amplitude of the reference beam returns to its normal value of A once the LCLV output reaches its steady state. By doing this, the upper bound of R can be increased by a factor of A_p/A.

3.6.3 Experimental Demonstration

In the experimental apparatus, a 5 mm cubic $BaTiO_3$ crystal with its c axis oriented 45° from the cut face was used, and the writing beams ($\lambda = 514.5$ nm) were extraordinary polarized. The stabilization of the feedback loop is demonstrated in Fig. 3.11. Fig. 3.11(a) shows the steady-state hologram when the initial one was recorded with a single exposure of the letter A and the sampling was used. The quality of the stored image remained virtually unchanged during the dynamic copying. Fig. 3.11(b) shows the same thing except that the sampling was removed. For the unsampled case, careful alignment simply slowed the transition from a recognizable image to the smeared final state.

A series of experiments were performed to demonstrate the capabilities of



(a) (b)

Figure 3.11 Examples of images dynamically stored in the crystal: (a) with sampling; (b) without sampling.

dynamic enhancement/equalization and selective erasure/rewriting of this system. Ten angularly multiplexed image holograms were first recorded with equal exposure and their diffracted signals were measured with an oscilloscope (shown in Fig. 3.12(a)). As expected, the diffraction efficiencies are highly nonuniform. Dynamic copying was then performed for 27 cycles until these holograms were finally equalized and enhanced (see Fig. 3.12(b)). Next the fifth hologram was erased by introducing an 180 degree phase shift in the corresponding reference beam with an electrically controlled liquid crystal phase retarder and writing a hologram that is out of phase with the one to be erased. The measured diffracted signals from the stored holograms are illustrated in Fig. 3.12(c), which shows that the fifth hologram was almost completely erased and the other nine holgrams were also partially erased. Subsequent dynamic copying rejuvenated the other nine holograms while

the fifth hologram got completely erased because it was too weak to satisfy the switching condition (see Fig. 3.12(d)). Finally, a new hologram was re-recorded at the fifth position, and all the holograms were again dynamically equalized (see Fig. 3.12(e)). The final diffraction efficiency of each hologram is approximately 0.12%.

Another experimental nonlinear copying system, with a feedback loop consisting of a CCD camera and a liquid crystal TV (LCTV) spatial light modulator, was demonstrated recently by Boj *et al.* [83]. The CCD camera detects the reconstructed signal from the weak primary hologram, and the signal is amplified electronically using a computer and then displayed on the LCTV to rewrite the hologram.

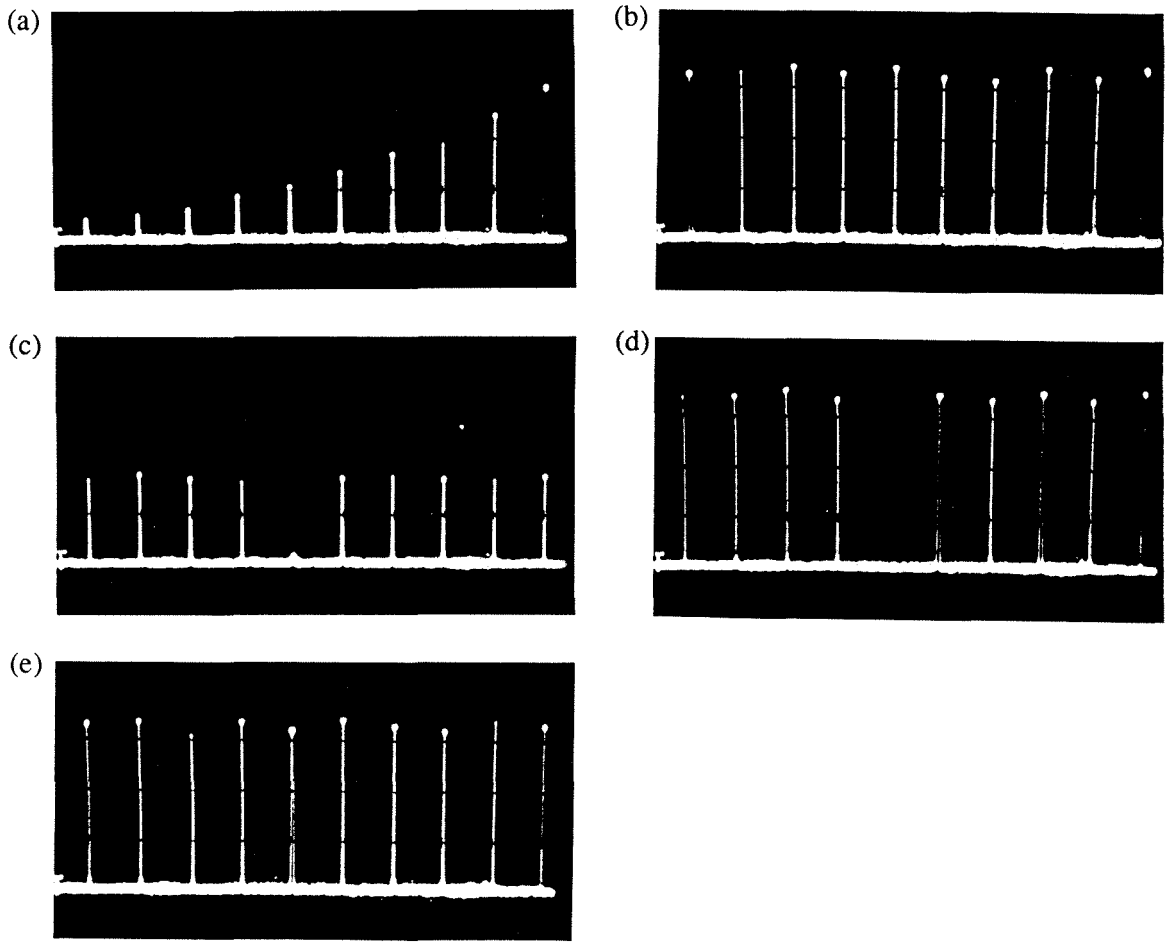


Figure 3.12 Experimental results: (a) the diffracted signals of ten angularly multiplexed holograms recorded with equal exposure; (b) the equalized holograms after dynamic copying; (c) selective erasure of the fifth hologram; (d) the other nine holograms rejuvenated by dynamic copying after being partially erased during selective erasure; (e) rewriting of the fifth hologram.

4. ELECTRICAL FIXING OF PHOTOREFRACTIVE HOLOGRAMS

4.1 Introduction

The dynamic copying methods described in the previous chapter can be used not only for refreshing multiply exposed holograms, but also for rejuvenating holograms that are partially erased by the readout light. The latter case can be thought of as a dynamic fixing scheme. Nondestructive readout can be also achieved by permanent hologram fixing and several fixing methods have been reported. Thermal fixing of holograms was demonstrated in $LiNbO_3$ [87], $Bi_{12}SiO_{20}$ [88], $KNbO_3$ [89], and $BaTiO_3$ [90], where a compensating ionic charge grating (which cannot be erased optically) is formed at an elevated crystal temperature. Micheron *et al.* demonstrated hologram fixing in $Sr_{0.75}Ba_{0.25}Nb_2O_6$ (SBN:75) [91] and $BaTiO_3$ [92] through the creation of a ferroelectric domain pattern by applying an external field at room temperature. Hologram fixing in $Sr_{0.75}Ba_{0.25}Nb_2O_6$ was also achieved by cooling the exposed crystal through the ferroelectric phase transition [93]. Leyva *et al.* demonstrated hologram fixing in $KTa_{1-x}Nb_xO_3$ by cooling the exposed crystal under an applied field through the ferroelectric phase transition [94]. In general, electrical fixing is preferable from a practical point of view because of its relative simplicity.

This chapter presents the results of our investigation on room-temperature electrical fixing of photorefractive holograms recorded in SBN:75 crystals [95]. We were able to reproduce some of the effects Micheron and Bismuth reported in Ref. 91, but our observations were different in several important respects. In addition, two novel ways of electrically fixing holograms in SBN:75 are presented that ex-

hibited improved performance, and it has been demonstrated that holograms of images can be fixed and faithfully reconstructed.

4.2 Experiments

The crystal sample used in the experiment was grown and poled at Rockwell International Science Center. It has a dimension of $6 \times 6 \times 6$ mm^3, with its c axis parallel to the edges. An external electric field can be applied along the c axis, and it is called positive (negative) if its direction is the same as (opposite to) that of the initial poling field. In our experimental setup (Fig. 4.1), an ordinary-polarized plane wave from an argon laser ($\lambda = 488$ nm) is split into three beams, two of which are used for recording a grating in the crystal, and the third is used as a non-Bragg-matched erasing beam. The grating vectors are approximately parallel to the c axis and the total recording intensity is about 10 mW/cm^2. The diffraction efficiency η is monitored with a low-intensity, extraordinary polarized HeNe laser beam incident at the Bragg angle. The diffraction efficiency is calculated by subtracting the background noise level from the measured diffracted light and dividing the difference by the transmitted light power.

In the first experiment, a holographic grating with a grating spacing $\Lambda = 11.6$ μm was recorded in the completely poled crystal without any applied field. After the diffraction efficiency η reached its saturation value ($\eta = 11\%$), the recording beams were blocked and a negative voltage pulse with amplitude $V =$ - 1 kV and duration $t = 0.5$ s was applied to the crystal, which caused η to fall quickly. After the voltage pulse was removed, η recovered a portion of its initial value before the pulse. Then the crystal was illuminated with the non-Bragg-matched erasing

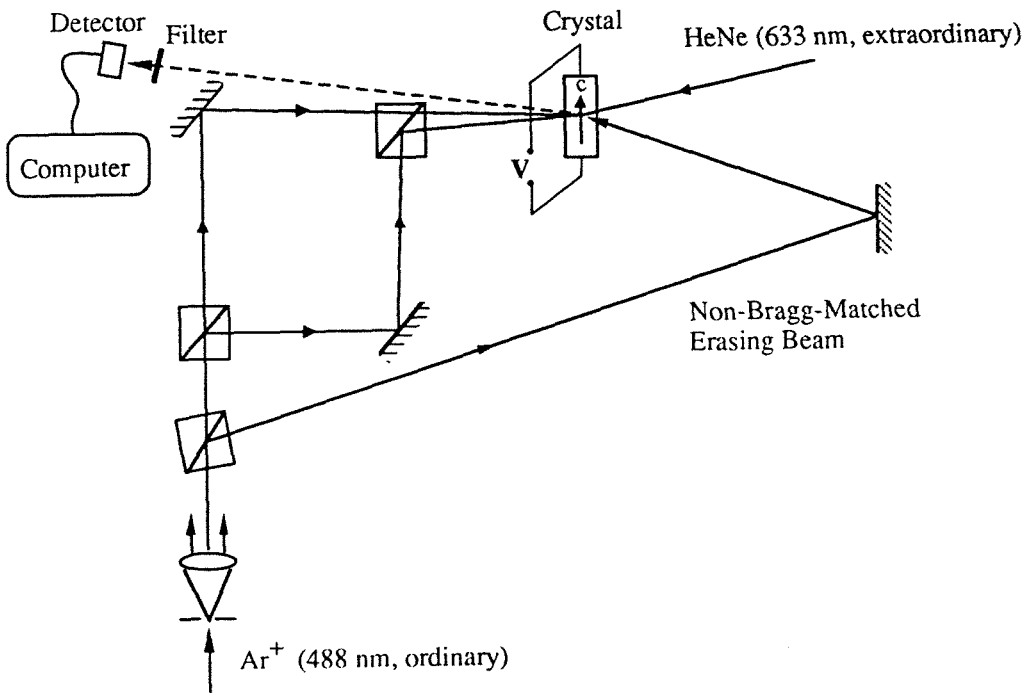


Figure 4.1. Optical Set-up.

beam, and η decreased further until it reached a steady-state value of $\eta \approx 0.06\%$. This fixed grating could not be erased by the erasing beam. Then the erasing beam was blocked and a positive voltage pulse, with amplitude $V = +\ 2$ kV and duration of a few seconds, was applied to the crystal. During the positive voltage pulse, $\eta = 1.8\%$ and after the pulse, a grating with $\eta = 0.4\%$ was revealed. This revealed grating can be optically erased. The experimental results of recording, fixing, revealing, and erasure are shown in Fig. 4.2, where the diffracted signal is plotted as a function of time during the different stages of the experiment.

Comparing the results in Fig. 4.2 with Ref. 91, we see that we were able to achieve electrical fixing, but the fixed grating obtained in our experiment is much weaker than in Ref. 91. Another important difference is our observation of a revealed grating upon applying a strong positive voltage across the crystal at relatively large grating spacings.

We have found that in general the strengths of the fixed grating and revealed grating can be greatly enhanced by applying a constant negative voltage during recording. With an applied voltage of $V =$ - 500 V, we recorded a grating with $\Lambda = 11.6\ \mu$m using an exposure time of 50 s. Then the recording beams were blocked, the applied voltage was removed and the grating was allowed to settle down to a steady level. Upon illuminating the crystal with the erasing beam, η decreased first and then rose back to a steady state, indicating a fixed grating of $\eta = 0.8\%$. Finally, a positive voltage pulse same as the one used in the previous experiment was applied, revealing a grating of $\eta = 0.5\%$. This experimental result is shown in Fig. 4.3.

The same experiment described in the last paragraph was repeated for several

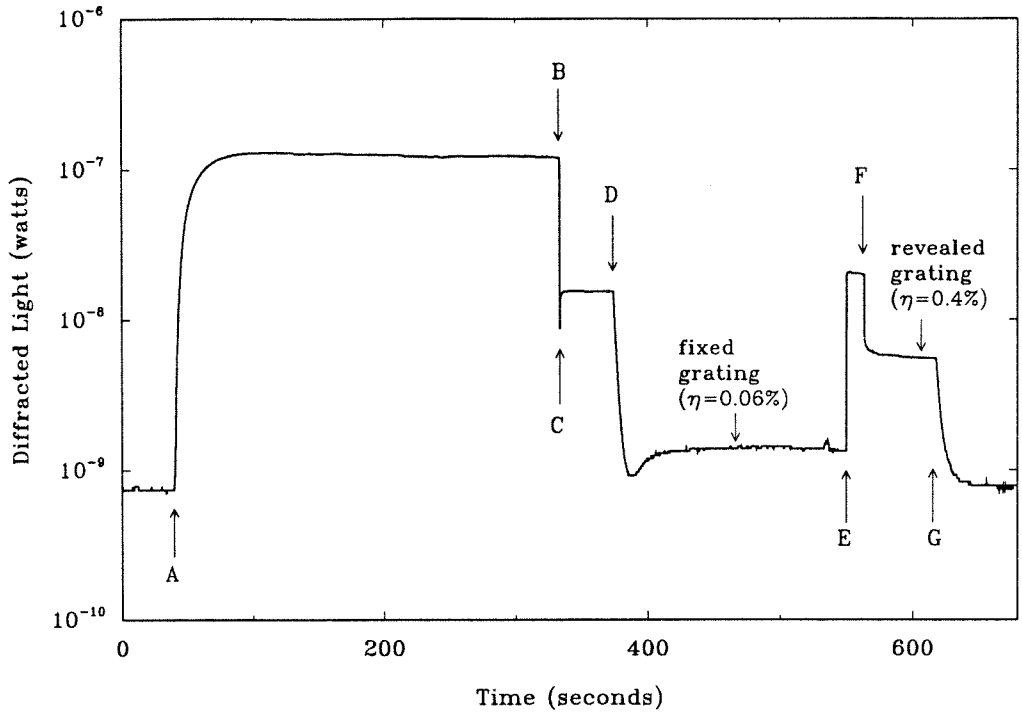


Figure 4.2. Diffracted light as a function of time for fixing with a negative voltage pulse ($\Lambda = 11.6\ \mu$m). The transmitted light power in the absence of the grating is 1.1 μW. (A) Hologram recording begins. (B) Negative pulse is applied. (C) End of negative pulse. (D) Optical erasure begins. (E) Erasing beam is blocked and positive voltage pulse is applied. (F) End of positive pulse. (G) Optical erasure begins.

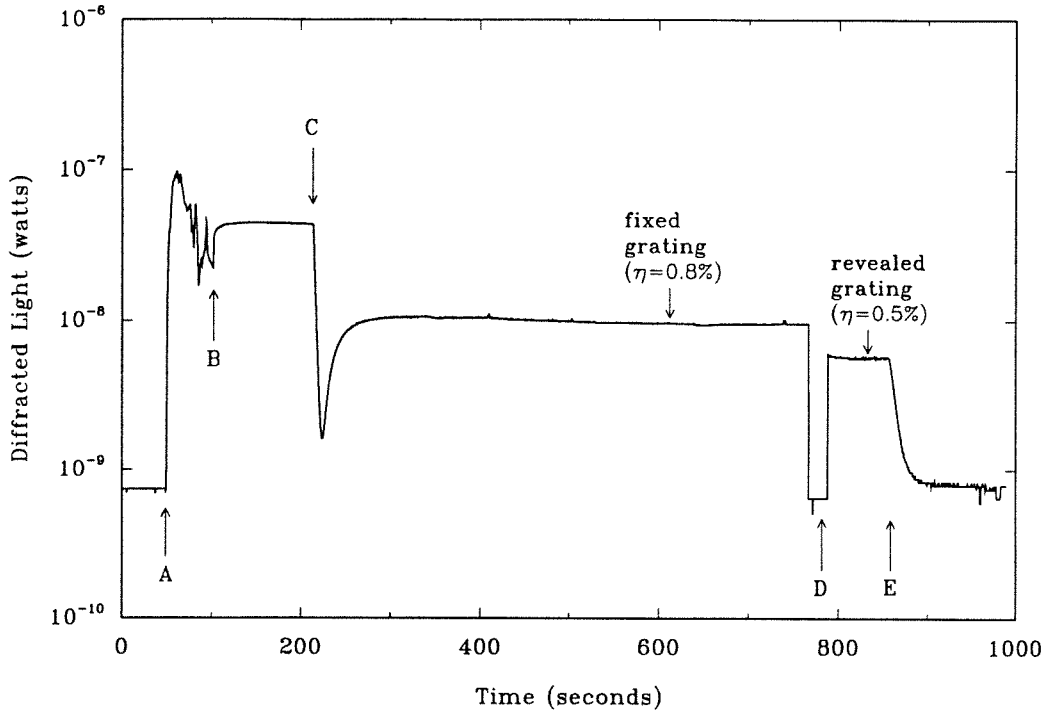


Figure 4.3. Diffracted light as a function of time for fixing with a constant negative voltage during recording ($\Lambda = 11.6$ μm). The transmitted light power in the absence of the grating is 1.1 μW. (A) Hologram recording begins and negative voltage is applied. (B) Negative voltage is removed and recording beams are blocked. (C) Optical erasure begins. (D) Positive voltage pulse is applied (Probe beam is blocked) (E) Optical erasure begins.

different grating spacings, and the results are summarized in Fig. 4.4. Under our experimental conditions, the maximum diffraction efficiency of the fixed grating is found to be 0.8% at $\Lambda = 11.6$ μm. The revealed grating is observed only for relatively large Λ and its strength increases drastically with the increase in Λ, achieving about 17% diffraction efficiency at $\Lambda = 29$ μm.

Finally, an image was recorded in the crystal as an image plane hologram, and Fig. 4.5a shows the reconstruction of the hologram. Figure 4.5b shows the reconstruction of the fixed hologram. The fixed hologram did not show any sign of degradation under illumination with the non-Bragg-matched erasing beam for two and a half hours.

4.3 Physical Model

The fixing and revealing process we described was not observed when a positive rather than negative fixing voltage was applied. If we assume that the mechanism responsible for fixing involves compensating ions, then we have no obvious explanation for the asymmetry regarding the polarity of the fixing voltage. Similarly, if the mechanism involves two types of photorefractive species [96], the effect would not be sensitive to the polarity of the fixing voltage. The fact that the fixed grating can be erased nearly instantaneously (less than 0.1 s) in the dark (i.e., without redistribution of charge carriers among the trap sites) also indicates that the effect does not involve two types of trap sites that compensate each other. Therefore, we believe that the mechanism responsible for this effect is the polarization grating formation suggested in Ref. 91.

Specifically, our observations can be explained as follows. With the aid of

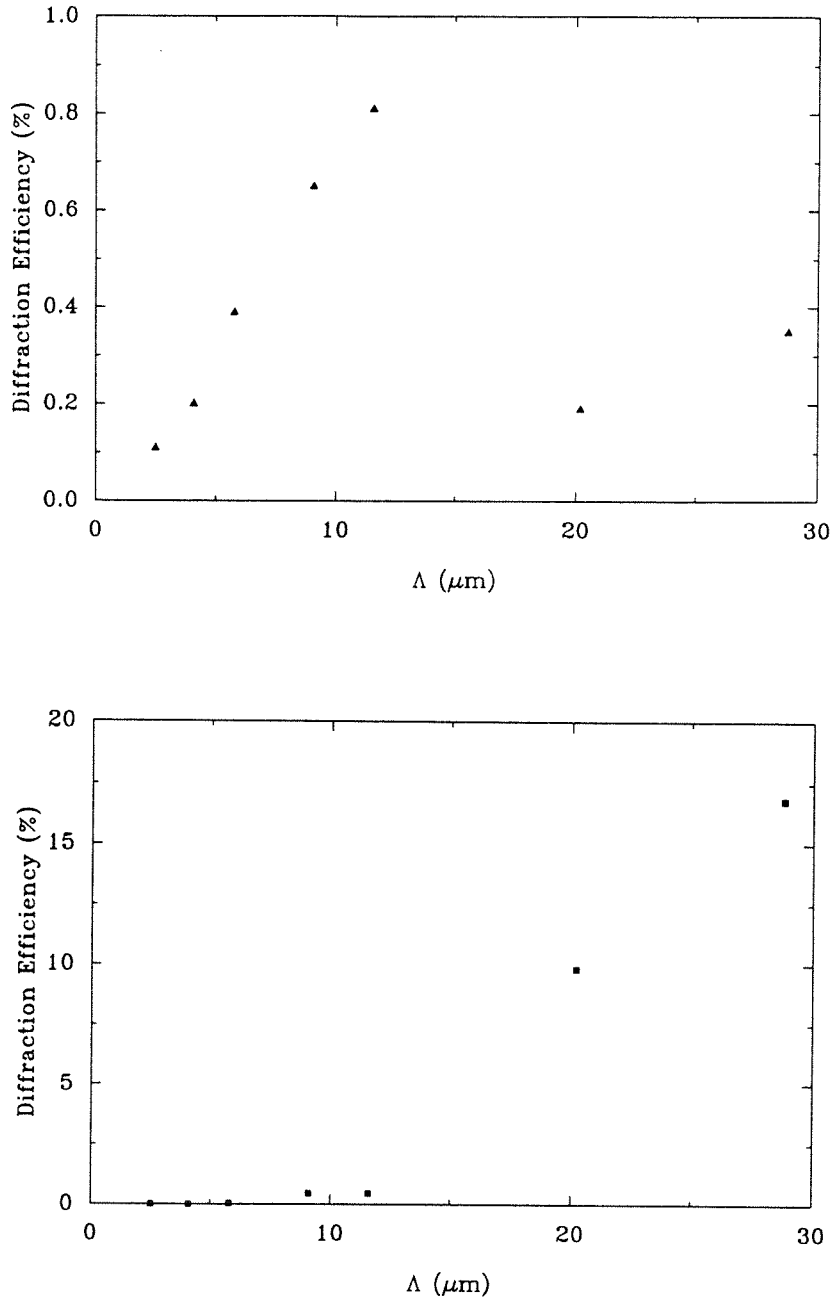


Figure 4.4 Diffraction efficiency as a function of grating spacing for (a) the fixed grating and (b) the revealed grating.

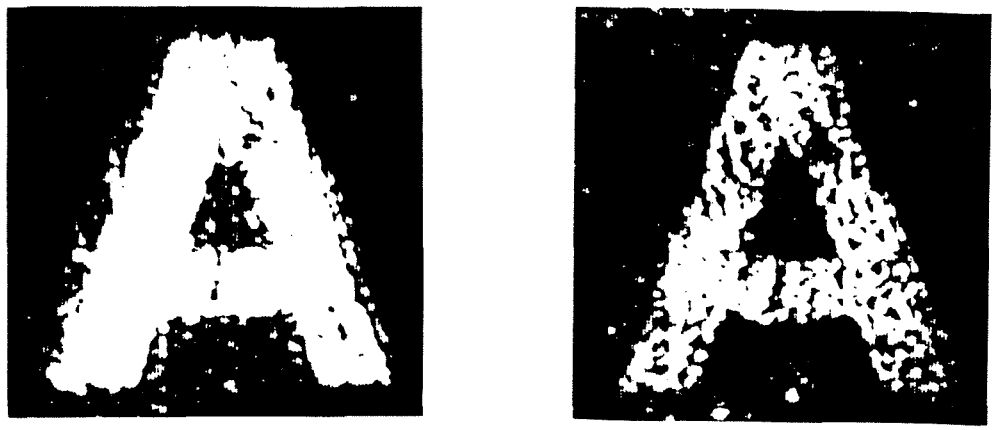


Figure 4.5 (a) Reconstruction of the recorded hologram. (b) Reconstruction of the fixed hologram.

the negative external field, the electronic space-charge field established during the holographic recording causes a spatial modulation of the ferroelectric polarization. In the areas where the space-charge field is negative, the local polarization is modified (which corresponds to local depoling or possible re-poling in the opposite direction). This causes the decrease of both the mean linear electrooptic coefficient and the amplitude of the space-charge field modulation. As a result, the amplitude of the refractive index modulation decreases upon applying a negative voltage pulse, causing the drop in the monitored diffraction efficiency. Under the illumination of the erasing beam, the electronic grating is erased further until a balance between the polarization grating and the electronic grating is reached.

The fixed grating observed is attributed to the polarization grating that is partially compensated by the electronic grating. Finally, a strong positive voltage applied to the crystal erases the polarization grating (i.e., the ferroelectric domains are re-aligned), revealing the compensating electronic grating which of course can be erased optically. The enhancement of the diffraction efficiency of the revealed grating during the positive voltage pulse (see Fig. 2) is attributed to the nonlinear electro-optic effect in SBN:75 [97].

A physical model describing the above process can be derived based on the band transport equations given in Chapter 2. Specifically, a spatial modulation of the ferroelectric polarization is equivalent to a fixed space charge distribution according to

$$\rho_p = -\vec{\nabla} \cdot \vec{P}. \tag{4.1}$$

Therefore the Poisson equation (Eq. (2.29)) becomes

$$\vec{\nabla} \cdot \epsilon \vec{E} = e(N_D^+ - N_A^- - n) - \vec{\nabla} \cdot \vec{P}. \tag{4.2}$$

Assuming small polarization mudulation and following the standard linearization procedure, the steady-state values for the net space-charge field and the compensating electronic component can be found. In the absence of externally applied field, the ratio of the diffraction efficiencies of the fixed and compensating (or "revealed") gratings is found to obey the following relationship [98]

$$\ln \frac{\eta_{reveal}}{\eta_{fix}} = 4ln\frac{\Lambda}{\Lambda_e} + c, \tag{4.3}$$

where c is a constant and $\Lambda_e(= 2\pi\sqrt{\frac{\epsilon k_B T}{e^2 N_A^-}})$ is the Debye length. It should be pointed out that Eq. (4.3) is independent of the actual value of the polarization modulation, which allows us to compare the theory and experiment without accurate measurement of the polarization modulation. Figure 4.6 shows the experimental data of $ln(\eta_{reveal}/\eta_{fix})$ versus $\ln\Lambda$, and a linear fit indicates slope of 4.2 ± 0.2 which closely matches the theoretical value of 4.

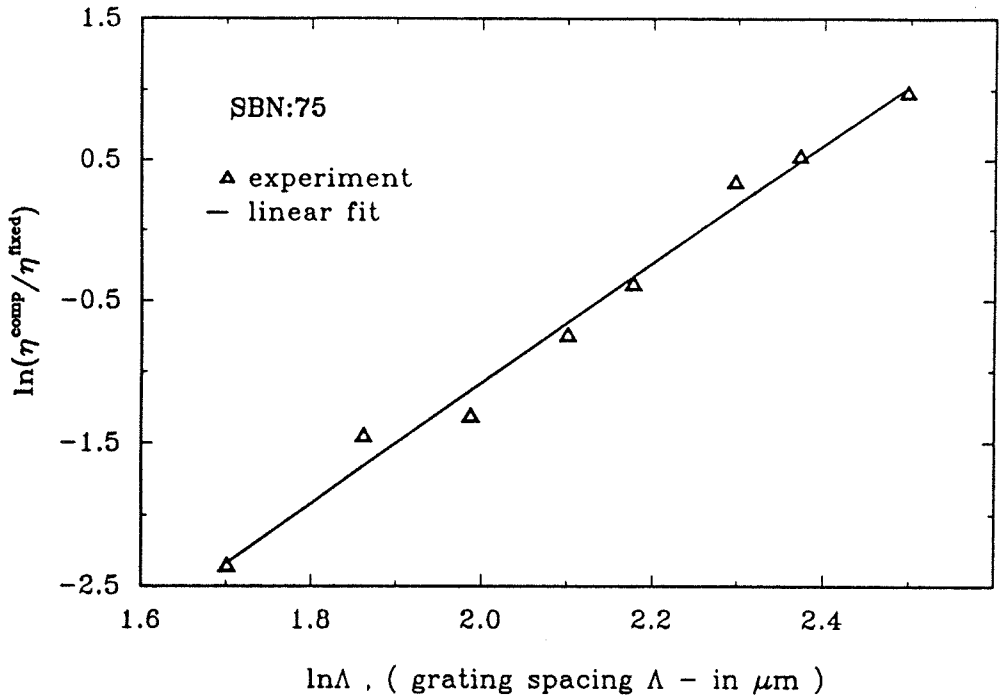


Figure 4.6 Log-log plot of the ratio of the diffraction efficiencies of the revealed compensating grating and the fixed grating vs. the grating spacing Λ. The linear fit has a slope of 4.2 ± 0.2, which is in good agreement with the theoretical value of 4.

The above experiments suggest two modes of hologram fixing in SBN:75 crystals: the fixed polarization grating and the revealed electronic grating. Although the revealing process is destructive to the polarization grating, it is possible to recreate this grating by applying a negative voltage pulse after the electronic grating is revealed. These revealing/fixing cycles were repeatedly performed (shown in Fig. 4.7) and the diffraction efficiency of the revealed grating (and thus the polarization grating) remained unchanged.

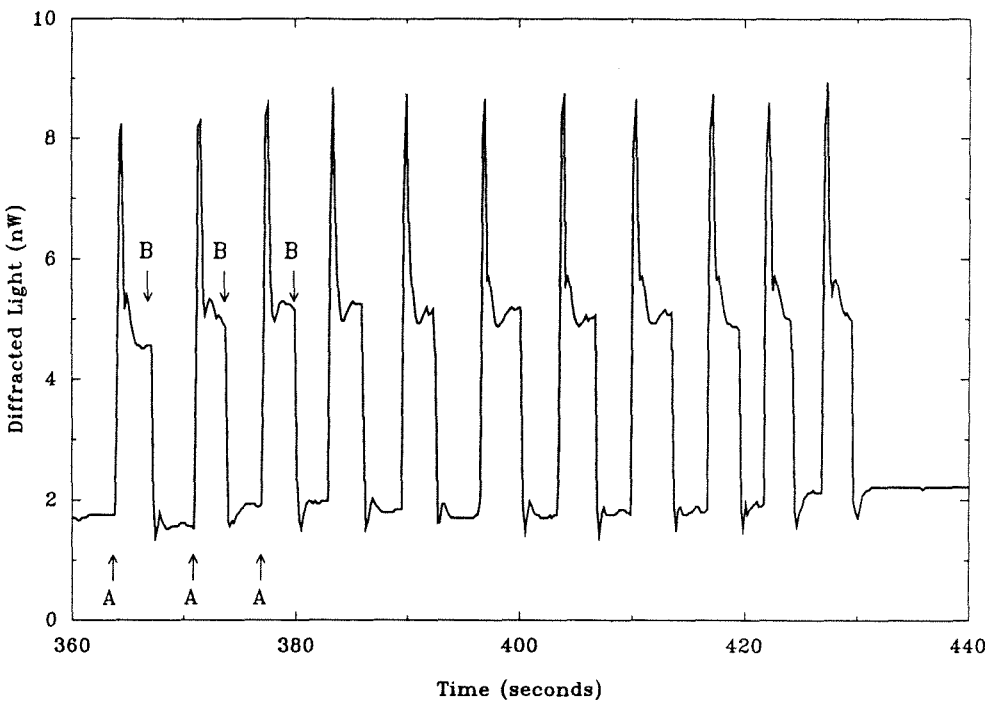


Figure 4.7 Cyclic fixing and revealing of the stored hologram ($\Lambda = 20.2\ \mu$m), with the erasing beam off during the entire process. (A) Positive (revealing) pulse is applied, with $V = +\ 1$ kV. (B) Negative (fixing) pulse is applied, with $V =$ - 1 kV.

5. LEARNING ALGORITHMS FOR OPTICAL MULTILAYER NEURAL NETWORKS

5.1 Introduction

One of the most important issues in the optical implementation of neural network models is the determination of a proper algorithm to train the network. The criterion is that they must effectively utilize the massive parallel processing capability provided by optics, while the operations involved in learning should be easily implementable with existing optical and optoelectronic devices and components.

The back error propagation (BEP) algorithm [28] and its variants are the most popular procedures for training multilayer optical networks [56,61,62]. Back error propagation is an example of a learning algorithm that yields *distributed* representations in the hidden layers of a network. In a distributed representation, a large portion (typically half) of the hidden units respond when the input is one of the training samples. In contrast, in a *local*-distribution learning algorithm, each hidden unit is trained to respond to only a small number of training examples. The Radial basis function (RBF) classifier is an example of a commonly used local-representation algorithm. An optical RBF system has been recently demonstrated [99].

The advantage of local-representation algorithms is the fact that the training process is relatively easy. If an input training sample does not cause any of the existing hidden units to respond sufficiently, a new hidden unit is added and devoted to the new sample. The disadvantage of local-representation algorithms

is the large network size that is typically obtained. On the other hand, distributed-representation learning algorithms have the advantage of a smaller network size. However, the training of such networks is difficult, typically requiring a large number of training cycles. The optical implementation of algorithms such as BEP requires a dynamic holographic medium such as a photorefractive crystal, the behavior of which can be accurately controlled. When a new hologram is recorded in a photorefractive crystal the previously recorded signal is partially erased. This "weight decay" in effect limits the number of cycles a training algorithm can run on an optical system, since earlier exposures are erased as the training progresses. Dynamic copying [79–83] can overcome this problem by restoring the strength of the hologram through feedback. Another way for bypassing the weight decay problem is to use local-representation algorithms since they do not require long training sequences. In this case the large storage capacity of 3-D holograms can be used to synthesize the large networks that are required.

The distributed-representation learning algorithms are discussed in section 4.2. We start with the single Perceptron network as a basic building block of feedforward multilayer networks. Then the BEP algorithm and its optical implementation is reviewed. An Anti-Hebbian Local Learning algorithm is described which eliminates the error propagation required by BEP, and therefore greatly simplifies the optical system. The RBF network is described in section 4.3 as an example of the local-representation networks.

Section 4.4 presents another type of multilayer networks, the structures of which are not fixed in advance. Rather the structures of these networks are determined during the learning based on the difficulty level of the given problems. A

typical example is the tiling algorithm, which will be discussed in some detail.

5.2 Distributed-Representation Networks

5.2.1 Perceptron

Shown in Fig. 5.1 is a diagram depicting the most basic one-layer network with N input elements and one output. The weighted sum of the input elements is thresholded to yield the output

$$y = g(\sum_{i=1}^{N} w_i x_i), \qquad g(z) = \begin{cases} 1, & z > 0; \\ 0, & z \leq 0, \end{cases} \tag{5.1}$$

where $g(z)$ is the nonlinear thresholding function, w_i is the weight of the ith connection, and x_i is the ith element of the input pattern. Such a network can be used to dichotomize a set of patterns into two prescribed classes, and more complex, multilayer networks can be built up using this as the basic building block.

The perceptron [24] is a powerful learning algorithm for the single-layer network of Fig. 5.1 in that if there exists a set of weights that could solve the dichotomization problem, the algorithm guarantees it will find this solution set. The perceptron algorithm is given by the update equation:

$$w_i(n+1) = w_i(n) + \alpha(n)x_i(n), \tag{5.2}$$

where $w_i(n)$ is the ith weight at the nth iteration, and $x_i(n)$ is the ith input at the nth iteration. $\alpha(n)$ is given by

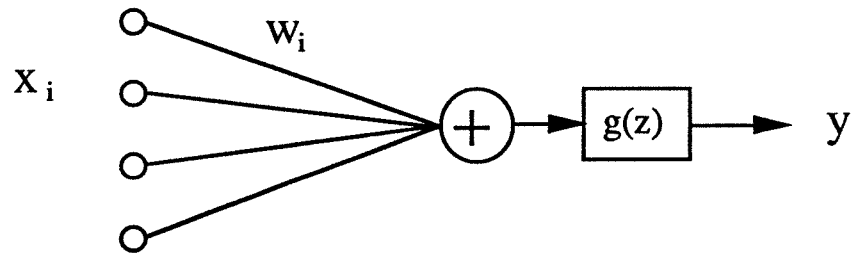


Figure 5.1. Single-layer neural network.

$$\alpha(n) = \begin{cases} 0, & \text{if } y(n) \text{ is correct;} \\ 1, & \text{if } y(n) = 0 \text{ but should have been 1;} \\ -1. & \text{if } y(n) = 1 \text{ but should have been 0.} \end{cases} \tag{5.3}$$

The threshold bias can be absorbed into the input patterns by choosing one element of each pattern to be always equal to a nonzero constant [24]. Note that both additive and subtractive changes to the weights w_i must be made to implement the algorithm directly.

The optical perceptron can be implemented with the correlator setup shown in Fig. 1.4. By virtue of their dynamic nature, photorefractive crystals are ideal candidates for the holographic medium. If the hologram is exposed with a weight pattern **w** at the input plane and a training plane wave, then the inner product between a new input **x** and the weights **w** stored in the hologram can be obtained at the center of the correlation pattern at the output plane. The holographic

process records both the magnitude and phase of patterns and thus it can be used to store bipolar valued weights and adaptable changes, which are either subtractive or additive, are possible.

Specifically, it requires that a grating can be recorded in the photorefractive crystal with a phase of either 0 or π (relative phase difference is π. After writing a hologram with one phase setting, hologram subtraction can be performed by using the other phase setting in the subsequent exposure and addition can be performed by using the original phase setting. A method to acheve this is shown in Fig. 5.2 in which the training beam passes through a liquid crystal phase modulator to acquire a phase of either 0 or π relative to the input signal beam.

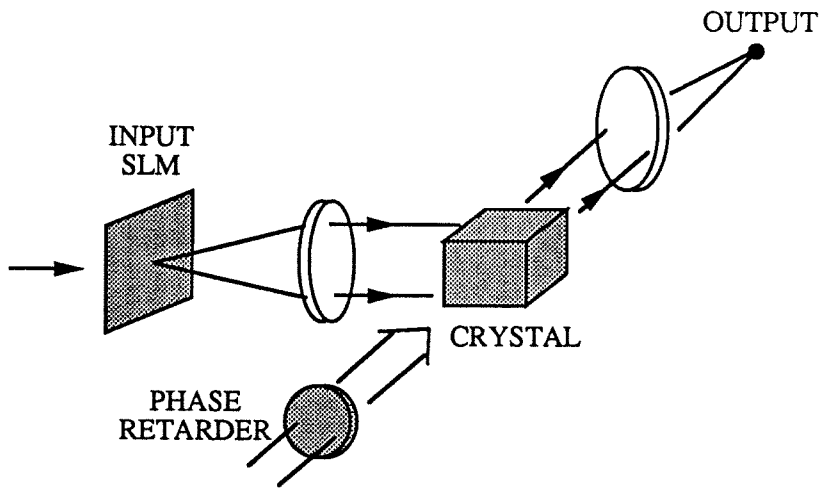


Figure 5.2. Optical implementation of the perceptron.

5.2.2 Back Error Propagation Algorithm

The single-layer system of Fig. 5.1 can be cascaded to form a multilayer netork, and one of the most widely used learning algorithms for training fully adaptive multi-layer neural networks is the Backward Error Propagation (BEP) algorithm [27,28], which is a steepest descent algorithm that minimizes the error at the output of the network with respect to the weight values of the connections between the neurons. Optical architectures capable of implementing the BEP algorithm have been proposed [53,61,68]. One of these architectures ([61]) is shown in Fig. 5.3. It has two layers but an arbitrary number of layers can be implemented as a straightforward extension. A training pattern $\mathbf{o}^{(0)}$ is placed at the input plane N_0. The pattern is then interconnected to the intermediate (hidden) layer N_1 via the dynamic volume hologram H_1. Simulating the action of an array of hidden neurons, a spatial light modulator (SLM) placed at plane N_1 performs a soft thresholding operation on the light incident on it to produce $\mathbf{o}^{(1)}$, the output of the hidden layer of neurons. Hologram H_2 connects N_1 to the output plane N_2 where another SLM performs the final thresholding to produce the response of the network to the particular input pattern. This network output $\mathbf{o}^{(2)}$ is compared to the desired response $\mathbf{t}$ and an error signal $\boldsymbol{\delta} = \mathbf{t} - \mathbf{o}^{(2)}$ is generated at N_2. The undiffracted beams from N_0 and N_1 are recorded on SLMs at T_1 and T_2 respectively. The SLMs at T_1, T_2 and N_2 are then illuminated from the right to read out the stored signals and the modulated light propagates back toward the left. Let $s_j^{(n)}$ be the total input to the jth neuron in plane N_n, and $w_{ji}^{(n)}$ be the weight of the interconnection between the jth neuron at N_n and the ith neuron at N_{n-1}, for $n = 1, 2$. Let the function $f[\cdot]$ be the thresholding function which operates on the input to each neuron in the forward path. According to the BEP

algorithm, the change of the interconnection matrix stored in H_2 is given by

$$\Delta w_{kj}^{(2)} \propto \delta_k f'[s_k^{(2)}] o_j^{(1)}, \tag{5.4}$$

where $f'[\cdot]$ is the derivative of $f[\cdot]$. Each neuron in N_2 is illuminated from the right by the error signal δ_k and the backward transmittance of each neuron is proportional to the derivative of the forward output. The hologram recorded in H_2 is the outer product of the activity patterns on planes N_2 and T_2, which means the change made in H_2 is that described by Eq. (5.4).

The change in the interconnection matrix stored in H_1 is

$$\Delta w_{ji}^{(1)} \propto \left[\sum_k \delta_k f'[s_k^{(2)}] w_{kj}^{(2)}\right] f'[s_j^{(1)}] o_i^{(0)}. \tag{5.5}$$

The error signal applied to N_2 produces a diffracted signal at the jth neuron at N_1 which is proportional to $\sum_k \delta_k f'[s_k^{(2)}] w_{kj}^{(2)}$. By setting the backward transmittance of the jth hidden neuron to be proportional to $f'[s_j^{(1)}]$, the interconnection matrix in H_1 is modified as described by Eq. (5.5).

The hardware implementation of the BEP algorithm is complicated by the need to realize error backpropagation through the network and the need for bidirectional optical devices with different forward and backward characteristics. To overcome this problem, we describe an Anti-Hebbian Local Learning (ALL) algorithm for two-layer networks [56]. With this rule, weight update for a certain layer depends only on the input and output of that layer and a global, scalar error signal. We show that this learning procedure still guarantees that the network is trained by error descent. The fact that error signals need not back-propagate through the network makes this local learning rule easy to implement.

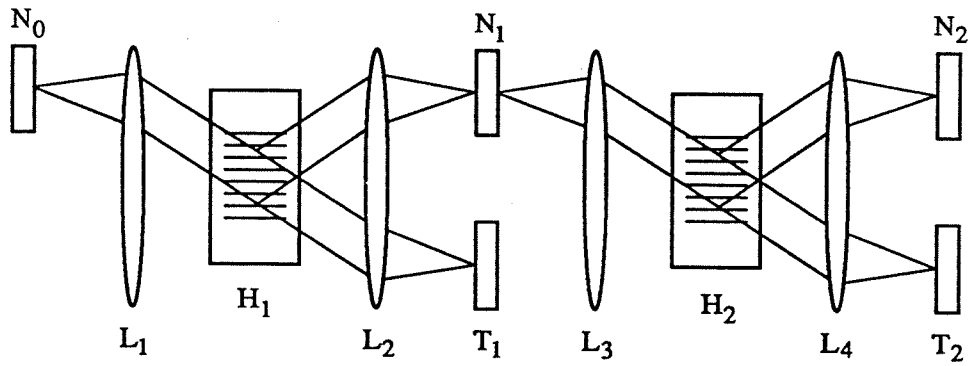


Figure 5.3. Optical architecture for backward error propagation (BEP) learning.

5.2.3 Anti-Hebbian Local Learning Algorithm

As pointed in the previous section, the practical implementation of the BEP algorithm is complicated by the fact that BEP is a non-local learning rule. Consider the two layer network shown in Fig. 5.4. If we change any one of the weights in the first layer (decrease or increase), the effect of this change on the output of the network depends on the value and sign of the weights of the second layer. BEP

is a steepest descent method and therefore it attempts to decrease the output error as quickly as possible. To accomplish this, we require knowledge of the second layer weights to change the weights in the first layer. From an implementation point of view, this complicates matters since we must communicate information in both directions and have bi-directional neurons with different functionality in the forward and backward directions [61]. In what follows, we describe an algorithm for training two-layer optical neural networks, in which the weight updates are calculated from the signal at the input of each connection, the signal at the other end of the same connection, and a global, scalar error signal. The advantage of this algorithm is that it can be implemented with signals that are locally available which simplifies the optical system.

We can get an intuitive feel for how our algorithm works by considering the network of Fig. 5.4 with only a single output neuron. Suppose further that the output of that neuron takes values +1 or -1. The weights of the second layer are trained using the same procedure as in BEP. For the weights of the first layer, if the output is wrong for a particular input, then we can correct it by adjusting the weights of the first layer to produce the negative of the current response of the hidden layer [100]. Therefore, we can treat the first layer as a single layer net with known desired output and thus it can be trained with one of the existing algorithms for training single layer nets. We will show in what follows that it is possible to select the training algorithm for the first layer to guarantee that the output error will decrease at each iteration.

Let the numbers of neurons for the input, first and second layers of the two-layer network shown in Fig. 5.4 be N_0, N_1 and N_2, respectively. The inputs to

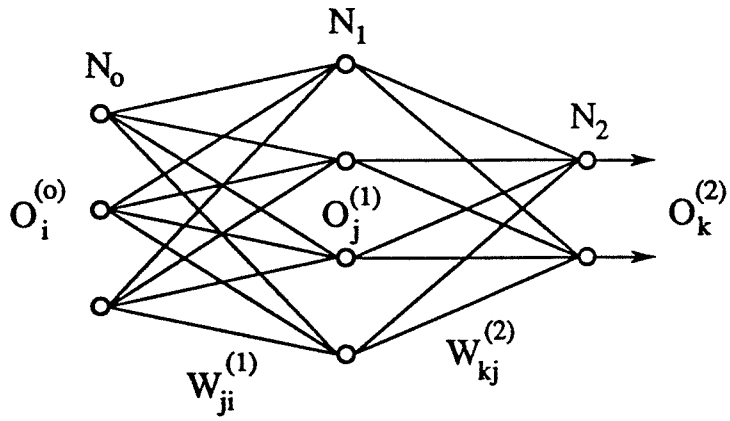


Figure 5.4. Schematic diagram of a feedforward two-layer neural network.

the neurons of the nth layer are:

$$s_j^{(n)} = \sum_{i=1}^{N_{n-1}} w_{ji}^{(n)} o_i^{(n-1)}, \tag{5.6}$$

where $w_{ji}^{(n)}$ is the weight of the interconnection between the jth neuron in the nth layer and the ith neuron in the previous layer, and $o_i^{(n)}$ is the output of the ith neuron in the nth layer. The signal at the input layer is denoted by $o_i^{(0)}$. The first and second layers of neurons perform a soft thresholding operation on their inputs to produce the outputs:

$$o_j^{(n)} = f[s_j^{(n)}], \tag{5.7}$$

where the function f is chosen as $f[x] = tanh[x]$ for the analysis.

The desired response corresponding to the input $o_i^{(0)}$ is the vector $\mathbf{t}$, whose elements are binary, $t_k \in \{1, -1\}$ for $k = 1, 2, \cdots, N_2$. The output error of the network is measured by the logarithmic energy function [101,102]:

$$E = \sum_{k=1}^{N_2} \{(1 + t_k) ln \frac{1 + t_k}{1 + o_k^{(2)}} + (1 - t_k) ln \frac{1 - t_k}{1 - o_k^{(2)}}\}. \tag{5.8}$$

The above function has its global minimum at $E = 0$ and it reaches this minimum if and only if the network output is the same as the desired response. We chose this form of error function instead of the more commonly used quadratic error function, because we found in simulations that for the problems we investigated this energy function gave better performance. The algorithm we describe in this paper also works with the quadratic error function with a slight, straightforward modification of the weight update rule we derive below.

The BEP rule changes the weights via gradient descent, i.e.,

$$\Delta w_{kj}^{(2)} \propto -\frac{\partial E}{\partial w_{kj}^{(2)}} = 2\delta_k o_j^{(1)}, \tag{5.9}$$

$$\Delta w_{ji}^{(1)} \propto -\frac{\partial E}{\partial w_{ji}^{(1)}} = 2(1 - o_j^{(1)^2}) o_i^{(0)} \sum_{k=1}^{N_2} \delta_k w_{kj}^{(2)}, \tag{5.10}$$

where $\delta_k = t_k - o_k^{(2)}$ is the output error signal. The non-local nature of the BEP algorithm is due to the $\sum_{k=1}^{N_2} \delta_k w_{kj}^{(2)}$ factor in Eq. (5.10), which contains the values of the weights of the second layer.

Let $\gamma_k = \delta_k s_k^{(2)}$. Then γ_k is positive if and only if the sign of the kth output unit matches the sign of t_k. For example, let us assume that t_k = -1 and $o_k^{(2)} > 0$. Then $\delta_k < 0$. The input signal to the kth output unit, $s_k^{(2)}$, has the same sign as the output of that unit and is therefore positive also. This gives us $\gamma_k < 0$ for this case. Let us now define the quantity $\gamma = \sum_{k=1}^{N_2} \gamma_k$ which will be positive if the sign of most of the output units matches the target sign and will be negative if the reverse is true. We can construct a learning rule for the first layer of weights using γ as a performance metric. Notice that γ is a scalar quantity that can be calculated from signals that are available at the output layer of the network. The basic idea is to modify the first layer of weights so that we reinforce the production of the current hidden layer response if γ is positive or reinforce the production of the negative of the current hidden layer response if γ is negative. We know that the simple Hebbian learning rule

$$\Delta w_{ji}^{(1)} \propto o_j^{(1)} o_i^{(0)}, \tag{5.11}$$

reinforces the reproduction of the current response when presented with the same input. Therefore, if we want to reinforce the negative of the current response when the output is incorrect, then we can simply adopt an "anti-Hebbian" rule by multiplying the right-hand side of Eq. (5.11) by γ. This idea leads to the following Anti-Hebbian Local Learning (ALL) algorithm for the first layer:

$$\Delta w_{ji}^{(1)} \propto \frac{\gamma o_j^{(1)} o_i^{(0)}}{(1 - o_j^{(1)^2})}. \tag{5.12}$$

The denominator in Eq. (5.12) is an additional term that is needed to guarantee that this learning rule always decreases the overall error at the output. Using Eqs.

(5.10) and (5.12), and assuming that the weights of interconnections between any input neuron and all hidden neurons are updated simultaneously (true in most practical situations), we obtain:

$$\begin{aligned}
\Delta E &= \sum_{j=1}^{N_1} \frac{\partial E}{\partial w_{ji}^{(1)}} \Delta w_{ji}^{(1)} \\
&\propto -\gamma o_i^{(0)^2} \sum_{k=1}^{N_2} \delta_k \sum_{j=1}^{N_1} w_{kj}^{(2)} o_j^{(1)} \\
&= -\gamma o_i^{(0)^2} \sum_{k=1}^{N_2} \delta_k s_k^{(2)} \\
&= -\gamma^2 o_i^{(0)^2} \\
&\leq 0, \qquad (5.13)
\end{aligned}$$

which proves our claim. Thus we proved that even though the ALL algorithm is not a steepest descent rule, it is still a descent rule.

In some cases, it may happen that during training, the output of certain hidden neuron becomes so close to +1 or -1 that the denominator $(1 - o_j^{(1)^2})$ in Eq. (5.12) is close to zero. This causes numerical instability. One way to avoid this is to find $o_{max}^{(1)}$, the hidden neuron output that has the maximum magnitude, and normalize the right-hand side of Eq. (7) by the factor $(1 - o_{max}^{(1)^2})$.

Computer simulations of the ALL algorithm were performed for the problem of recognition of handwritten zipcode digits provided by the U.S. Postal Service. For comparison, the BEP algorithm was also used to solve the same problem. The handwritten zipcodes were first segmented into single digits and then each digit

was reduced to fit a 10 × 10 binary pixel grid. A network of 100 input neurons (to match the 10 × 10 pixel grid), 5 hidden neurons and 3 output neurons was selected and trained to perform classification on 3 classes of handwritten digits: **3**, **6**, and **8**. Each output neuron responds to only one class. 600 digit patterns, with 200 patterns from each class, were selected. These 600 patterns were partitioned into 300 training samples, 150 validation samples, and 150 test samples. The validation samples were used after each learning iteration (i.e., presentation of the whole training set) to calculate the classification error of the network. The network training stops when the classification error of the network on the validation set stops decreasing with further iteration. After the network was trained, the test samples were presented to the network to find its generalization error. For the ALL algorithm, the first layer was trained only in 1 out of 40 iterations. By doing this, we rely more on the steepest descent training in the second layer and it improved learning convergence for this particular classification problem and this particular network. For given training, validation, and test sets, the network was trained 4 times with different initial conditions for both the ALL and the BEP algorithms. The same step size was used for the two algorithms for the purpose of comparison. The same simulations were repeated using different training, validation, and test sets obtained from different partitioning of the 600 digit patterns (the numbers of the training, validation, and test samples were still 300, 150, and 150, respectively). Therefore, there were a total of 8 runs for each algorithm.

For the ALL algorithm, the network was able to converge (meaning all the training patterns were classified correctly) in 7 cases. In only one case the network fell into a local minimum and gave a training error of 1%. For the BEP algorithm,

the network was able to converge in all the 8 cases. The average generalization errors for the ALL and the BEP algorithms are 9% and 8%, respectively. As for the average convergence rate of the two algorithms, it took 581 iterations for the BEP algorithm to converge, and 3,665 iterations for the ALL algorithm. However, since the amount of computation involved in each learning iteration is different for the two algorithms, we should also compare the convergence rate in terms of number of computational steps. It turns out that for this size of network, the number of computational steps in each iteration for BEP is about 2.4 times of that for ALL. In this sense, ALL is only about 2.6 times slower than BEP when implemented by a serial digital computer. In the optical implementation which is fully parallel, the appropriate speed comparison between BEP and ALL should be based on the number of cycles since the time required to complete each cycle is roughly the same for both algorithms. Therefore for the problem we studied in our experiment, a parallel optical implementation of ALL would be approximately 6 times slower than BEP. However, the optical system that implements ALL is much simpler than the BEP system, and this makes it much more likely that an ALL system can be built in practice. We should point out that the relative convergence rates of the two algorithms quoted above apply only to the problem we have tried and the relative performance will be generally problem dependent.

Shown in Fig. 5.5 is one possible implementation of the ALL algorithm. This architecture is quite similar to the architecture described in Refs. 53 and 61 for the implementation of BEP, with a few key differences. The input images are recorded on an electrically addressed spatial light modulator (EASLM1) and hologram #1 interconnects the pixels at the input plane to the pixels at the intermediate or

hidden layer plane. The non-linear response of the neurons at the hidden layer is simulated by an optically addressed spatial light modulator (OASLM). The second layer is very similar, with hologram #2 interconnecting pixels from the OASLM to the output plane where a 2-D CCD detector (CCD1) is placed to detect the light. At the input stage there is a second electrically addressed SLM (EASLM2) on which the reference signals necessary for the adaptation of hologram #1 are recorded. Similarly at the hidden layer there is EASLM3 to record the reference for hologram #2, which is simply the error signal δ_k (see Eq. (3.9)). This error signal is produced by subtracting the network output from the desired target signal, a point operation that can be accomplished either optically or electronically. The reference for hologram #1 is more difficult to derive (see Eq. (5.12)). It involves the global error signal γ and the response of the hidden layer $\mathbf{o}^{(1)}$. γ can be calculated with point operations from signals already available at the output of the system but we need a second detector (CCD2) at the output to record $\mathbf{o}^{(1)}$ as it is imaged through hologram #2. Once the reference signals are calculated and recorded on EASLM2 and EASLM3, hologram #1 (hologram #2) is exposed to the interference between the signal recorded on EASLM2 (EASLM3) and $\mathbf{o}^{(0)}$ ($\mathbf{o}^{(1)}$). This requires that a latching device, such as the microchannel spatial light modulator, is used as the OASLM. The key difference between the ALL architecture and the BEP architectures described in Refs. 53 and 61 is that the light always travels in the same direction throughout the system (left to right in Fig. 5.5). This simplifies the construction and the alignment of the system, and most importantly it does not require a device at the hidden layer (OASLM) that operates in both directions and has a different response function in each direction.

Thus, the ALL architecture is much more likely to be constructed in the foreseeable future.

5.3 Local-Representation Networks

A typical example of local-representation networks is the radial basis function (RBF) classifier [103]. The RBF classifier is somewhat similar to sample-based systems such as K-nearest neighbor (KNN) classifier [104] where learning time and learning algorithm complexity are traded for memory requirements and classification time. The moltivation for using a RBF classifier to perform pattern recognition tasks comes from the regularization and approximation theory [105]. The basic idea is to approximate a prescribed function $f(\mathbf{x})$ with a parameterized function $\hat{f}(\mathbf{w}, \mathbf{x})$, where $\mathbf{w}$ is a parameter vector used to tune the estimate $\hat{f}$. The problem becomes that given a set of training samples $\{\mathbf{x}_i \rightarrow f(\mathbf{x}_i); i = 1, ..., M\}$, how to choose the form of $\hat{f}$ and its parameters $\mathbf{w}$ such that $\hat{f}(\mathbf{w}, \mathbf{x}_i) = f(\mathbf{x}_i)$ for $i = 1, ..., M$. The RBF approach defines $\hat{f}(\mathbf{w}, \mathbf{x})$ as a weighted sum of radially symmetric basis functions:

$$\hat{f}(\mathbf{w}, \mathbf{x}) = \sum_{i=1}^{M} a_i \exp(-|\mathbf{x} - \mathbf{t}_i|^2 / \sigma^2), \tag{5.14}$$

where $\{\mathbf{t}_i\}$, called the centers, the widths $\{\sigma_i\}$, and the weights $\{a_i\}$ comprise the parameter vector $\mathbf{w}$, and are determined from the training set using any supervised or unsupervised learning algorithms.

The RBF classifier can be considered as a two-layer neuron network. We define the RBF unit, shown in Fig. 5.6(a), as a neuron with response given by

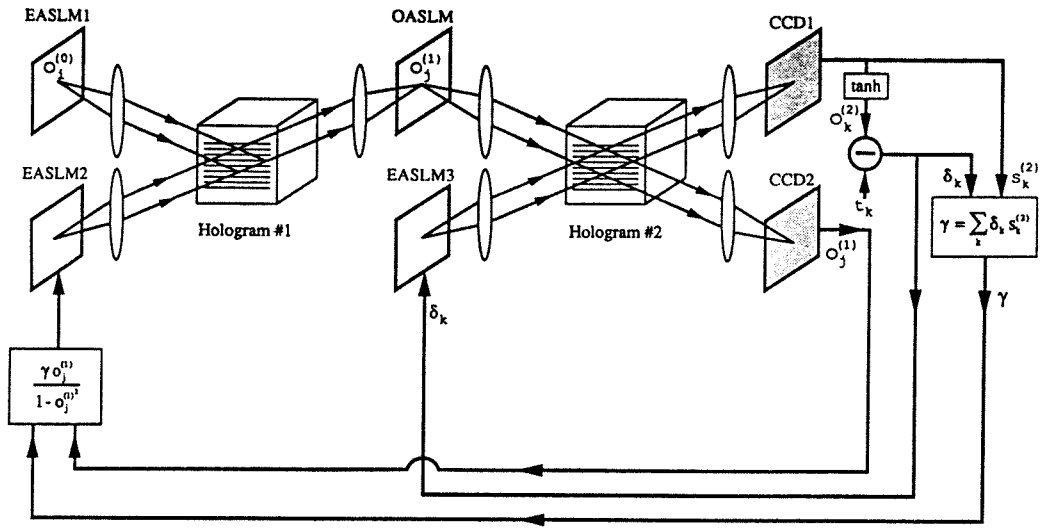


Figure 5.5. Optical architecture that implements the ALL algorithm.

$$y_i = \exp(-|\mathbf{x} - \mathbf{t}_i|^2/\sigma^2). \tag{5.15}$$

Then, the RBF classifier can be viewed as a two-layer shown in Fig. 5.6(b), where the hidden RBF units perform the nonlinear operation given by Eq. (5.15) to the weighted sum of the inputs, and the output layer consists of a single linear unit whose output is simply the weighted sum of its inputs.

The optical implementation of the RBF network has recently been demonstrated [99]. One important issue is to find an optical way of calculating the Euclidean distance $|\mathbf{x} - \mathbf{t}_i|^2$. As shown previously, an optical correlator is capable of computing the inner product of two vectors in a parallel fashion. If $\mathbf{x}$ and $\mathbf{t}_i$ are both binary (with their elements equal to either 0 or 1), then we can write

$$|\mathbf{x}|^2 = \mathbf{x} \cdot (\mathbf{t}_i + \bar{\mathbf{t}}_i), \tag{5.16}$$

so that

$$\begin{aligned} |\mathbf{x} - \mathbf{t}_i|^2 &= |\mathbf{x}|^2 + |\mathbf{t}_i|^2 - 2\mathbf{x} \cdot \mathbf{t}_i \\ &= |\mathbf{t}_i|^2 + \mathbf{x} \cdot \bar{\mathbf{t}}_i - \mathbf{x} \cdot \mathbf{t}_i. \end{aligned} \tag{5.17}$$

$\mathbf{x} \cdot \bar{\mathbf{t}}_i$ and $\mathbf{x} \cdot \mathbf{t}_i$ can be calculated with optical correlators, and $|\mathbf{t}_i|^2$'s can be stored in a memory and read out during the postprocessing stage to form $|\mathbf{x} - \mathbf{t}_i|^2$.

The advantage of the RBF network may be trained with some clustering algorithm rather than an error-driven procedure, and therefore the training time is reduced. One of the experimental systems described in the next chapter is

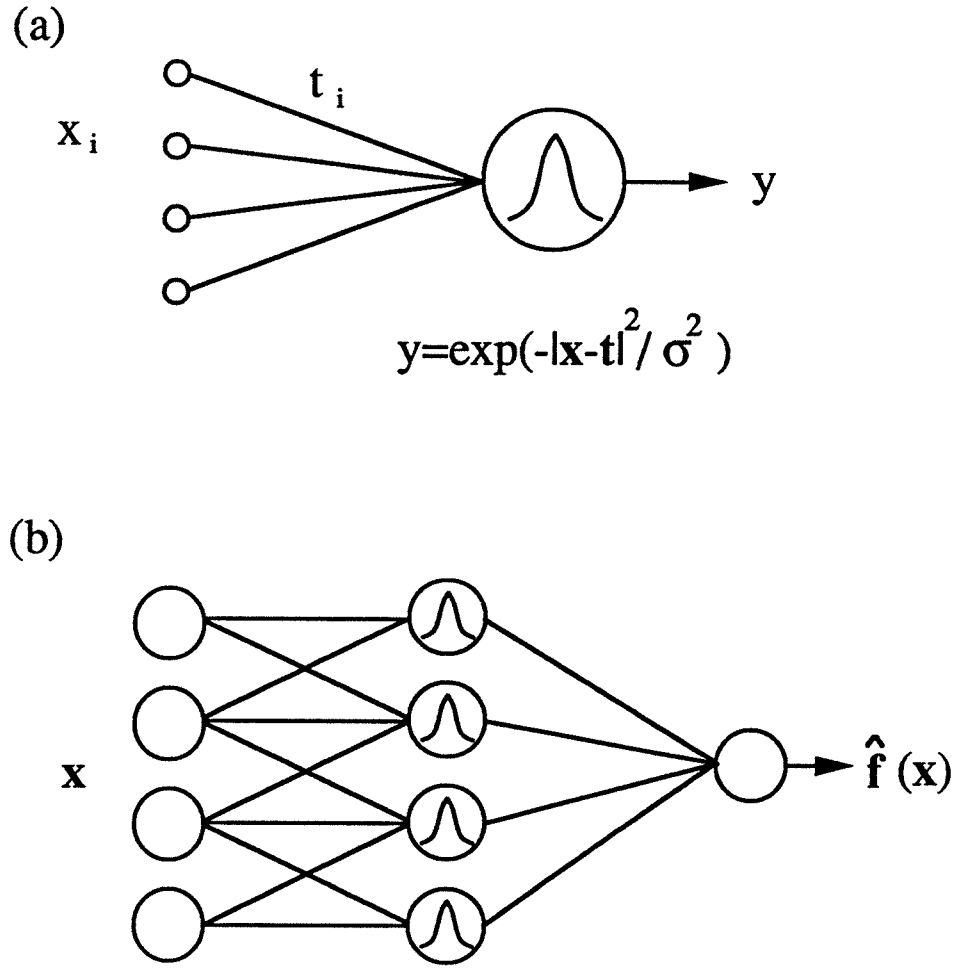


Figure 5.6. (a) Definition of a RBF unit; (b) Schematic of the RBF network.

trained with an algorithm partially moltivated by the RBF approach, so that the network can be trained relatively easily and the decay problem associated with multiply exposed holograms is less severe.

5.4 Multilayer Networks with Variable Structures

The neural network models discussed above have the same feature that their structures, including the number of layers and number of neurons for each layer, are fixed *a priori*. The choice of the appropriate network size for a given problem is usually based either on experience or on trial and error. There exist one type of learning algorithms [106–108] that does not require that the size of the network be fixed in advance. Rather, the size of a network is changed dynamically during learning, i.e., the network size is determined, as part of the learning, based on the difficulty level of the given problem. The network either grows or shrinks until it reaches a size which enables it to implement the desird mapping.

As an example, we will describe the tiling algorithm [106] in which case units are added like tiles whenever they are needed. The first unit of each layer is called a master unit. During the growth of the network, the master unit of the more recently built layer gives a strictly better approximation of the desired output than the previous one. Therefore, it guarantees convergence with a finite number of layers.

Let us consider layered nets, made of binary units whose output can be either +1 or -1. The ith unit in the Lth layer is connected to the N_{i-1} units of the preceding layer, and its output $o_i^{(L)}$ is obtained according to

$$o_i^{(L)} = \mathrm{sgn}(\sum_{j=0}^{N_{L-1}} w_{i,j}^{(L)} o_j^{(L-1)}), \tag{5.18}$$

where $w_{i,j}^{(L)}$ is the weight of interconnection between the jth unit in the $(L-1)$th layer and the ith unit in the Lth layer. The threshold bias is taken into account by a zeroth unit in each layer, clamped in the $+1$ state ($o_0^{(L)} = 1$), so that $w_{i,0}^{(L)}$ is the bias. For a given training set of p patterns of N_0 binary units, $\{\mathbf{x}^\mu = (x_j^\mu = \pm 1, j = 1, ..., N_0), \mu = 1, ..., p\}$, we want to learn a given mapping $\mathbf{x}^\mu \rightarrow y^\mu$ where y^μ is scalar, binary number.

To solve this problem, the tiling algorithm build a network layer by layer. In each layer, once the master unit has been obtained, one checks if it does the classification on the training set correctly. If yes, the training is complete. Otherwise, new auxiliary units are added to the layer until this layer gives a 'faithful representation' of the problem: any two input patterns with distinct outputs have distinct internal representations. Once this condition is fulfilled, the current layer is considered to have been achieved, and one proceeds to build the master unit of the next layer.

Each new unit is trained with a variant of the perceptron algorithm, the 'pocket algorithm' [109]. It consists of runing the standard perceptron algorithm, with a random presentation of the training patterns, but keeping in memory (in one's pocket) the set of weights which has produced the smallest number of errors so far. It has been shown that with probability as close to one as desired, it will give the set of weights with the least number of errors.

Let us denote the number of errors that the master unit of the Lth layer makes by e_L. It has been proved that if all the classes in layer $L-1$ are faithful,

and that e_{L-1} is non-zero, then there exists at least one set of weights connecting the $(L-1)$th layer to the master unit of the Lth layer such that $e_L \leq e_{L-1} - 1$. This ensures learning convergence with a finite number of layers.

The optical implementation of the tiling algorithm is straightforward since it involves only two tasks: (a) to implement the pocket algorithm as a variant of the perceptron algorithm; and (b) to build up the network unit by unit and layer by layer. The pocket algorithm can be implemented with the same optical setup shown in Fig. 5.2, except that a memory buffer should be added for temporary storage of the currently best set of weights, which can be reconstructed with the corresponding training beam and imaged onto the memory buffer. Adding new units to a layer is achieved by changing the angle of the training beam, so that new sets of weights can be multiplexed within the same hologram volume. Shown in Fig. 5.7 is one layer of an optical tiling network, and this module can be cascaded as new layers are added to the network as training proceeds.

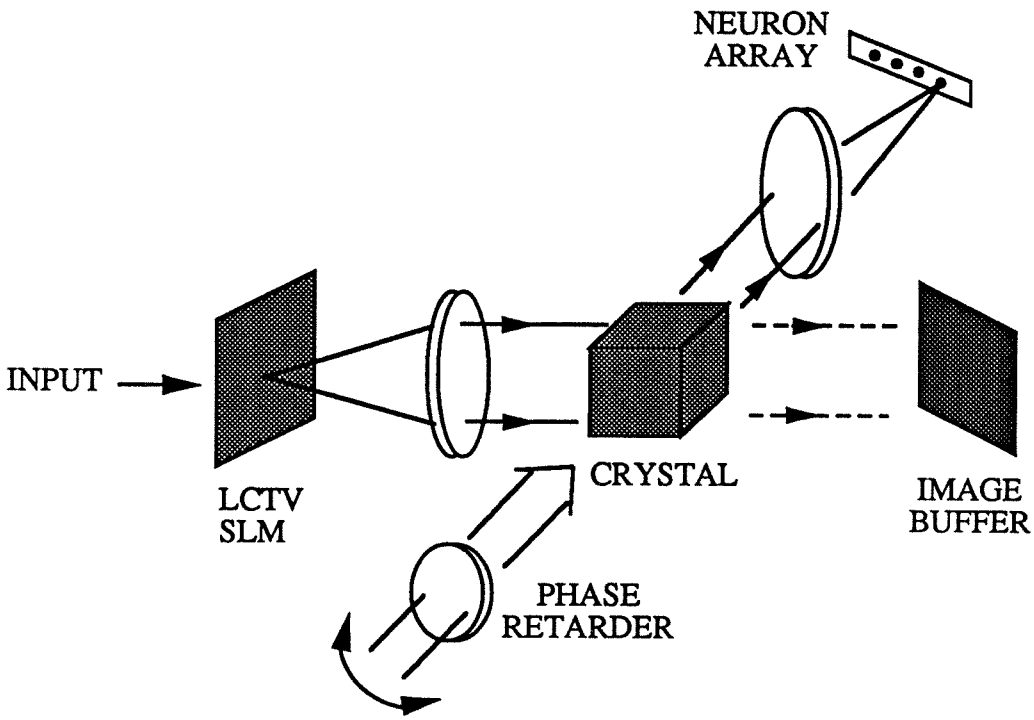


Figure 5.7. Basic module of the optical tiling network

6. EXPERIMENTAL DEMONSTRATIONS OF OPTICAL NEURAL NETWORKS

6.1 Optical Perceptron Using Dynamic Holographic Memory

In the optical perceptron network described in Chapter 5, learning is achieved by modifying the holographic weights with new exposures. As pointed out earlier, new exposures made to a photorefractive crystal will partially erase the previously recorded holograms. The weight update equation in comparison to the ideal rule of Eq. (5.2) is given by

$$w_i(n+1) = \exp(-t_e/\tau)w_i(n) + [1 - \exp(-t_e/\tau)]\alpha(n)x_i(n), \qquad (6.1)$$

where t_e is the exposure time. The partial erasure of the old weights may slow down the learning convergence, or even cause the network never to converge. This problem can be solved with the dynamic copying methods introduced in chapter 3. We have built an optical perceptron network based on the feedback copying scheme shown in Fig. 3.7, and the optical setup is illustrated in Fig. 6.1. The input patterns are displayed on a spatial light modulator (SLM), amplified by a liquid crystal light valve (LCLV), and fed to an optical correlator with a photorefractive hologram placed in the Fourier plane. The inner product between the input and the interconnection weights stored in the hologram is detected and then thresholded to produce the response of the output neuron. Real-time learning is achieved by modifying the hologram according to the perceptron algorithm, which says that the input is either added to or subtracted from the weights based on whether the actual output is less or greater than the desired output. The operations of addition

and subtraction are implemented in our system by appropriately setting the phase of the reference beam with an electrically controlled liquid crystal phase retarder.

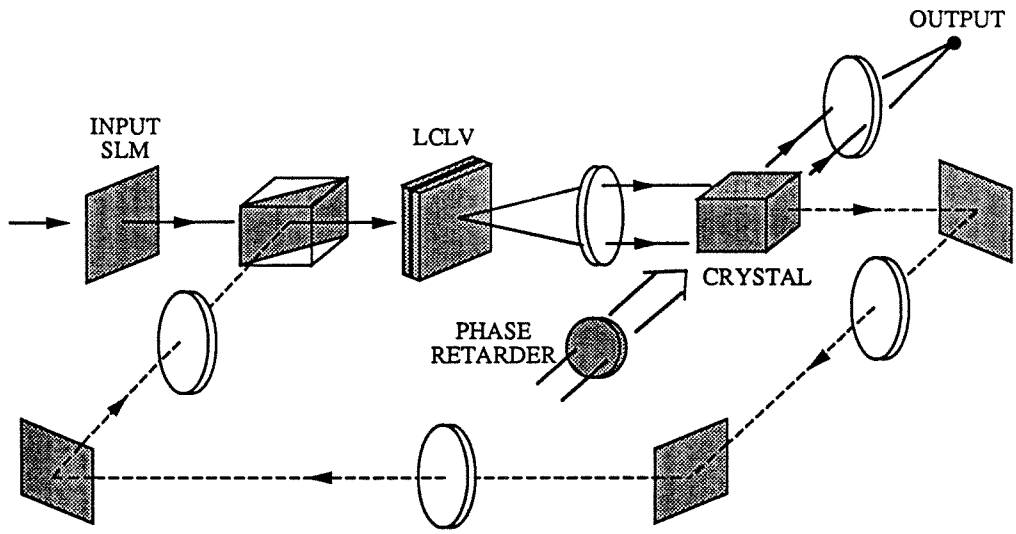


Figure 6.1. Dynamic holographic memory based optical perceptron network.

Dynamic copying is performed every few learning iterations by reconstructing the hologram with its reference beam and closing the feedback loop (denoted by the dashed lines in Fig. 6.1). The reconstructed signal is fed back to the writing side of the LCLV, and quickly amplified by the LCLV and the feedback loop. The original hologram is then refreshed by rewriting it with the amplified reconstructed signal.

The experimental perceptron network was trained to distinguish the faces of two people. The training set, selected from a video clip of the two people, consists of ten frames for each person. Figure 6.2 shows some of the training samples. The network was trained both with and without dynamic copying, and the experimental results are shown in Fig. 6.3. The optical network took 37 cycles to learn the task without dynamic copying, while it took only 20 cycles to converge with dynamic copying. It is expected that for larger training sets and larger networks, this difference in the convergence rate can be more significant. By cascading such optical learning modules, a multilayer neural network can be constructed that is capable of an arbitrarily long sequence of real-time adaptations.

6.2 Optical Two-Layer Network for Real-Time Handwritten Character Recognition

6.2.1 Introduction

This section describes an experiment in which commonly available optical devices are used to implement an adaptive multilayer network [62,110]. The system is a two-layer network that was trained based on Kanerva's model of Sparse, Distributed Memory (SDM) [111]. This learning model was chosen primarily because it is relatively easy to implement. The system uses photorefractive holograms as synaptic interconnections and liquid crystal light valves (LCLVs) to perform nonlinear thresholding. The first layer has random interconnection weights, which map each input pattern into a very large sparse, distributed internal representation. The second layer is trained by the sum-of-outer-products rule [112], which associates internal representations of different classes of characters to different re-

Figure 6.2. Examples of the training samples.

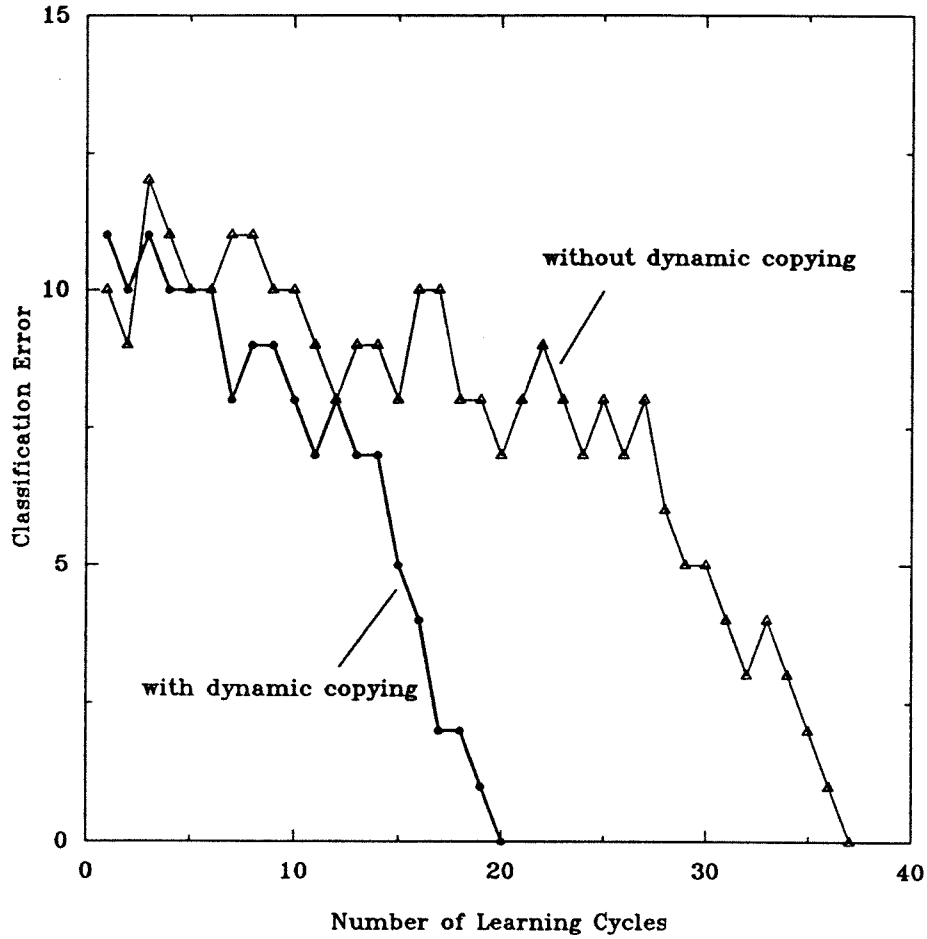


Figure 6.3. Experimental learning curves of the optical perceptron network.

sponses of output neurons. The trained network can recognize not only all the training patterns but also a fairly large percentage of test patterns that it has never seen.

6.2.2 System Architecture

We will briefly review the SDM model to point out the necessary characteristics that the optical system must incorporate. A schematic representation of a two-layer network is shown in Fig. 6.4, which consists of an input layer globally interconnected to a hidden layer, which is interconected through a second weighted network to an output layer. The system is trained so that the desired outputs $\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(M)}$ are produced for the respective input patterns $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(M)}$. Moreover, the output $\mathbf{o}^{(2)}$ of the network should be close to $\mathbf{y}^{(j)}$ when the system is presented with the input $\mathbf{o}^{(0)}$ close to $\mathbf{x}^{(j)}$. $\mathbf{y}^{(j)}$ and $\mathbf{x}^{(j)}$ are real vectors of length N_2 and N_0, respectively, with components restricted to the binary set $\mathbf{B} = \{-1, +1\}$. The weights of the connections between the input and hidden layers form an $N_0 \times N_1$ matrix denoted by $\mathrm{W}^{(1)}$, whereas the weights of the hidden to output layer connections form an $N_1 \times N_2$ matrix $\mathrm{W}^{(2)}$. In general, the interconnection weights of both layers are modifiable, so that the system can be trained to perform a desired transformation from the input space to the output space. In SDM, however, the first layer acts as a fixed-weight preprocessor, encoding each N_0-bit input into a very large N_1-bit internal representation, $N_1 \gg N_0$. The second layer is a trainable sum-of-outer-products network, which is programmed to recognize the higher-dimensional internal representations. Kanerva's primary contribution is the specification of the preprocessor, i.e., how to map each N_0-bit

input into a very large N_1-bit internal representation in such a way as to permit the network capacity to exceed by far any linear relationship with the input dimension. This is important because in most applications, the dimension of the input (which is approximately equal to the capacity of a single-layer machine) is much smaller than the number of patterns we wish to recognize.

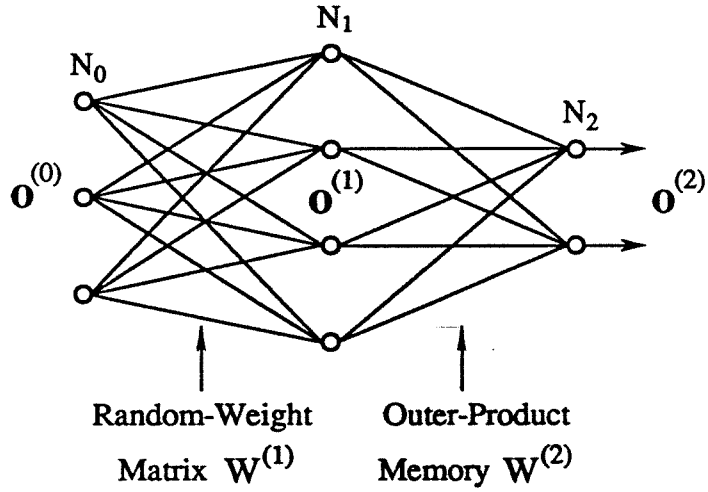


Figure 6.4. Kanerva's Sparse, Distributed Memory (SDM) model.

The operation performed by each hidden neuron is thresholding. Specifically, if we denote by $f_\theta(U)$ the neuron response function with U being the input to the neuron and θ being the threshold, then $f_\theta(U)$ is 1 if $U \geq \theta$ and 0 if $U < \theta$. The weight matrix $\mathrm{W}^{(1)}$ is populated at random by +1's and -1's. The input vector to the hidden neurons is given by the matrix-vector product $\mathrm{W}^{(1)}\mathbf{o}^{(0)}$, which is thresholded by the function f_θ to become the output vector $\mathbf{o}^{(1)} = f_\theta(\mathrm{W}^{(1)}\mathbf{o}^{(0)})$

of the hidden neurons. With $\theta = N_0 - 2r$, the N_1-bit word $\mathbf{o}^{(1)}$ contains a 1 in the ith coordinate if and only if $\mathbf{o}^{(0)}$ is within Hamming distance r of the ith row of $\mathrm{W}^{(1)}$. If the parameters r and N_1 are set correctly, then the number of 1's in the representation $\mathbf{o}^{(1)}$ will be very small in comparison to the number of 0's. Hence $\mathbf{o}^{(1)}$ can be considered as a sparse, distributed representation of $\mathbf{o}^{(0)}$: sparse because there are few 1's, distributed because several 1's share in the representation of $\mathbf{o}^{(0)}$.

The overall SDM can be regarded as a sum-of-outer-products associative memory operating on the sparse, distributed representation of $\mathbf{o}^{(0)}$. We define a function $g : R^{N_2} \rightarrow R^{N_2}$ to be the vector signum function, which takes the sign of each coordinate independently. Then the response of the output neuron is $\mathbf{o}^{(2)} = g(\mathrm{W}^{(2)}\mathbf{o}^{(1)})$, where the synaptic weight matrix $\mathrm{W}^{(2)}$ is given by

$$\mathrm{W}^{(2)} = \sum_{j=1}^{M} \mathbf{y}^{(j)}[f_\theta(\mathrm{W}^{(1)}\mathbf{x}^{(j)})]^t. \tag{6.2}$$

It was shown by Chou [113] that by allowing N_1, the dimension of hidden layer, to grow exponentially with the input dimension N_0, the capacity of the SDM can grow exponentially in N_0, achieving the universal upper bound of any associative memory. This is in sharp contrast to the capacity of a single-layer associative memory, which grows at most linearly with the input dimension. In terms of pattern recognition, large N_1 implies mapping input vectors into a higher dimensional space so that it is much easier to find the appropriate decision boundaries. In this way, a linearly nonseparable problem can be converted into a linearly separable one at the hidden layer [114].

The optical implementation of a two-layer neural network trained by SDM

requires both fixed and modifiable interconnection matrices. Dynamic volume holograms recorded in photorefractive crystals are very promising candidates for the implementation of such interconnection matrices because of the three dimensional storage capacity possible within the volume of a crystal [7], the well-studied dynamic response of photorefractive crystals and the ability to fix photorefractive holograms. Nonlinear effects, such as fanning in photorefractive crystals, generally a nuisance, are helpful for the implementation of the random interconnection matrix in the first layer. Optically addressed spatial light modulators (SLMs) with nonlinear thresholding and amplification functions can be used to simulate neural response. In our experiment, liquid crystal light valves (LCLVs) manufactured by Hughes are used both for providing the input and gain, as well as for use as thresholding devices.

6.2.3 Handwritten Character Recognition Application

The problem that was selected to test the operation of the system is handwritten character recognition. 104 training patterns were created by drawing characters on a 10×10 pixel grid (4 character patterns for each letter of the alphabet A-Z). The optical system architecture is shown in Fig. 6.5. The interconnections between the layers were implemented with Fourier transform holograms recorded in two $LiNbO_3$ photorefractive crystals (PR1 and PR2) using an argon-ion laser ($\lambda = 514$ nm). The input layer consists of a video monitor (VM) and a liquid crystal light valve (LCLV1). There are 100 input units, matching the size of the 10×10 grid for the character patterns. Input patterns are presented on VM by a computer, imaged onto the LCLV1 by an imaging lens (L1), and read out by the

laser beam on the other side of the LCLV1. The hidden layer, implemented by a second liquid crystal light valve (LCLV2), consists of an array of approximately 300 $\times$ 300 neurons. There are 26 output neurons for this system, represented by 26 pixels in a CCD detector array, each responding to one letter of the alphabet.

The method used to train the network is a modification of the SDM model. According to this method, the weights of the first layer are selected at random. The weights of the second layer are trained by presenting the training patterns at the input of the network that induce a response at the hidden layer through the random connections. If, for the current input, the desired response for an output neuron is high, the response of the hidden layer is added to the second layer weights leading to that particular output neuron. This simple procedure is repeated for all the patterns in the training set.

During the training of the first layer, random dot patterns were presented at the input as training patterns. Each random dot pattern was split into two parts, and both were Fourier transformed by the lenses L2 and L3. These two Fourier transformed random patterns were used to record a hologram which consists of gratings of random strength. This process was repeated many times so that a volume hologram with random interconnection weights was recorded. Furthermore, in the crystal we used, the photorefractive nonlinearity is sufficiently strong that a laser beam passing through the crystal loses much of its power to a broad fan of light resulting from amplification of radiation scattered by imperfections in the crystal [115]. This phenomenon, called beam fanning, further randomized the recorded interconnections and at the same time drastically increased the number of hidden neurons that input neurons are connected to. The writing beams in the

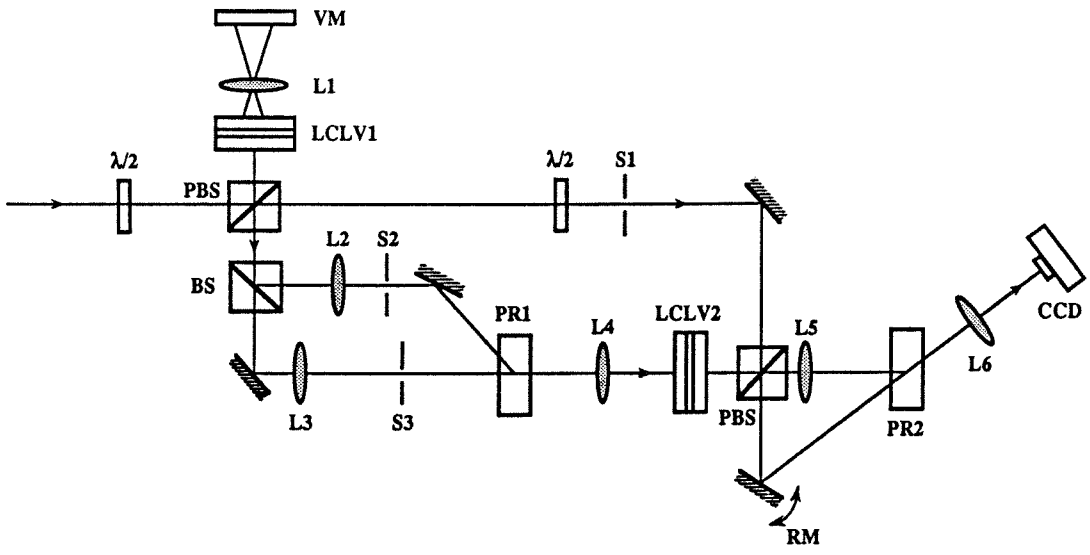


Figure 6.5. Optical two-layer network. VM = video monitor, LCLV = liquid crystal light valve, PR = photorefractive crystal, (P)BS = (polarizing) beam splitter, RM = rotating mirror, L = lens, S = shutter.

first layer were polarized in the extraordinary direction with respect to the crystal in order to obtain maximum fanning. In our experiment, each of the input neurons were connected to about 10^5 hidden neurons. Therefore the resulting weight matrix performs a dimensionality-expanding random mapping, which is exactly what is needed in the implementation of the SDM model. After the first layer training was completed, the random interconnection hologram was thermally fixed by heating the crystal to 100 °C for 30 minutes [87] and training of the second layer was then started.

The goal of the second layer training is to ensure that when a character pattern is presented at the input of the network, one of the 26 output neurons, with spatial position proportional to the order of that letter in the alphabet, will be switched on. This was achieved by training the second layer using the sum-of-outer-products rule. During this process, the 104 training patterns were sequentially presented at the network input and randomly mapped into higher-dimensional hidden representations. These hidden representations were amplified and thresholded by LCLV2 and Fourier transformed by lens L5. Their Fourier transform holograms were recorded in association with plane wave references with appropriate propagation directions. The directions of these reference beams are chosen according to the identity of the input patterns. The reference beam transmitted through the crystal is focused by lens L6 onto a different position on the CCD, with the angle of the reference beam determining the position of the focused spot. Therefore, by selecting the proper angle for the reference beam, the response of the hidden layer is added to the weights of the interconnections leading to the output neuron that is responsible for the current input pattern. The reference

beam angle is selected by rotating a mirror, which was mounted on a motorized rotary stage controlled by the computer. The photorefractive crystal was exposed 104 times to record the desired interconnection pattern.

In order to compensate for the hologram decay associated with multiple exposures in photorefractive crystals, the exposure schedule described in chapter 2 was followed during the learning process so that all the holograms were formed with equal strength. This leads to the implementation of the sum of outer products in Eq. (6.2). The crystal we used for the second layer was an 8 mm thick $LiNbO_3$, doped with 0.01% Fe. Under our experimental condition, the time constant τ was measured to be 425 seconds. During the network training, internal representations of the 104 training patterns need to be recorded in the second-layer crystal with roughly equal diffraction efficiencies. The exposure time for each of these holograms except the first one can be calculated from Eq. (2.37). For example, $t_2 =$ 295 s and $t_{50} = 8.6$ s. t_1 was chosen to be 25 min so that $t_1 \gg \tau$ and the first hologram reached the saturation diffraction efficiency. Therefore, with $M = 104$, the total exposure time is $T \approx 58$ min.

Another important issue is the finite angular bandwidth of volume holograms. If the angular separation between the reference plane waves is too small, the presentation of any character pattern at the input may reconstruct several plane waves so that several output neurons (corresponding to these reference waves) will be turned on. This leads to crosstalk and possible misclassification. The angular separation, however, cannot be too large because of the limitation of optics. To find an appropriate angular separation, we need to examine the angular bandwidth of volume holograms in the crystal, which is given by [116]

$$\Delta\theta_c \approx \frac{\lambda}{2n_c d sin\theta_c}, \tag{6.3}$$

where λ is the laser wavelength in vacuum, θ_c is the angle between the normal of the crystal surface and the propagation direction of the reference beam inside the crystal, and d is the hologram thickness. In our experiment, the angle of incidence of the writing beams in the air is $\theta_0 = 20^o$ and the index of refraction of the $LiNbO_3$ crystal is $n_c = 2.20$. Therefore θ_c can be solved from

$$n_c sin\theta_c = sin\theta_0, \tag{6.4}$$

which gives $\theta_c = 8.94^o$. With $\lambda = 0.514\mu m$, $d = 8mm$ and using Eq. (6.3), $\Delta\theta_c = 0.0054^o$. Finally, we can find the angular bandwidth in the air by differentiating Eq. (6.4), which yields

$$\Delta\theta_0 = \Delta\theta_c n_c cos\theta_c / cos\theta_0 = 0.0125^o. \tag{6.5}$$

To make sure that crosstalk due to the finite angular bandwidth is greatly suppressed, we chose the angular separation between reference beams to be 0.03^o. Therefore the total angular sweep of the reference beam is $26 \times 0.03^o = 0.78^o$, which is reasonable for the motorized rotary stage and at the same time guarantees that the two writing beams overlaps in the crystal for all reference beam angles.

Once the training is complete, the presentation of any one of the training patterns causes the second hologram to reproduce the reference beam with which it was recorded. This reconstructed beam codes, in the angle of propagation, the

identity of the pattern. The final lens in the system focuses the reconstructed beam to an output neuron whose position in the output array is proportional to the angle of the reconstructed beam.

6.2.4 Experimental Results

A photograph of the experimental system is shown in Fig. 6.6. After training, all the 104 training patterns were tested and recognized correctly by the system. Figure 6.7 shows three examples of the input patterns, their internal representations, and the responses at the output of the optical system. The input patterns shown in Fig. 6.7 were among those used for training the network. The bright dot in each example indicates the position of the switched-on output neuron. As can be seen, crosstalk was completely suppressed in these cases, mainly due to the drastically expanded dimensionality of hidden representations and the nonlinear thresholding operation of the neurons. We can also observe the differences between hidden representations for different input patterns.

To check the generalization property of this trained network, 520 handwritten character patterns (20 patterns from each class) that were not in the training set were presented to the optical network and the identity of each pattern was determined from the position of the output neuron that had the maximum response. Figure 6.8 shows some of the testing patterns and the result is summarized in Fig. 6.9, which gives the number of correct classifications out of 20 tests for each class. It turned out that 311 out of the 520 testing patterns were correctly classified, giving an average recognition rate of about 60%. This recognition rate is much better than random guessing (4%), but far below what is required for a useful

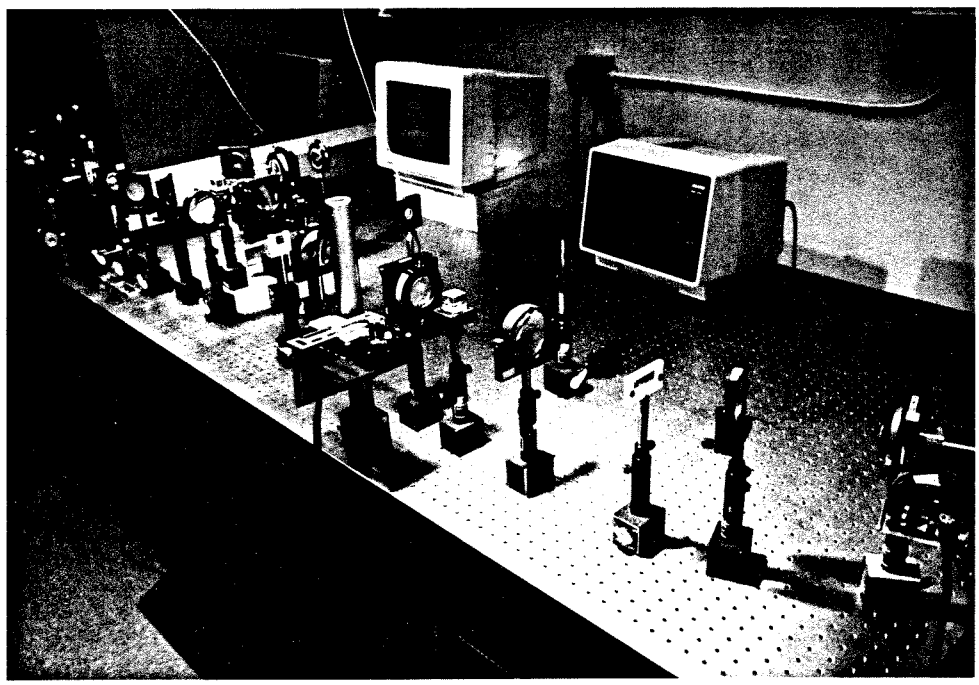


Figure 6.6. Experimental apparatus.

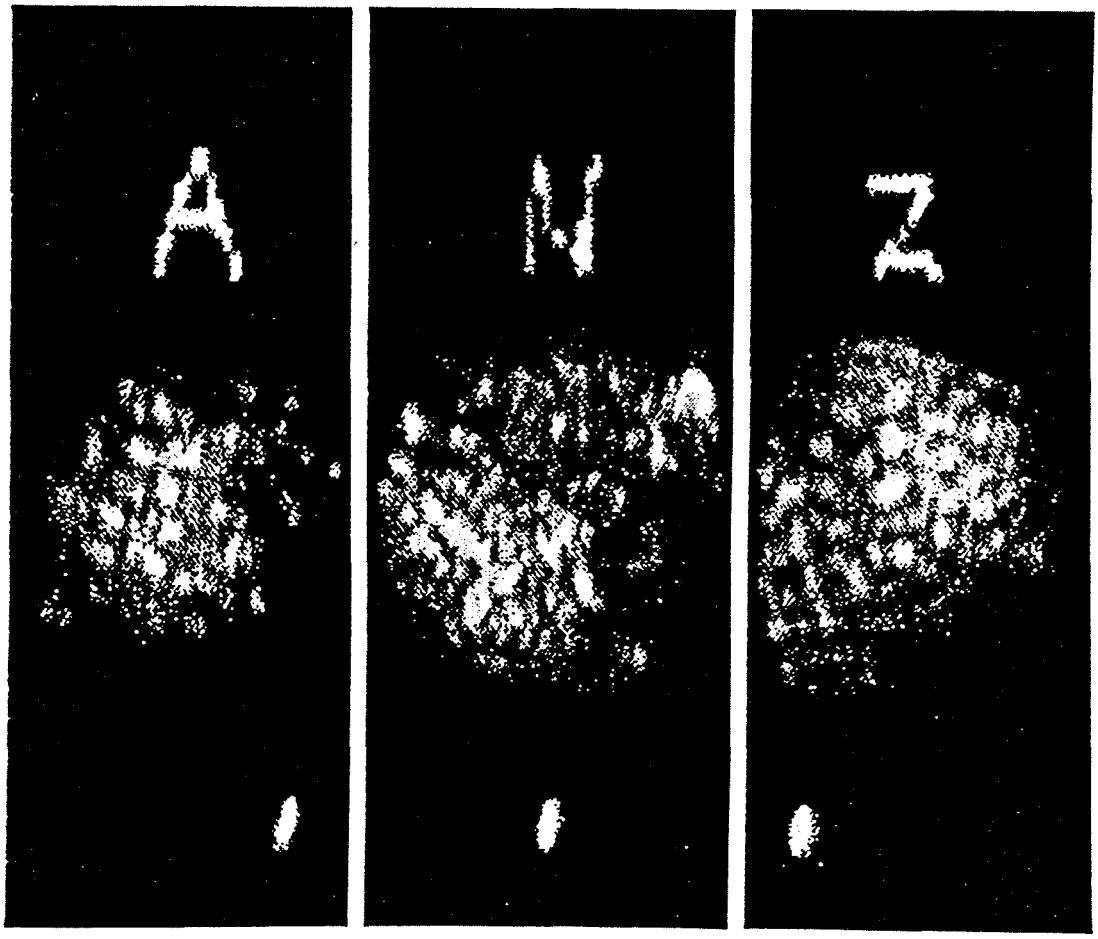


Figure 6.7. Examples of the signals at the input (top), hidden (middle), and output (bottom) layers in the experimental system.

character recognition system. The reason for the relatively poor performance on the test set is the choice of training algorithm used, specifically the fixed first layer weights and the limited number of training cycles for the second layer. This same system can be used to implement algorithms in which both layers are fully trained in response to the training patterns, which, in computer simulations, give much better performance.

6.3 Optical Two-Layer Network for Real-Time Face Recognition

6.3.1 Introduction

This section describes the experimental demonstration of a two-layer optical network that accepts input images of faces at standard video rates and classifies them in real time. The optical system is a modified version of the standard holographic multilayer architecture [61,62]. Its second layer has fixed weights, whereas the first layer interconnections, implemented with holograms stored in a photorefractive crystal, are modifiable.

The algorithm that we use is a hybrid. It has features of local-representation algorithms in that each hidden unit is trained separately and the training method is not iterative. On the other hand, the resulting representations are distributed. It was found that distributed representations are crucial for two reasons. First, when the optical network was trained with purely local representations, we found that it became extremely susceptible to noise and the performance deteriorated very rapidly as the number of hidden units increased. This is because in a purely local representation, only one hidden unit is on at a time. Since the output is formed as a linear combination of all the hidden units, a small amount of noise from

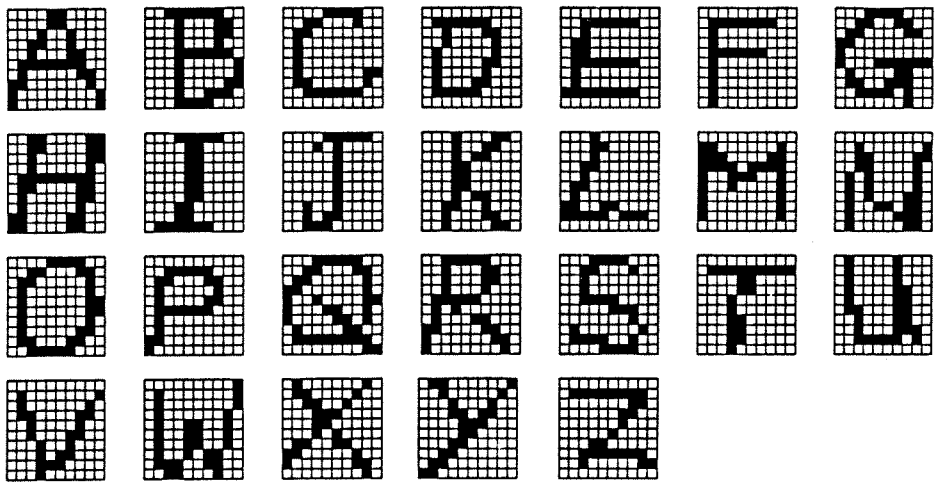


Figure 6.8. Examples of the test patterns.

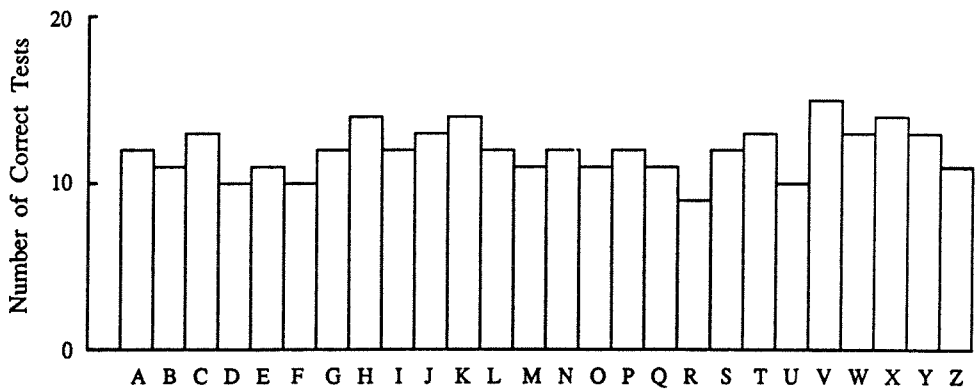


Figure 6.9. Histogram of the test results.

each hidden unit will ultimately overwhelm the signal term as more hidden units are added. Poor generalization performance is the second reason to avoid purely local representations. We found that by switching to distributed representations the system performed much better when presented with images it had never seen before.

In the next section the optical architecture and the overall experimental set-up are described. Section 6.3.3 describes the training algorithm and the details of the training procedure. Section 6.3.4 describes the performance obtained with the network.

6.3.2 Experimental Apparatus

The optical set-up is shown in Fig. 6.10. It is a two-layer network with an optical preprocessing stage that performs edge enhancement. The input device to the network is a liquid crystal TV (LCTV) that has a resolution of 320 by 240 pixels and 2×2.5 cm^2 clear aperture . This device was extracted from an EPSON television projector. The LCTV is illuminated with collimated light from an Argon laser ($\lambda = 488$nm). Lens L1 produces the Fourier transform of the input image at plane P2. A spatial filter is placed at P2 to accomplish two goals. It blocks the higher diffracted orders that result from the pixelation of the LCTV. The removal of the higher orders gives a smoother, less noisy image but it reduces the light efficiency of the LCTV. The second function of the spatial filter in plane P2 is to block the low frequency components of the input image which enhances the edges of the input image and dramatically improves the ability of the system to discriminate between inputs from different classes. A photograph of the spatial

filter is shown in Fig. 6.11. It consists of a cross-hair and a DC block for high pass filtering. The purpose of the cross-hair is to remove the diffraction pattern at P2 due to the sharp edges formed at the boundary of the actual area of the LCTV. This boundary, when edge enhanced, yields a very strong rectangle that is common to all inputs and makes discrimination difficult. The diameter of the DC block is 260 μm. Given the wavelength of light and the focal length of L1 (F_{L1} = 50 cm) we can find the cut-off frequency to be .533 lines/mm. Roughly speaking, features in the input plane that are smaller than 1.9 mm are highlighted in the edge-enhanced image. An iris (not shown in Fig. 6.11) is used to block the higher orders not blocked by the crosshair. An example of an image of a face and the edge enhanced version of it that was produced by the optical system is shown in Fig. 6.12.

Lens L2 images with magnification 1 plane P2 onto plane P3, the plane of the hologram. The size of the spectrum on the hologram is approximately 5 mm in diameter. The hologram is formed by introducing a plane-wave reference. The angle between the signal and reference beams varies from 29 to 31 degrees, outside the crystal. The reference beam is reflected off a mirror mounted on a computer controlled rotation stage. The plane of the rotating mirror is imaged onto the crystal with a unit magnification 4-f system that allows the angle of the reference beam to be scanned without moving the position of the reference beam on the crystal. The crystal is an iron doped $LiNbO_3$, with doping level 0.01%. The c-axis of the crystal is in the horizontal direction in Fig. 6.10. The crystal dimensions are $20 \times 20 \times 8$ mm^3.

Lens L4 is a Fourier transform lens that produces an image of the edge

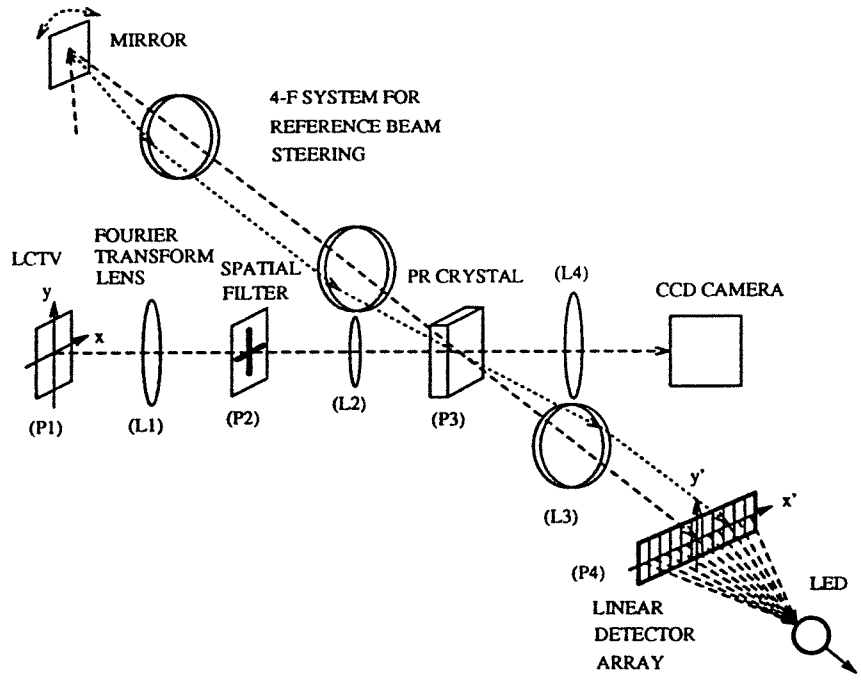


Figure 6.10. Optical set-up of the face-recognition system.

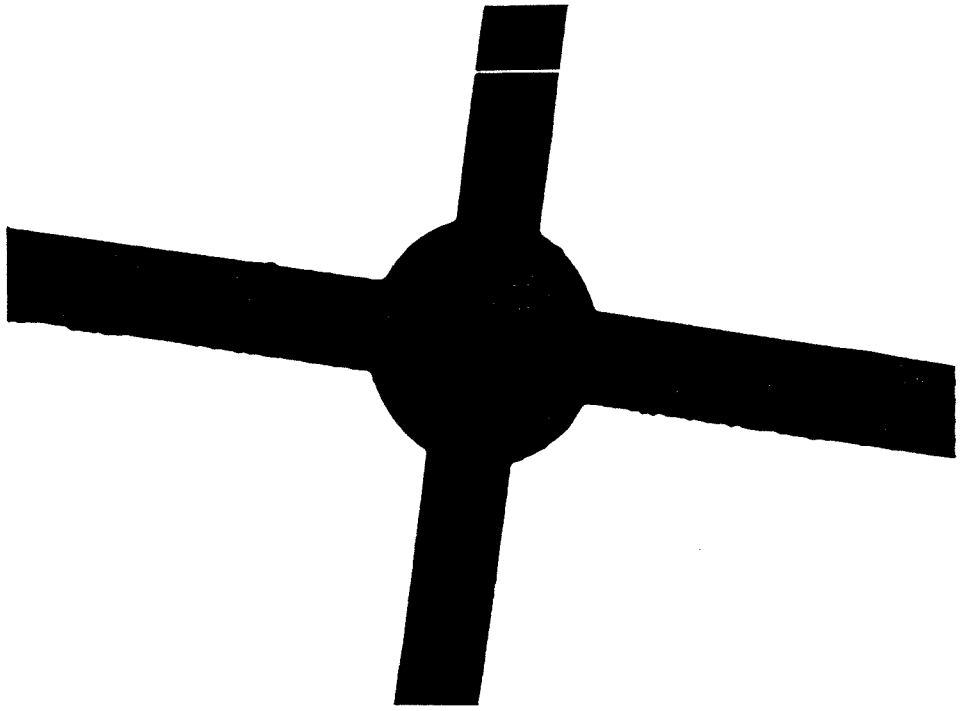


Figure 6.11. Spatial filter used in plane P1 of Fig. 6.10.

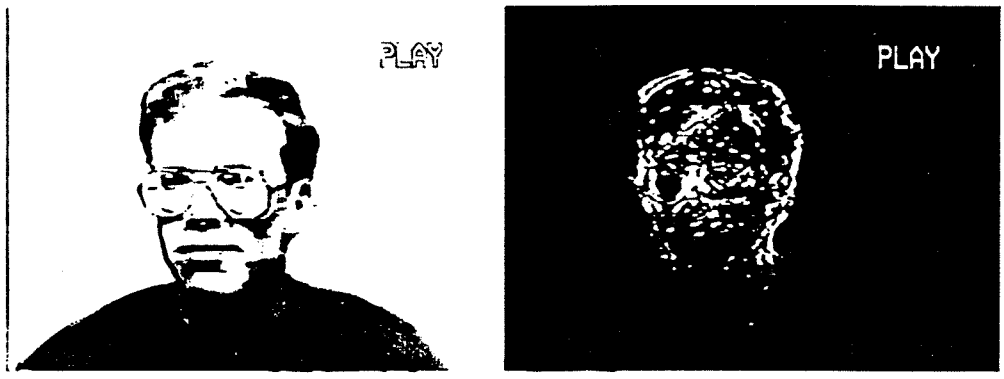


Figure 6.12. Edge enhanced image and original face.

enhanced input image on a CCD for visual assessment. Lens L3 is also a Fourier transforming lens that produces at the output plane P4 the response of the first layer where it is sensed by a linear detector array. A beamsplitter placed in front of the array diverts a portion of the light to a CCD camera so that the output of the first layer can be visually monitored. Functionally, the system from the input plane P1 to P4 is an array of image correlators with 1-D shift invariance. To understand this, consider the case where a single hologram is recorded in the crystal at a particular angle of the reference beam. In this case the system is a classic VanderLugt [26] correlator except that a volume hologram is used and the input has been high-pass filtered. The effect of the volume hologram is to eliminate shift invariance in the horizontal direction in Fig. 6.10. This happens because a horizontal shift at plane P1 will change the angle of incidence at plane P3 and cause the hologram to be Bragg mismatched [117–119]. Specifically, the light distribution at plane P4 is given by [117]

$$g(x',y') = \int\int f(x,y)h(x-x',y-y')dxdy \ \ \mathrm{sinc}(\alpha x'), \tag{6.6}$$

where f(x,y) and h(x,y) are the input and filter functions, respectively. The input coordinates are (x,y) and the output coordinates are (x',y'). The thickness of the crystal is L, θ is the angle of the reference beam and

$$\alpha = \frac{L\sin\theta}{2\lambda F}. \tag{6.7}$$

We see from Eq. (6.6) that the effect of the thick hologram is to mask off the 2-D correlation pattern except for one vertical strip whose position depends on the angle of the reference beam. The amount of shift invariance that can be tolerated

in the horizontal direction is approximately equal to $1/\alpha$ plus the width of the correlation peak in the horizontal direction. The system retains its shift invariance in the vertical direction. If we change the angle of the reference beam and record a different hologram at each angle then the 1-D strip of the 2-D correlation function will be produced at a different horizontal location. In the experiment that we will describe, holograms are recorded at 40 separate angles separated by 0.05°, yielding a system that has 40 correlators with 1-D shift invariance.

The experiment in Fig. 6.13 demonstrates the operation of this part of the system. In this case each filter was a recording of the face of the same person at different scales. What is shown in Fig. 6.13 is the input to the network for 4 different size images, along with the corresponding response at the right-hand side of each picture. We see that as the size increases the strongest response of the system is at different vertical positions. In the optical set-up, the correlation responses shown in the right-hand side of each picture is actually horizontal, and the display was created by simply rotating the CCD camera by 90 degrees.

The role of the second layer is two-fold. The first task is to take advantage of the vertical shift invariance of the first layer and the second task is to combine the outputs of the 40 correlators and make the final classification. We will discuss first the shift invariance. Suppose that an image at a particular location at the input produces a strong correlation peak somewhere at the output. If the input is horizontally translated by approximately 0.4 mm then the correlation peak disappears. If the input is translated vertically then the correlation peak moves vertically also. What we really need for shift invariant recognition is a system whose output does not change as the input shifts. To accomplish this we use long

Figure 6.13. Experiment showing the position of the correlation peak to be proportional to the size of the input face.

detector elements in the vertical direction as shown in Figure 1. These long detectors collect the correlation peak and continue to produce a strong output signal as the input image shifts vertically. Unfortunately, we cannot use an arbitrarily long detector element to obtain full shift invariance vertically because then the detector would simply collect all the diffracted energy from the corresponding filter stored in the hologram. Roughly speaking, all input signals with the same total energy would yield the same response. A shorter detector responds more selectively to the correlation peak, and hence the degree of match between the input and the reference but it sacrifices shift invariance. Thus there is a basic trade-off between shift invariance and discrimination capability. In our network we made this compromise by trial and error. By repeating the experiment with a horizontal slit of varying width placed in front of the detector array, the amount of shift invariance in the vertical direction is roughly 3 mm or equivalently 12 percent of the size of the input image. As we will see later on this choice yields good discrimination capability.

The second layer also puts together all the vertically integrated responses from the first layer and produces the final output. Since the output of the detector array in plane P4 is electronically available we can implement the second layer either electronically or optically. We have done both with comparable performance. The optical implementation of the second layer is realized by thresholding the output of the detector array and then feeeding it to a second LCTV. The inner product between the signal recorded on the LCTV and a weight vector stored in the form of a transparency is then optically formed. This inner product is electronically thresholded to produce the final output. In the current system we

describe in this paper, the operations of the second layer are so simple that it was easier to do them electronically. Specifically, all the weights of the second layer have the same value. In other words , the second layer simply integrates the output of the first layer. The electric signal from each detector is the square of the light amplitude of the total signal incident at each element. The signal from the detector can be thresholded electronically. However, we get the best performance by simply using the square-law non-linearity. In this case, the system becomes similiar to a quadratic associative memory [48,120]. Notice that the nonlinearity performed at plane P4 is crucial in this system. If the outputs of all the correlators from the first layer were somehow coherently added without the inclusion of the nonlinearity then the overall system would simply be equivalent to a single correlator.

A schematic diagram of the overall system is shown in Fig. 6.14. The input images are detected by a standard television camera. The video signal is either stored on a video cassette recorder (VCR) to form a training set or fed directly to the LCTV during real time operation. The 2-layer optical network is the system we described above. A personal computer controls the experiment during the training phase by instructing the VCR to advance the video by one frame and pause so that the training algorithm can be executed in the optical system. The output of the hidden layer determines whether the hologram should be modified by the current input image. If a holographic exposure is needed the computer opens two shutters (one for the signal and one for the reference beam) for a specified time and the hologram is recorded. During the execution of the algorithm the computer also controls the angle of the reference beam, so that different hidden units can

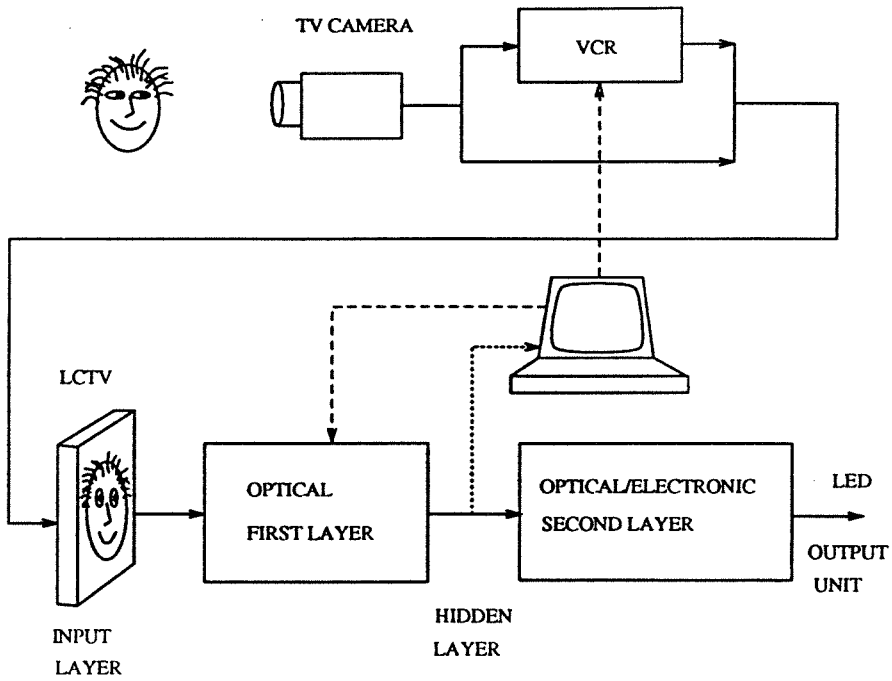


Figure 6.14. Schematic diagram of overall system.

be trained. After the training is completed, the computer is no longer involved in the operation of the system except to record the output data if desired.

6.3.3 Training Procedure

The training algorithm that we use is partially motivated by the tiling algorithm [106]. In the tiling algorithm, individual units are trained separately for a fixed number of iterations. Once a unit is trained the algorithm moves on to a new unit and trains it to make up for the deficiencies in the performance obtained with the previous units. In this way networks with multiple layers and many neurons per layer can be built-up and trained. In the standard tiling algorithm each unit is trained with the perceptron algorithm with the entire training set. In our algorithm each unit is trained by a subset of the training set that consists of similar images. This similarity measure is enforced by training each unit to respond to a contiguous short segment of the training video. In this way, each unit is trained to respond to a specific aspect of the input face. This simplifies the training of individual units and the overall training procedure results in networks of predictable size.

The flow chart for the algorithm we use is shown in Fig. 6.15. We describe more specifically the algorithm. Let $\mathbf{f}^k$ denote the k-th image in the training sequence stored in the VCR and let w_{ij} denote the weight of the first layer connecting the i-th input pixel to the j-th hidden unit. The training algorithm is as follows:

set $e = 0$ (e is the number of exposures per hidden unit)
set $j = 1$ (j enumerates the hidden units)
while ("there are more training examples")

do { (go through the training set one frame at a time)
$h = 0$ (h is the number of hidden units turned on)
for $j' = 1$ to j, if $\sum_{i'=-I/2}^{I/2} |\sum_i f_i^k w_{i-i',j}|^2 > \theta$ then $h = h + 1$
(count the number of hidden units that are on)
if ($h < H$ and $\sum_{i'=-I/2}^{I/2} |\sum_i f_i^k w_{i-i',j}|^2 < \theta$)
(less than H hidden units are on, and the current unit is off)
then $w_{ij} = w_{ij} + f_i^k$ and $e = e + 1$ (make an exposure)
if ($e > E$) (more than E exposures on current unit)
then $j = j + 1$ and $e = 0$ (create new hidden unit)
"go to next frame"
}

The user must select the parameter θ, H, and E before the algorithm begins. In what follows we will explain the role of each parameter and how it affects the performance of the trained network. The variable j counts the hidden units. We begin training the first unit ($j = 1$) by presenting frames to the system in sequence (incrementing k). The k-th input is added to the weights of the first unit if the response of the first hidden unit is below a threshold θ. Notice that in the optical system the response of the hidden unit is not simply the inner product between the input and the weight vector but an integration over I pixels of the center of the correlation function, as we described earlier. If θ is set too high then the units become very highly tuned to respond to the particular images they are trained for. If the threshold is too low then too much cross talk with unfamiliar faces results leading to erroneous classifications. Ideally, θ should be lowered as the training proceeds and hidden units are added, since this weakens all the stored holograms. In the experiment we describe we used a constant θ. The first unit continues to accumulate training examples in this way until a total of E exposures have been made to it. At that point a new hidden unit is created (j is incremented) by

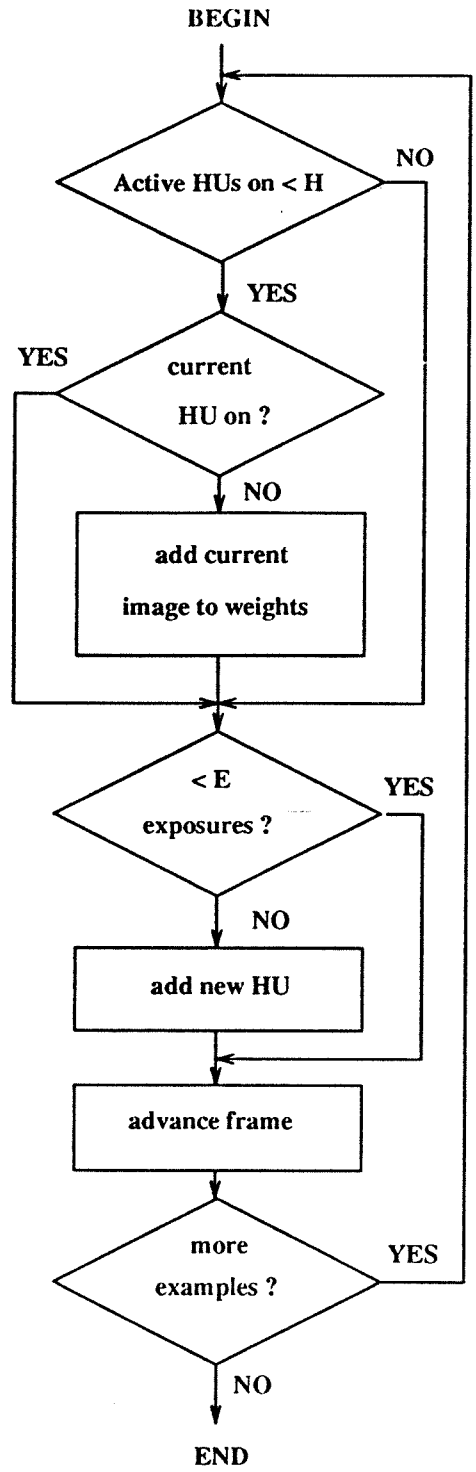


Figure 6.15. Flow chart for the algorithm used to train the network. (HU means "hidden unit").

rotating the mirror that controls the angle of the reference beam. We would like to have E large in order to have each unit be responsive to as many training examples as possible. However, since we are only presenting positive examples to the system (i.e., we never subtract anything from the weights but always add to them), if too many examples are accumulated the weight is simply the average of the subject's faces, which is similar to the average of anybody's face, and loses its discrimination capability. The first H hidden units are trained in exactly the same manner as the first. When j exceeds H the current input frame is added into the weights of the j-th hidden unit only if fewer than H units are above threshold. If H is set to 1, then the training of the early units is identical to the rest. However, this results in a hidden layer response that has only one unit on at a time. We have already commented that we found that this results in poor performance on the training set due to susceptibility to noise and poor generalization. By requiring that at least H hidden units are on at any one time for the training set, we improve the robustness of the system and improve generalization. If H becomes too large, we would need too many hidden units to enforce this requirement, and the encoding becomes inefficient.

The discussion in the above paragraph describes the basic trends that we predict and experimentally observe as the parameters E, H, and θ are adjusted. The experiment that we will describe in this paper was carried out with $H = 3$, $E = 6$, and θ was set equal to 3 times above the noise background level. These values were arrived at empirically by running the experiment several times and measuring the generalization performance. The system performance is sensitive to the setting of θ (it should be set relatively low), but not as sensitive to changes

in H and E. These settings worked best for all the face recognition experiments we tried. Unfortunately, there is no guarantee that these settings are the best for other problems.

The most attractive feature of this algorithm is that it can be easily implemented with the optical system described in the previous section while yielding remarkably good classification performance, as we will see in the next section. The algorithm requires two basic operations from the optical system: Evaluation of the response of the hidden units to an input image so that the computer can compare it to a threshold and addition of the current image into the hologram that specifies the weights of that unit. We have already described how the system evaluates the response of the hidden units. We will discuss here how the weight updates are performed. When a hologram is exposed to light the strength of an individual holographic grating (or connection w_{ij} is modified according to the following equation :

$$\tau \frac{dw_{ij}}{dt} + w_{ij} = \beta m_{ij}, \tag{6.8}$$

where τ is the time constant of the holographic recording in the photorefractive crystal, β is a constant that depends on the crystal properties, and m_{ij} is the modulation depth of the frequency component of the illuminating light that matches the grating w_{ij}. For a short light exposure of duration Δt, we can approximate the change in the hologram by

$$\Delta w_{ij} \approx -\frac{\Delta t}{\tau} w_{ij} + \frac{\Delta t}{\tau} \beta m_{ij}. \tag{6.9}$$

In other words, each exposure reinforces each weight in proportion to the strength

of the corresponding frequency component of the illuminating light. However, each exposure also erases all the weights in proportion to their current strength. This is the well known weight decay problem that plagues photorefractive memories [73] and photorefractive neural networks [53]. Several solutions to this problem have been proposed [79–84]. We use a simple exposure schedule in our experiment, in which the later exposures are linearly shortened to compensate for the decay of the earlier holograms resulting in an approximately uniform final recording. Specifically the m-th exposure, t_m, is set equal to $t_m = 3 - m/240$ seconds. Thus the exposures varied from 3 seconds at the beginning of the exposure sequence to 2 seconds at the end, with a total light intensity equal to 10 mW/cm^2, a modulation depth approximately 0.1.

The training set for the experiment was a video recording of the face of one of the authors (Yong Qiao) moving his head in front of the camera, turning, nodding, tilting his head, smiling, etc. The total number of images in the training set is 5,400 frames. The execution of the algorithm modified the hologram with only 240 of these images. The rest produced an acceptable hidden layer response. Since each hidden unit receives 6 exposures, a total of 40 hidden units were created. The maximum number of hidden units that the system can support is limited by two factors. One is the dynamic range of the photorefractive hologram. In this case a total of 240 holograms are superimposed. If we assume that all these exposures are statistically uncorrelated (i.e., each exposure simply erases all the previously recorded holograms and does not ever reinforce them) then the diffraction efficiency of each hologram would fall by a factor equal to $(240)^2$ [53] compared to the efficiency with which a single hologram is stored. Since up to 5,000 [12] holograms

have been superimposed in lithium niobate crystals, the dynamic range was not a problem in our experiment. The second limitation is the numerical aperture of the optical system to allow all the reference beams to enter the crystal. The system we used in the experiment had the capability to implement in excess of 100 units and it is possible to build systems with more than 1,000 units. Therefore, this particular training set did not stretch the limits of the system's capabilities. The entire training cycle lasted about 40 minutes, which includes the time for hologram exposure and controlling the system by computer.

Shown in Fig. 6.16 is a composite photograph showing a short sequence of the training session. Each picture in the composite shows the current input frame and on the right, vertically displayed, is the optical response of the hidden units. The first event in the sequence is on the top left in Fig. 6.16 and it shows the frame shortly after the hologram is exposed. As time progresses the hidden layer response changes (upper right corner) and gradually dims (lower right corner). Ultimately, there are fewer than 3 units on and the system is triggered to make another exposure (lower right corner). The white ribbon on the left of the input image where the hidden layer normally appears, indicates that the hologram is being exposed to light and the camera that monitors the hidden layer response is flooded with light.

6.3.4 Classification Performance

In this section we describe the performance of the trained network. Once the network is trained it operates in real time, processing 30 frames per second directly from the input TV camera. The outputs from the detector array are

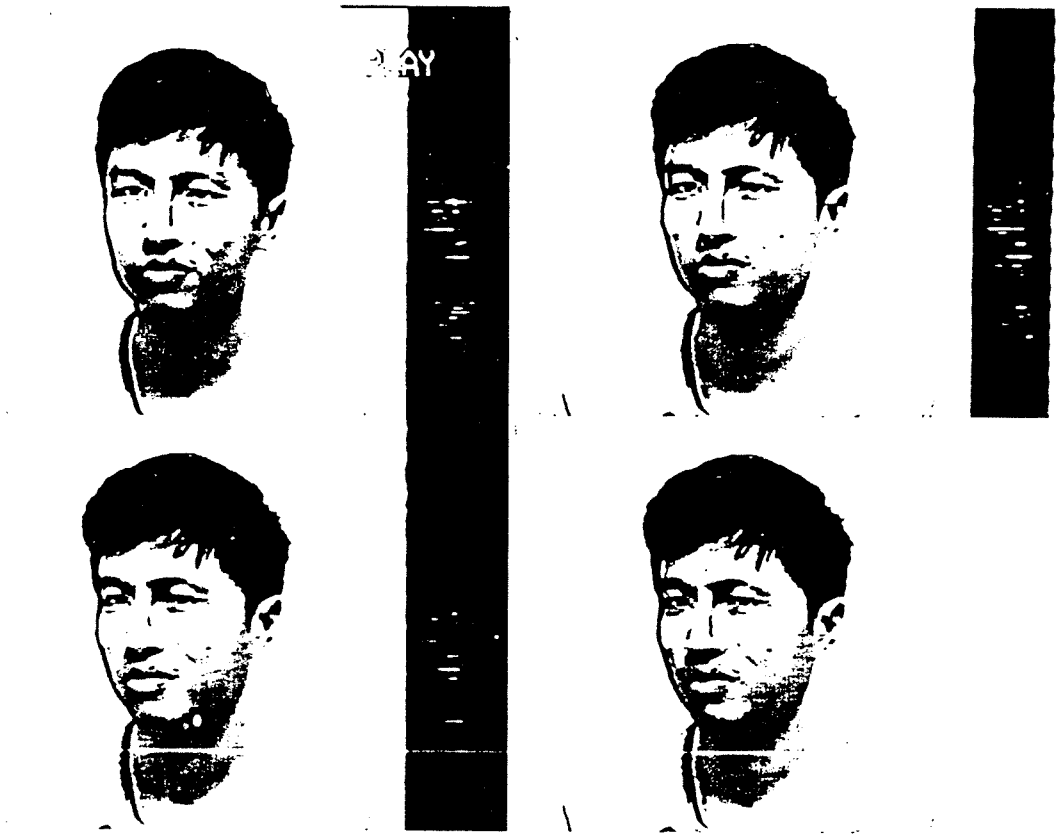


Figure 6.16. Photographs showing part of the training session.

simply added together electronically and this sum is then thresholded to produce the final output. The holograms will decay when exposed to light during the testing phase. We can overcome this by either thermally fixing the hologram [87] or by using dynamic copying described in chapter 3. In this experiment we adopted a simpler route that temporary overcomes this problem. By reducing the readout light intensity by a factor of 20, compared to the total writing intensity, we can calculate that the holograms will decay after several hours of constant illumination.

The holograms were sufficiently strong that the reduction in the readout intensity yielded sufficient signal at the detector. The system was tested with the training set and a wide variety of test sets, including Yong, presented to the system under various conditions and others attempting to confuse the system. Shown in Fig. 6.17 is the signal at the output of the system before final thresholding. The entire recorded presentation shown in Fig. 6.17 lasts for about 10 minutes. The first minute is a portion of the training set. The next 2 minutes is a real-time input of Yong who looks into the TV camera and moves around in a manner similar to the training set. While he does this, he does not have access to any information from the network. The rest of the sequence is the response of the system to two other persons (Sid and Allen). We can see that the average response is highest for the training set, and it remains almost as high for the rest of the time when Yong is the input. The average response for the other two subjects is markedly lower. The variance of the response is higher for Yong, because he was exhibiting a wider range of head perspectives, compared to Sid and Allen, to test the limits of the system. Similar behaviors were observed for all 14 members of our group.

To make the final classification, we need to threshold the signal shown in Fig. 6.17. In the actual system, this is done electronically in real time. The optimum threshold was determined from the data shown in Fig. 6.17. Shown in Fig. 6.18 is a plot of probability of error as a function of the output threshold level. The three curves correspond to the probability of error for Yong, Sid and Allen, estimated by classifying the data in Fig. 6.17 with different thresholds. If we want to minimize the overall probability of error, the optimum threshold level is approximately 2.5 nW, giving a probability of error of about 12%. If we set

the threshold slightly above 3 nW, then we almost never make a false recognition while correctly identifying Yong approximately 70% of the time.

We can improve the performance of the system further by using the time domain. If the input face is moving and presents different views to the system, we can eliminate many of the errors by using a period of time longer than the duration of a single frame to do the classification. Specifically, we classify the current frame to be Yong if M out of the N previous frames give us a positive response. In implementing such an algorithm, we need to select N, M and the threshold level. Shown in Fig. 6.19 is a plot of probability of error on the same three data sets as before as a function of the threshold level for M = 7 and N = 25. Notice that if the threshold level is selected in the range of 2.75 nW to 3 nW, the estimated probability of error is zero. In this example, the decision is made based on observation of the input video for 6 seconds (the computer sampled the output at 4 samples/second). In general, there is a tradeoff between performance and observation time.

The next sequence of experiments we describe were carried out to evaluate the kind of generalization obtained by the network. In this case, the subjects (Yong and others) were allowed to look at the output of the network and adjustments were made to test the limits of the system. Examples from this series of experiments are displayed in the composite of Fig. 6.20. The pictures are arranged in a 4×4 matrix. We will assign to each picture a pair of numbers (i,j) with the picture at the upper left corner being (1,1), and the one at the upper right corner being (1,4). The small black circle within each picture displays the final output of the system after thresholding. If the bright dot appears in the circle, the system

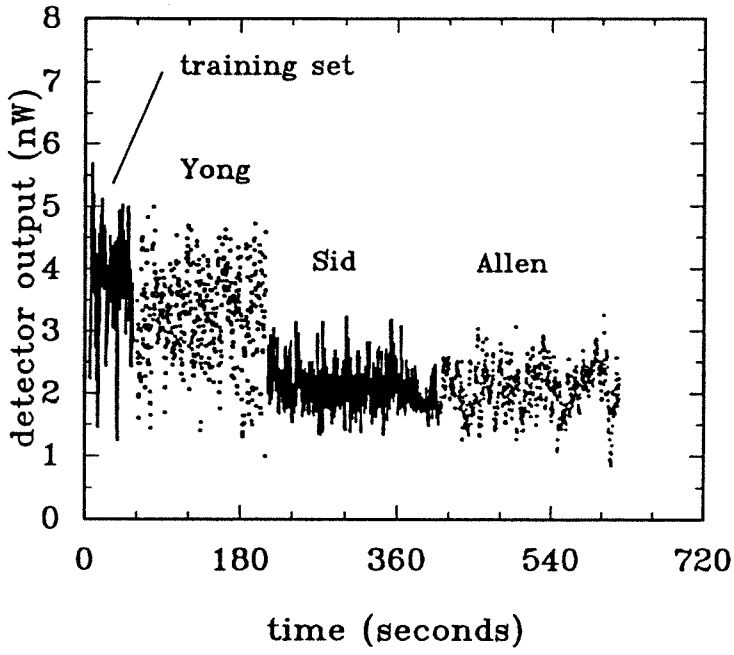


Figure 6.17. System response before thresholding.

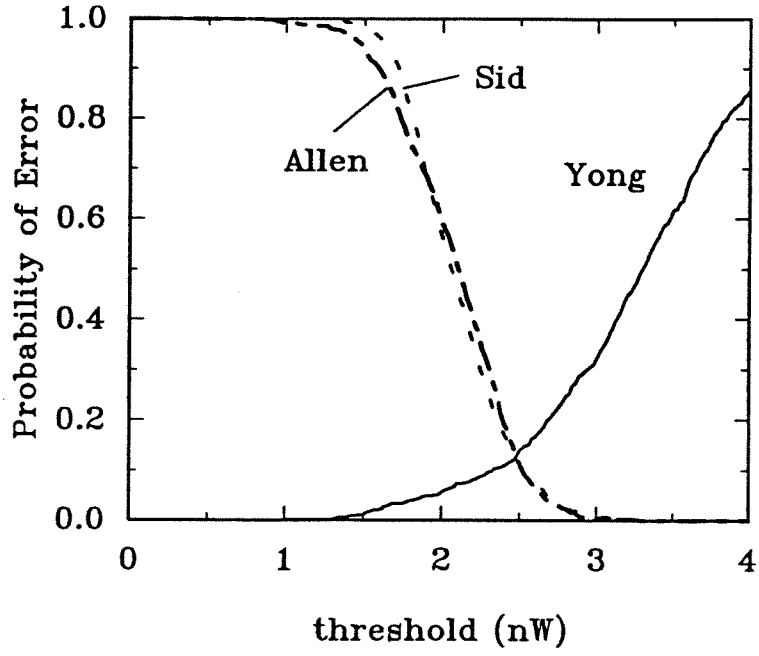


Figure 6.18. Probability of error as function of the output threshold level.

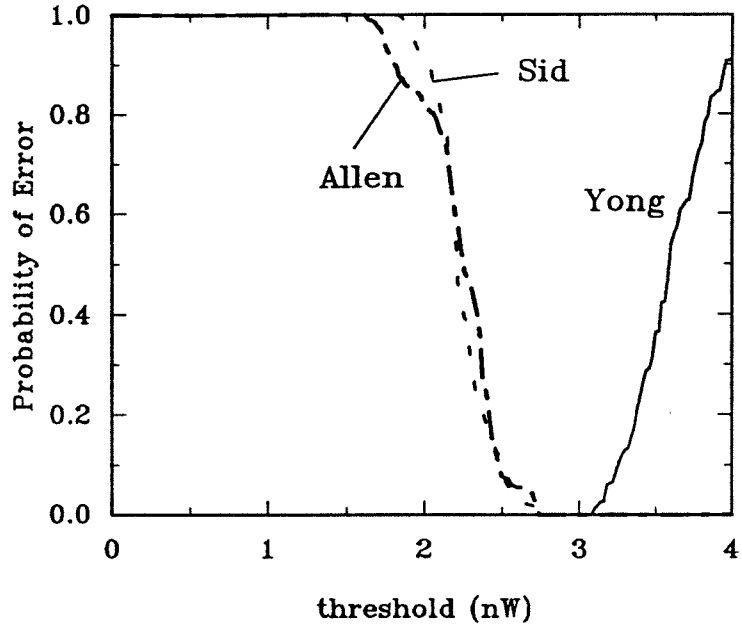


Figure 6.19. Probability of error as function of the output threshold level when the output is observed for 6 seconds to perform the classification.

makes a positive identification of Yong. Picture (1,1) is an example of Yong being correctly recognized by the system. Picture (1,2) shows Yong illuminated from below and the side, whereas during training the illumination was from above. We can see that the system is sensitive to the direction of illumination because of the edge-enhancement that is performed by the system. As the direction of illumination changes, the edges move around. To obtain invariance to illumination

direction, we need to include in the training set examples of different lighting. Picture (1,4) and (2,1) show that key features such as the mouth and the eye are crucial for recognition. However, as picture (2,2) shows, the eyes alone are not enough for a positive identification. Picture (2,3) is meant to display the invariance of the system to up and down motion. It is difficult to assess this from the still photo. However we measured a tolerance to vertical shifts of about 5% of the whole scene. The optical system was arranged such that vertical shifts of the input image become horizontal shifts on the LCTV. We did this because we need more tolerance to horizontal input shifts (people move side to side much more than up and down) and the optical system provides shift invariance in the vertical direction at the LCTV plane. Prior to the training, the tolerance to vertical input shifts was 2% of the whole scene. Training more than doubled the tolerance of the vertical shift. The tolerance of the system to nodding up and down was recorded by measuring the vertical motion on the screen of a fixed point on Yong's forehead, as he nodded up and down. According to this measure, the spot on his forehead can move by 1 cm without loss of recognition. From this measurement, and by measuring the dimensions of Yong's head, we obtain a crude estimate of 5 degrees for the maximum tolerable angle of forward head tilt. Picture (3,3) shows an example of the tolerance of the system to horizontal shifts of the input image. In this direction the optical correlator provides considerable shift invariance. We measured the maximum horizontal shift to be about 13 percent of the total horizontal extent of the input frame. Overall, the system has more than 3 times better tolerance to shifts in the horizontal than the vertical direction. Pictures (3,4) and (4,1) demonstrate the system's ability to tolerate turning of the

head, which we measured to be 30 degrees in either direction. The maximum tilt of the head (picture (4,2)) was measured to be 12 degrees in either direction. We did not seriously test the response of the system to scale changes.

6.4 Discussion and Conclusion

The power of the optical neural networks is indicated by calculating their processing speed. Take the character recognition network for an example. It has approximately 10^7 synapses or weights and the response time of the network is approximately 10 milliseconds. This corresponds to 10^9 analog multiplications per second. If the size of the input image is increased to the full resolution of the input spatial light modulator (approximately 10^5 pixels), then the rate increases to 10^{12} multiplications per second. Finally, if a ferroelectric liquid crystal SLM is used at the input and hidden layers, the response time can be improved to approximately 10 microseconds [121], which yields the very impressive rate of 10^{15} multiplications per second. This processing speed is achievable with currently available optical components and it would be extremely difficult to duplicate with electronics. The issue therefore is not whether optics can in practice outperform electronics in terms of computational speed, but rather whether the high speed of optics can be put to practical use. The main issue is whether these large optical networks can be trained effectively to solve useful practical problems. There are both algorithmic and device related problems to address. The hologram copying method we described earlier is a promising solution for the device dynamic range problems. As for the training of multilayer networks, it is well known that large networks require very large training sets [122] but it is not known how the training

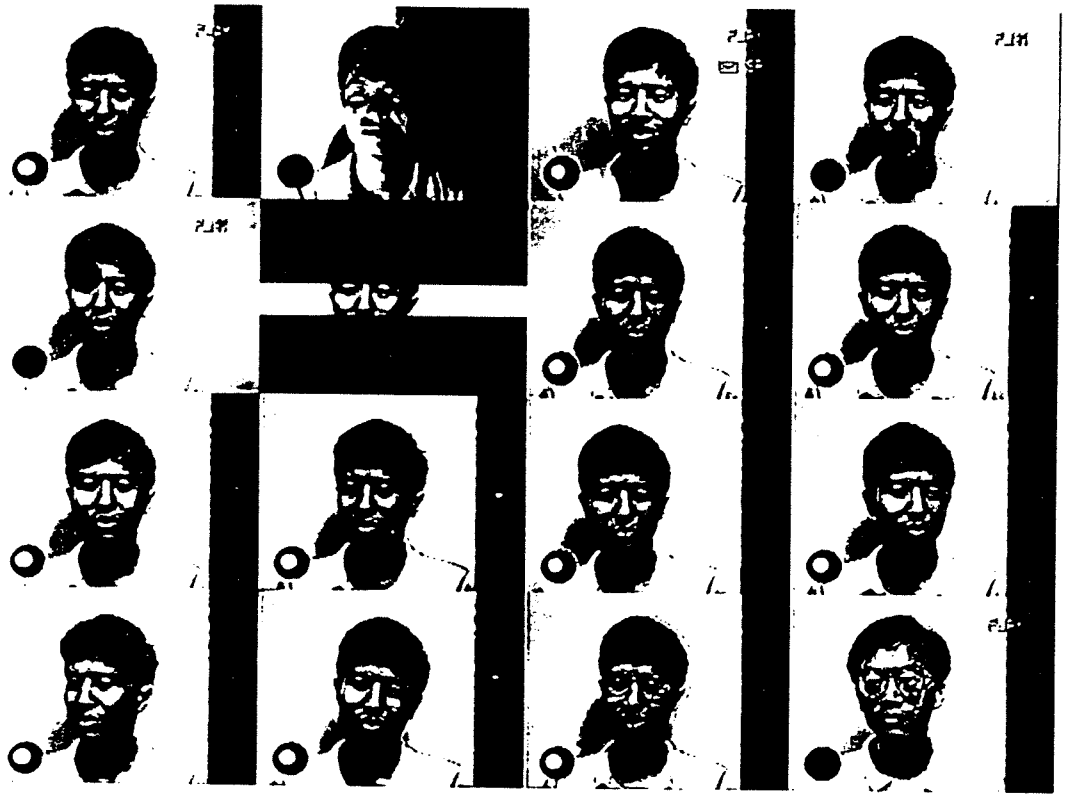


Figure 6.20. Examples demonstrating the generalization capabilities of the system. A bright dot in the circle at the lower-right corner of each photograph indicates that the system classifies the input image as the person it was trained to recognize.

time scales with the size of the network and/or the training set. Understanding such issues about the training of large optical networks remains a major challenge before these systems can have a practical impact.

REFERENCES

1. D. Gabor, "A new microscopic principle," Nature **161**, 777 (1948).

2. W. L. Bragg, "A new type of 'x-ray microscope'," Nature **143**, 678 (1939).

3. E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory," J. Opt. Soc. Am. **52**, 1123 (1962).

4. E. N. Leith and J. Upatnieks, "Wavefront reconstruction with continuous-tone objects," J. Opt. Soc. Am. **53**, 1377 (1963).

5. E. N. Leith and J. Upatnieks, "Wavefront reconstruction with diffused illumination and three-dimensional objects," J. Opt. Soc. Am. **54**, 1295 (1964).

6. Y. N. Denisyuk, "Photographic reconstruction of the optical properties of an object in its own scattered radiation field," Sov. Phys. Dokl. **7**, 543 (1962).

7. P. J. van Heerden, "Theory of optical information storage in solids," Appl. Opt. **2**, 393(1963).

8. H. Kogelnik, "Coupled wave theory for thick hologram gratings," Bell Syst. Tech. J. **48**, 2909 (1969).

9. L. D'Auria, J. P. Huignard, and E. Spitz,"Holographic read-write memory and capacity enhancement by 3-D storage," IEEE Trans. Mag. **MAG-9**, 83 (1973).

10. D. L. Staebler, W. J. Burke, W. Phillips, and J. J. Amodei, "Multiple storage and erasure of fixed holograms in Fe-doped $LiNbO_3$," Appl. Phys. Lett. **26**, 182 (1975).

11. F. H. Mok, M. C. Tackitt, and H. M. Stoll, "Storage of 500 high-resolution holograms in a $LiNbO_3$ crystal," Opt. Lett. **16**, 605 (1991).

12. F. H. Mok, "Angle-multiplexed storage of 5000 holograms in lithium niobate," Opt. Lett. **18**, 915 (1993).

13. S. Tao, D. R. Selviah, and J. E. Midwinter,"Spatioangular multiplexed storage of 750 holograms in an Fe:$LiNbO_3$ crystal," Opt. Lett. **18**, 912 (1993).

14. Y. Taketomi, J. Ford, H. Sasaki, J. Ma, Y. Fainman, and S. H. Lee, "Multimode operations of a holographic memory using orthogonal phase codes," in *Technical Digest on Photorefractive Materials, Effects, and Devices, 1991* (Optical Society of America, Washington, D.C., 1991), Vol. 14, pp. 126-129.

15. C. Denz, G. Pauliat, G. Roosen, and T. Tschudi, "Phase-coded hologram multiplexing for high capacity optical data storage," in *Technical Digest on Photorefractive Materials, Effects, and Devices, 1991* (Optical Society of America, Washington, D.C., 1991), Vol. 14, pp. 122-125.

16. S. K. Case, Appl. Opt. **18**, 1890 (1979).

17. F. T. S. Yu, S. Wu, A. W. Mayers, and S. Rajan, "Wavelength multiplexed reflection matched spatial filters using $LiNbO_3$," Opt. Comm. **81**, 343 (1991).

18. G. A. Rakuljic, V. Leyva, and A. Yariv, "Optical data storage by using orthogonal wavelength-multiplexed volume holograms," Opt. Lett. **17**, 1471 (1992).

19. T. J. Hall, R. Jaura, L. M. Connors, and P. D. Foote, "The photorefractive effect – a review," Pro. Quant. Electr. **10**, 77 (1985).

20. N. V. Kukhtarev, V. B. Markov, S. G. Odulov, M. S. Soskin, and V. L. Vinetskii, "Holographic storage in electro-optic crystals. I. Steady State," Ferroelectrics **22**, 949(1979).

21. D. Brady and D. Psaltis, "Control of volume holograms," J. OPt. Soc. Am. A **9**, 1167 (1992).

22. See, for example, J. Anderson and E. Rosenfeld, eds., *Neurocomputing*, MIT Press, Cambridge, Mass, 1988.

23. D. Hebb, *The Organization of Behavior*, Wiley, New York, 1949.

24. F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms,* Spartan Books, Washington, D.C., 1962.

25. B. Widrow and M. E. Hoff, "Adaptive switching circuits," IRE Wescon Conv. Rec. **4**, 96 (1960).

26. A. B. Vander Lugt, "Signal detection by complex spatial filtering," IEEE Trans. Inform. Theory **IT-10**, 139 (1964).

27. P. Werbos, Ph.D. Thesis, Harvard University, 1974.

28. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing, Vol. 1,* D. E. Rumelhart and J. L. McClelland, eds., MIT Press, Cambridge, Mass, 1986.

29. K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," Neural Networks **2**, 359 (1989).

30. W. P. Bleha, L. T. Lipton, E. Wiener-Avnear, J. Grinberg, P. G. Reif, D. Casasent, H. B. Brown, and B. V. Markevitch, "Application of the liquid

crystal light valve to real-time optical data processing," Opt. Eng. **17**, 371 (1978).

31. S. Fukushima *et al.*, "Bistable spatial light modulator using a ferroelectric liquid crystal," Opt. Lett. **15**, 285 (1990).
32. A. L. Lentine *et al.*, "Symmetric self-electro-optic effect device: optical set-reset latch," Appl. Phys. Lett. **52**, 1419 (1988).
33. Y. H. Lee *et al.*, "Room-temperature continuous-wave vertical-cavity single-quantum-well microlaser diodes," Electron. Lett. **25**, 1377 (1989).
34. N. Streibl et al., "Digital optics," Proc. IEEE **77**, 1954 (1989).
35. J. W. Goodman, F. J. Leonberger, S. Y. Kung, and R. A. Athale, "Optical interconnections for VLSI systems," Proc. IEEE **72**, 850 (1984).
36. D. Psaltis and N. H. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," Opt. Lett. **10**, 98 (1985).
37. N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical implementation of the Hopfield model," Appl. Opt. **24**, 1469 (1985).
38. B. H. Soffer, G. J. Dunning, Y. Owechko, and E. Marom, "Associative holographic memory with feedback using phase-conjugate mirrors," Opt. Lett. **11**, 118 (1986).
39. R. A. Athale, H. H. Szu, and C. B. Friedlander, "Optical implementation of associative memory with controlled nonlinearity in the correlation domain," Opt. Lett. **11**, 482 (1986).
40. A. Yariv and S. K. Kwong, "Associative memories based on message-bearing optical modes in phase-conjugate resonators," Opt. Lett. **11**, 186 (1986).

41. D. Z. Anderson, "Coherent optical eigenstate memory," Opt. Lett. **11**, 56 (1986).

42. Y. S. Abu-Mostafa and D. Psaltis, "Optical neural computers," Scientific American **256**, 88 (1987).

43. E. G. Paek and D. Psaltis, "Optical associative memory using Fourier transform holograms," Opt. Eng. **26**, 428 (1987).

44. C. C. Guest and R. TeKolste, "Designs and devices for optical bidirectional associative memories," Appl. Opt. **26**, 5055 (1987).

45. J. M. Kinser, H. J. Caulfield, and J. Shamir, "Design for a massive all-optical bidirectional associative memory: The big BAM," Appl. Opt. **27**, 3442 (1988).

46. L. -S. Lee, H. M. Stoll, and M. C. Tackitt, "Continuous-time optical neural network associative memory," Opt. Lett. **14**, 162 (1989).

47. E. G. Paek and A. Von Lehmen, "Holographic associative memory for word-break recognition," Opt. Lett. **14**, 205 (1989).

48. D. Psaltis, C. H. Park, and J. Hong, "Higher order associative memories and their optical implementations," Neural Networks, **1**, 149 (1988).

49. J. S. Jang, S. Y. Shin, and S. Y. Lee, "Programmable quadratic associative memory using holographic lenslet arrays, " Opt. Lett. **14**, 838 (1989).

50. L. Zhang, M. G. Robinson, and K. M. Johnson, "Optical implementation of a 2nd-order neural network," Opt. Lett. **16**, 45 (1991).

51. N. H. Farhat, "Optoelectronic analogs of self-programming neural nets: architecture and methodologies for implementing fast stochastic learning by

simulated annealing," Appl. Opt. **26**, 5093 (1987).

52. A. D. Fisher, W. L. Lippincott, and J. N. Lee, "Optical implementations of associative networks with versatile adaptive learning capabilities," Appl. Opt. **26**, 5039-5054, 1987.

53. D. Psaltis, D. Brady and K. Wagner, "Adaptive optical networks using photorefractive crystals," Appl. Opt. **27**, 1752-1759, 1988.

54. M. Ishikawa *et al.*, "Experimental studies on learning capabilities of optical associative memory," Appl. Opt. **29**, 289-295, 1990.

55. M. Oita *et al.*, "Optical implementation of large-scale neural networks using a time-division-multiplexing technique," Opt. Lett. **15**, 227 (1990).

56. Y. Qiao and D. Psaltis, "Local learning algorithm for optical neural networks," Appl. Opt. **31**, 3285 (1992).

57. E. G. Paek, J. R. Wullert, and J. S. Patel, "Holographic implementation of a learning-machine based on a multicategory perceptron algorithm," Opt. Lett. **14**, 1303-1305, 1989.

58. J. H. Hong, S. Campbell, and P. Yeh, "Optical pattern classifier with Perceptron learning," Appl. Opt. **29**, 3019-3025, 1990.

59. Y. Qiao, "Optical perceptron network uses dynamic holographic memory," Optical Processing & Computing **4**, 3 (1993).

60. C. Benkert, V. Hebler, J. -S. Jang, S. Rehman, M. Saffman, and D. Z. Anderson, "Feature extraction by a self-organizing photorefractive system," in *Technical Digest on Photorefractive Materials, Effects, and Devices, 1991* (Optical Society of America, Washington, D.C., 1991), Vol. 14, pp. 372-375.

61. K. Wagner and D. Psaltis, "Multilayer optical learning networks," Appl. Opt. **26**, 5061-5076, 1987.

62. D. Psaltis and Y. Qiao, "Adaptive multilayer optical networks," in *Progress in Optics, Vol. XXXI,* E. Wolf, ed., North-Holland, Amsterdam, 1993.

63. H. -Y. Li, Y. Qiao, and D. Psaltis, "An optical network for real time face recognition," to appear in Applied Optics.

64. X. G. Gu, *Optical neural networks using volume holograms*, Ph.D. thesis, California Institute of Technology, 1990.

65. A. Yariv and P. Yeh, *Optical Waves in Crystals*, John Wiley & Sons, New York, 1984.

66. D. Psaltis, J. Yu, X. Gu, H. Lee, "Optical neural nets implemented with volume holograms," in *2nd Topical Meeting on Optical Computing* (Optical Society of America, Washington, D.C., 1987), p. 129.

67. H. Lee, X. Gu, and D. Psaltis, "Volume holographic interconnections with maximal capacity and minimal crosstalk," J. Appl. Phys. **65**, 2191 (1989).

68. Y. Owechko, "Cascaded-grating holography for artificial neural networks," Appl. Opt. **32**, 1380 (1993).

69. J. H. Hong and R. Saxena, "Diffraction efficiency of volume holograms written by coupled beams," Opt. Lett. **16**, 180 (1991).

70. N. V. Kuktarev, "Kinetics of hologram recording and erasure in electrooptic crystals," Soc. Tech. Phys. Lett. **2**, 438 (1976).

71. D. Brady, *Photorefractive Volume Holography in Artificial Neural Networks*, Ph.D. thesis, California Institute of Technology, 1990.

72. A. M. Glass, D. von der Linde, and T. J. Negran, "High-voltage bulk photovoltaic effect and the photorefractive process in $LiNbO_3$," Appl. Phys. Lett. **25**, 233 (1974).

73. K. Blϕtekjaer, "Limitations on holographic storage capacity of photochromic and photorefractive media," Appl. Opt., **18**, 57 (1979).

74. D. Psaltis, D. Brady, X. -G. Gu, and S. Lin, "Holography in artificial neural networks," Nature **343**, 325 (1990).

75. D. Z. Anderson and D. M. Lininger, "Dynamic optical interconnects: volume holograms as optical two-port operators," Appl. Opt. **26**, 5031 (1987).

76. E. K. Gulanyan, I. R. Dorosh, V. D. Iskin, A. L. Mikaelyan, and M. A. Maiorchuk, "Nondestructive readout of holograms in iron-doped lithium niobate crystals," Sov. J. Quantum Electron., **9**, 647 (1979).

77. Y. Qiao and D. Psaltis, "Photorefractive holographic memories with dynamic copying," to appear in International Journal of Optical Computing.

78. C. -T. Chen, *Linear System Theory and Design*, Holt, Rinehart & Winston, New York, 1984.

79. D. Brady, K. Hsu and D. Psaltis, "Periodically refreshed multiply exposed photorefractive holograms," Opt. Lett. **15**, 817 (1990).

80. Y. Qiao, D. Psaltis, C. Gu, J. Hong, P. Yeh, and R. R. Neurgaonkar, "Phase-locked sustainment of photorefractive holograms using phase conjugation," J. Appl. Phys. **70**, 4646 (1991).

81. H. Sasaki, Y, Fainman, J. E. Ford, Y. Taketomi, and S. H. Lee, "Dynamic photorefractive optical memory," Opt. Lett. **16**, 1874 (1991).

82. Y. Qiao and D. Psaltis, "Sampled dynamic holographic memory," Opt. Lett. **17**, 1376 (1992).

83. S. Boj, G. Pauliat, and G. Roosen, "Dynamic holographic memory showing readout, refreshing, and updating capabilities," Opt. Lett. **17**, 438 (1992).

84. Y. Taketomi, J. E. Ford, H. Sasaki, J. Ma, Y. Fainman, and S. H. Lee, "Incremental recording for photorefractive hologram multiplexing," Opt. Lett., **16**, 1774 (1991).

85. D. L. Staebler and J. J. Amodei, "Coupled-wave analysis of holographic storage in $LiNbO_3$," J. Appl. Phys. **43**, 1042 (1972).

86. K. -Y. Hsu, H. -Y. Li, and D. Psaltis, "Holographic implementation of a fully connected neural network, " Proc. IEEE, **78**, 1637 (1990).

87. J. J. Amodei and D. L. Staebler, "Holographic pattern fixing in electro-optic crystals," Appl. Phys. Lett., **18**, 540 (1971).

88. L. Arizmendi, "Thermal fixing of holographic gratings in $Bi_{12}SiO_{20}$," J. Appl. Phys. **65**, 423 (1989).

89. G. Montemezzani and P. Gunter, "Thermal hologram fixing in pure and doped $KNbO_3$ crystals," J. Opt. Soc. Am. **B7**, 2323 (1990).

90. D. Kirillov and J. Feinberg, "Fixable complementary gratings in photorefractive $BaTiO_3$," Opt. Lett. **16**, 1520 (1991).

91. F. Micheron and G. Bismuth, "Field and time thresholds for the electrical fixation of holograms recorded in $Sr_{0.75}Ba_{0.25}Nb_2O_6$ crystals," Appl. Phys. Lett. **23**, 71 (1973).

92. F. Micheron and G. Bismuth, "Electrical control of fixation and erasure of

holographic patterns in ferroelectric materials," Appl. Phys. Lett., **20**, 79 (1972).

93. F. Micheron and J. Trotier, Ferroelectrics **8**, 441 (1974).

94. V. Leyva, A. Agranat, and A. Yariv, "Fixing of a photorefractive grating in $KTa_{1-x}Nb_xO_3$ by cooling through the ferroelectric phase transition," Opt. Lett. **16**, 554 (1991).

95. Y. Qiao, S. Orlov, D. Psaltis, R. R. Neurgaonkar, "Electrical fixing of photorefractive holograms in $Sr_{0.75}Ba_{0.25}Nb_2O_6$," Opt. Lett. **18**, 1004 (1993).

96. G. S. Trofimov and S. I. Stepanov, "Electrical development of a hologram in a $Bi_{12}SiO_{20}$ crystal," Sov. Tech. Phys. Lett. **10**, 282 (1984).

97. J. P. Wilde and L. Hesselink, "Electrical-field-controlled diffraction in photorefractive strontium barium niobate," Opt. Lett. **17**, 853 (1992).

98. S. Orlov, D. Psaltis, and R. R. Neurgaonkar, "Dynamic electronic compensation of fixed gratings in photorefractive media," submitted to Appl. Phys. Lett..

99. M. A. Neifeld and D. Psaltis, "Optical implementations of radial basis classifiers," Appl. Opt. **32**, 1370 (1993).

100. T. Grossman, R. Meir, and E. Domany, "Learning by choice of internal representations," Complex Systems **2**, 555 (1988).

101. J. J. Hopfield, "Learning algorithms and probability distributions in feed-forward and feed-back networks," Proc. Natl. Acad. Sci. USA **84**, 8429 (1987).

102. S. A. Solla, E. Levin, and M. Fleisher, "Accelerated learning in layered neural

networks," Complex Systems **2**, 625 (1988).

103. J. Moody and C. Darken, "Fast learning in networks of locally tuned processing units," Neural Computation **1**, 281 (1989).

104. T. M. Cover and P. E. Hart, "Nearest Neighbor pattern classification," IEEE Trans. Inf. Theory **IT-13**, 21 (1967).

105. T. Poggio and F. Girosi, "Networks for approximation and learning," Proc. IEEE **78**, 1481 (1990).

106. M. Mezard and J. P. Nadal, "Learning in feedforward layered networks – the tiling algorithm," J. Phys. A **22**, 2191 (1989).

107. M. Frean, "The upstart algorithm: a method for constructing and training feedforward neural networks," Neural Computation **2**, 198 (1990).

108. C. Ji, R. R. Snapp, and D. Psaltis, "Generalizing smoothness constraints from discrete samples," Neural Computation **2**, 188 (1990).

109. S. I. Gallant, "Optimal linear discriminants," IEEE Proc. 8th Conf. Pattern Recognition, Paris, 1986, pp. 849-852.

110. D. Psaltis, and Y. Qiao, "Optical neural networks," Opt. & Photonics News **1**, 17 (1990).

111. P. Kanerva, "Parallel structures in human and computer memory," in *Neural Networks for Computing,* J. S. Denker, ed., Am. Inst. Phys., New York, 1986, pp. 247-258.

112. T. Kohonen, *Self-Organization and Associative Memory*, Springer, Berlin, 1984.

113. P. A. Chou, "The capacity of the Kanerva associative memory," IEEE Trans.

Inform. Theory **IT-35**, 281 (1989).

114. R. Duda and P. Hart, *Pattern Classification and Scenes Analysis*, Wiley, New York, 1973.

115. M. Cronin-Golomb and A. Yariv, "Optical limiters using photorefractive nonlinearities," J. Appl. Phys. **57**, 4906 (1985).

116. R. Collier, C. B. Burckhardt, and L. H. Lin, *Optical Holography*, Academic Press, New York, 1971.

117. J. Yu, F. Mok, and D. Psaltis, "Capacity of optical correlators," Proc. SPIE **825**, 114 (1987).

118. C. Gu, J. Hong, and S. Cambell, "2-D shift-invariant volume holographic correlator," Opt. Comm. **88**, 309 (1992).

119. M. A. Neifeld and D. Psaltis, "Programmable image associative memory using an optical disk," submitted to Appl. Opt..

120. C. L. Giles and T. Maxwell, "Learning, invariance, and generalization in high-order neural networks," Appl. Opt. **26**, 4972 (1987).

121. G. Moddel, K. M. Johnson, W. Li, R. A. Rice, L. A. Pagano-Stauffer, and M. A. Handschy, "High-speed binary optically addressed spatial light modulator," Appl. Phys. Lett. **55**, 537 (1989).

122. Y. S. Abu-Mostafa, "The Vapnik-Chervonenkis dimension: information versus complexity in learning," Neural Computation **1**, 312 (1989).