Optical Expanders with Applications in Optical Computing

John H. Reif* Akitoshi Yoshida*

July 20, 1999

Abstract

We describe and investigate an optical system which we call an optical expander. An optical expander electrooptically expands an optical boolean pattern encoded in d bits into an optical pattern of size N bits. Each expanded pattern is one of the N mutually orthogonal boolean patterns. We wish the expansion to be exponential, so we have $d = c \log N$ for some constant c. An optical expander can be viewed as either an electrooptical line decoder which converts d bits of optically encoded binary information to up to N unique optical outputs; or a digital beam deflector which deflects an input laser beam into one of N distinct directions with a control signal of d bits. We show that an optical expander can not be constructed by using linear optical systems, so a non-linear optical filter must be used. We describe two different architectures to implement an optical expander. One uses an optical matrix-vector multiplier and an array of N threshold devices. The other uses $\log N$ novel reflection/transmission switching cells. We then analyze these architectures in terms of size, energy requirement, and speed.

Our optical expander can help develop various applications in electrooptical computing systems. In general, because of I/O constraints and the limited fan-in/fan-out of electrical circuits, the conventional VLSI technology is not suitable for building a large line decoder. On the other hand, conventional acoustooptic beam deflectors are bulky and limited by capacity-speed product. We show that as an electrooptical functional unit, the line decoder finds many applications in optical computing such as optical interconnects and optical memory. Thus, the design and development of optical expanders is vital to optical computing. Our optical expander utilizes high speed and high space-bandwidth product connections provided by optical beams in a volume, so it offers fast and accurate operations.

Holographic memory system and message routing systems are potential applications of the optical expanders, and are discussed to further motivate the design and development of optical expanders. *Key words:* Optical computing, electrooptical interconnections, holographic memory.

1 Introduction

1.1 Potential of Optical Computing

Optical Computing has recently become a very active research field. Optics has been used for image processing and long distance communications as well as for local area networks. Recently, much attention was given to the incorporation of optics into VLSI electrical circuits. Influenced by the success of optical local area network and the progress of optical emitters and detectors, there has been a large amount of attention focused on digital optical computing utilizing optics to offer global interconnections with large fan-outs. The possibilities of this approach are also contrasted to the limitation of current VLSI technology. The VLSI technology is not suitable for interconnection intensive circuits due to its two dimensionality, I/O constraints on the chip border, and electrical properties such as resistance, capacitance, and inductance. In contrast, optics can utilize free-space interconnects as well as guided wave technology, neither of which has the problems mentioned for the VLSI technology. Many researchers have been investigating suitable optical logic devices, interconnection schemes, and architectures for such optical computing systems. Surveys and introductions are found in [1] [2] [3] [4] [5].

*Department of Computer Science, Duke University, Durham, NC 27706. (919)-660-6568 (919)-660-6561 Email: reif@cs.duke.edu and ay@cs.duke.edu. Research supported in part by Air Force Contract No. AFOSR-87-0386, Office of Naval Research, Contract No. N00014-87-K0310, DARPA/ARO Contract No. DAAL03-88-K-0195, DARPA/ISTO Contract No. N00014-88-K-0458

In general, replacing electrical devices and interconnects with their functional equivalent in optics does not lead to a good result. We must find a particular area where optics can be advantageous over electronics. One such area is in high speed interconnection requiring large fan-outs among processing units, modules, or boards. Because optics can provide free interconnections in a volume without introducing various drawbacks found in the two dimensional electrical wiring, it can provide an alternative way to implement an interconnection network from a large number of source units to a large number of destination units. Another area in which optics may play an advantageous role is in holographic memory. Holographic memory may be advantageous over the conventional random access memory (RAM) in terms of storage capacity, nonvolatility, and parallel readout capability. However, in spite of their potential, practical implementation has been limited by the lack of electrooptical interface which creates a mutually orthogonal optical pattern of size N from an input of size $\log N$. We call this interface the optical expander. We discuss the optical expander and its applications in detail.

1.2 Disadvantage of VLSI implementation

Optical Expander can be viewed as an electrooptical implementation of a line decoder. The line decoder is a basic component of combinatory logic, and it is used to select one of N devices with an input control signal of size $\log N$ bits. The typical VLSI implementation of a large line decoder suffers from two problems. One is due to the topological properties of a VLSI chip. In VLSI, all wires must run on a two dimensional plane with a constant number of layers, and all I/O pads must be located on the border of the chip. As N becomes large, even though the logic gates occupies a small area, the interconnections and the I/O occupy a large area, resulting in a serious area consumption problem. The other problem is due to the small fan-in and fan-out of electrical devices. A typical implementation requires a tree structure of logic gates in $\log N$ stages. This will lead to $\log N$ time steps to generate an output. A speedup may be obtained by taking advantage of the capacitance of electrical wires. One can use a two-phase clock with a bus configuration. In phase one, N buses are charged with Vdd. Each bus has its corresponding $\log N$ pulldown gates to discharge the voltage. In phase two, the pulldown gates are enabled to discharge the buses, and the outputs are inverted to generate N outputs. However, this will still not solve the I/O constraints problem mentioned earlier. These I/O constraints may force the chip to transmit N bits in bit sequential manner. Optics may provide an alternative way to implement a large line decoder by using its free space three dimensional interconnection with large fan-in and fan-out capabilities.

1.3 Disadvantage of Beam Deflector

Analog beam deflectors based on acoustooptic effect have several drawbacks. First of all, they are bulky and acoustooptic modulators require high drive power. Secondly, they are limited by capacity-speed product.[6] The capacity-speed product of an acoustooptic modulator can be expressed as

$$N(1/t) = \triangle f \tag{1}$$

where N is the number of resolvable spatial points, $\triangle f$ is a frequency bandwidth of the acoustic wave, and t is a switching time, respectively. A frequency band width $\triangle f$ as high as 300 MHz can be obtained by the acoustooptic material such as alpha-iodic acid.[7] If we want to switch the deflector every 1μsec, then with a safety factor of 2, the number of resolvable points will be at most 150. In order to overcome the disadvantage of the acoustooptic beam deflector, a multistage digital beam deflector was designed.[8] They demonstrated a 20-stage deflector consisting of a series of nitrobenzene Kerr cells and birefringent calcite prisms. The laser beam was deflected into a two dimensional 1024×1024 plane in every 2μsec. This approach provided a great flexibility and accuracy in controlling the deflection angle. However, it required very high bias and switching voltage of several kilovolts, and also resulted in a large power consumption.

1.4 Description of Optical Expanders

An optical expander takes as an input a boolean pattern of size $d = c \log N$ bits, and expands it to a boolean pattern of size N bits, where c is a constant satisfying $1 \leq c \leq 2$. Each expanded boolean pattern must be mutually orthogonal to the others. Thus, the optical expander can be viewed as either of the following: an electrooptical line decoder which converts d bits of optically encoded binary information to up to N unique optical outputs, or a digital beam deflector which uses a control signal encoded in d bits to deflect an input laser beam into one of N directions.

More precisely, an optical expander takes as input one of N distinct boolean vectors $p_0, p_1, \ldots, p_{N-1}$ of length d. We call these vectors the input patterns. Each input pattern is optically encoded by using d pixels, each pixel being either ON (denoted by 1) or OFF (denoted by 0). We will require that each input pattern has exactly $d/2$ pixels ON. The optical expander produces a spatial output pattern r_i from given input pattern p_i. Each output pattern r_i is one of N mutually orthogonal boolean vectors of length N. In addition to our standard optical expander, we define a generalized optical expander. The generalized optical expander is similar to the standard optical expander except for one detail. Unlike the standard optical expander, each expanded boolean pattern may have more than one ON in its elements. In other words, the generalized optical expander creates a boolean pattern of size N which is a bitwise OR product of some subset of the N mutually orthogonal boolean patterns. The advantage of the generalized optical expander becomes clear in certain applications. It can be used for broadcasting messages in a message switching network. It can also be applied to a holographic memory system with a multiple readout capability, where the bitwise OR, AND, or XOR products of several images (data) can be directly obtained as a superimposed output on the detector array.

Our optical expander accepts an input pattern encoded in d bits, and expands it into a pattern encoded in N bits. We wish to have an exponential expansion, so d has to be represented by $d = c \log N$ for some constant c. First, we describe an optical expander with the constant $c = 2$. Later, we will look at an encoding scheme with the constant $c \approx 1$ for a large d. (Table 1) This allows us to produce a greater number of orthogonal patterns with the same number of input bits. However, setting $c = 2$ offers several advantages. First of all, it makes the coding scheme simple, since $d = 2 \log N$ offers a coding scheme where each p_i can be a concatenation of two binary strings: one representing i in binary format, and the other representing i in one's complement binary format. Thus, p_i can be easily produced from the binary-coded output from the electrical interface without any additional electrical mapping interfaces. Second, it also makes the optical interconnection patterns from d optical inputs to the threshold array regular, thus resulting in a simple implementation. Finally, it can provide an addressing scheme for the generalized optical expander. We will discuss this encoding issue later in detail.

1.5 Optical Expanders require Non-linear optical systems

A linear optical system can not be used as an optical expander, since any linear mapping from an input of size d creates no more than d linear independent output patterns. Thus, it is impossible to create a set of $N > d$ mutually orthogonal patterns by any linear optical system on d linear independent patterns.

1.5.1 Non Linear Optical Filters

Non-linearity can be introduced to an optical system by two methods. One can use a non-linear device. Thresholding the input intensity at a certain level to produce output is a non-linear operation. It can be implemented by optical non-linear devices such as optical logic etalon (OLE)[9] [10], or by electrooptical non-linear devices such as self electrooptic effect device (SEED)[11]. The other method is to translate an input into a spatial pattern, and then to apply a linear filter at the fourier plane. An example is Theta modulation, where data are encoded as a grating of different orientations. In our optical expanders, we use non-linear devices.

1.5.2 Outline of Paper

In section (2), we describe applications of our optical expanders. In particular, section (2.1) describes optical interconnects in general. Sections (2.2) and (2.3) discuss holographic memory and message routing, respectively. We start our discussion on the optical expanders in section (3). In section (3.1), we describe our first optical expander. It consists of two parts: a linear part and a non-linear part. The linear part is a matrix-vector multiplier, and the non-linear part is an array of thresholding devices. In section (3.2), we describe our second optical expander. It consists of $\log N$ identical switching cells. Finally, section (4) concludes the paper.

2 Applications of Optical Expanders

Before we discuss specific applications, we look at the general issues. There are various applications which can be efficiently implemented by use of optical expanders. Systems which require N distinct entry beams either to an N-superimposed hologram or to an array of N devices may use an optical expander to generate N beams from optical input of $c \log N$ bits. Without our optical expander, such systems require either a beam deflector to deflect a laser beam into one of N unique directions [12], or an electrically implemented line decoder which accepts $\log N$ bits of binary information and creates one of the N mutually orthogonal beams.

The conventional acouso- or electro-optical beam deflectors have several drawbacks described earlier. Electrically implemented large line decoders are not practical in terms of speed and wiring areas for a large N. As we mentioned earlier, the I/O constraints limit the size of system which can be practically implemented.

Our optical expander will provide an advantage to these devices by utilizing free space with flexibility and accuracy provided by digital operations. The following sections describe typical applications.

2.1 Optical Interconnects

Because of the availability of non-linear electrical devices as gates which are extensively used in the interconnection network, electrically implemented interconnections are widely seen among many computer organizations.[13] [14] However, the future of electric interconnections is not necessarily bright. The problem comes from its restricted two dimensionality and RC delay on interconnections.[15] This implies that with the current technology of electrical interconnects, the interconnection area will soon occupy a large portion of a chip and the interconnection delay will become a bottleneck in processing.

These drawbacks do not exist in optical interconnections. Light beams need not be confined in a wave guide such as an optical fiber, but can travel freely through space. Light beams can provide a great bandwidth, and the propagation of light traveling through space or in a fiber is not affected by resistance, capacitance, or inductance. Thus, optical interconnections may offer a high data transfer rate in a simple architecture by a set of light beams freely traveling through space. A large number of papers discuss the potential of optical interconnections. Surveys are found in [16] [2] [17].

Several theoretical studies have been made to investigate the advantage of free space optical interconnects.[18] [19] [20] The studies indicate that optical interconnects have an advantage over their electrical counterparts in terms of area (volume), speed, and power consumption for high speed communications except for the shortest distance within chips. Furthermore, they suggest that optical interconnects may be advantageous in large area VLSI circuits which require high data rates and/or large fan-out. These results as well as work on various implementations of optical interconnects, lead to our interest in designing our optical expander which can be efficiently used for a holographic memory storage and a message routing network.

Finally, from the purely theoretical computational point of view, for a given problem, there is a lower bound on the circuit area and its computational time. One such lower bound on VLSI model called "AT^2 bounds" states that $AT^2 = \Omega(I^2)$, where A is the circuit area, T is the time used by the circuit, and I is information content[1] of the problem.(See [21]) In a three dimensional electrooptical model described by Barakat and Reif called VLSIO, the similar lower bound can be expressed as $VT^{3/2} = \Omega(I^{3/2})$.[22] This implies that as the information content becomes larger, the VLSI circuit requires a larger and larger area to solve the problem in a fixed amount of time. Using optical interconnection as in VLSIO model overcomes this interconnection problem by utilizing space in a volume.

2.2 Holographic Memory Storage

Holograms can be used to implement random access memory storage systems. [23] [24] [25] [26] [27] The basic idea of holographic memory storage is that the data are arranged in blocks which are stored in holograms. A large block of memory can be retrieved at a any given time by using its corresponding reconstruction beam. This type of memory is particularly suited for read-only applications, since the holograms can be fixed. However, dynamically modifiable holograms such as photorefractive materials may be useful for active holographic memory storage systems. The work in the 70s promised the advantage of holographic memory over other types of memory in terms of bit/volume ratio, size, and throughput. However, the lack of appropriate recording materials and fast addressing methods kept holographic memory behind the progress of conventional bipolar or MOS based memory.

Recently, the advance in recording materials such as various photocrystals and the success in fabricating an array of large number of micro lasers have provided a chance for holographic memory to be efficiently implemented. Several prototypes of such a memory storage system have been developed at Microelectronics and Computer Technology [28] and Bellcore [29].

[1] Information content is the number of bits that must cross a boundary in order to solve the problem. The boundary separates the circuit into two sides, each of which holds approximately half the input bits.

In a typical holographic memory storage system, the data are organized in blocks.(Fig. 1) Each block is a two dimensional bit image consisting of $L \times L$ pixels (total of L^2 pixels per block) The detection of a bit can be made by using a PIN photodetector. N blocks of images are stored in either a single multiple-exposure volume hologram or an array of $N \times N$ holograms. If a multiple-exposure volume hologram is used as a storage medium, we need N mutually orthogonal patterns to retrieve each block. In other words, in order to read out each block, we need a total of N beams such that each beam has a distinct incident angle to the holographic medium. When one of these beams illuminates the holographic medium, the corresponding stored image is formed on the detector array. The detector array can be an array of $L \times L$ PIN photodetectors which convert optical signals to electrical signals. If instead an array of thin holograms is used, each hologram can be separately illuminated by a distinct beam. Again, N mutually orthogonal patterns are necessary to retrieve N blocks of image.

Using a laser beam with wavelength λ, the number of bits which can be stored in a unit volume is proportional to $1/\lambda^3$, whereas in a unit area it is proportional to $1/\lambda^2$. Therefore, the first approach can achieve a large bit/volume ratio. However, the diffraction efficiency of each image significantly decreases as the number of stored images increases. The second approach is limited by the maximum diffraction angle at which the image is formed on the detector array.[25] As the array becomes large, it becomes difficult for the hologram at the corner to form its image on the detector array. This may be improved by having several detector windows on the back of the hologram array, and merging these windows together into the detector array with some integrated optics. In any case, we must create N mutually orthogonal addressing patterns to address N images. These patterns can be electrooptically created by our optical expander with input size $d = c \log N$, where $1 \leq c \leq 2$. Thus, the electrical interface to the holographic memory system needs only a pattern encoded in d pixels: this pattern is electrooptically expanded into a pattern of N pixels to retrieve one of N blocks of images. Without our optical expanders, an alternative approach requires a set of N orthogonal addressing patterns which must be either electrically created by using a VLSI line decoder, or acousto- or electrooptically created by using a beam deflector. As we mentioned earlier, these previous approaches have disadvantages. Thus, our optical expanders offer an advantageous approach.

2.3 Message Routing

Interconnection networks in parallel processing computers are very important subjects. There are many interconnection networks for different applications, since different algorithms require different degree of globality of the interconnects.

Message routing is a task where messages are to be moved among the various processing units. We assume there are N processors and N messages, where each processor has a distinct message with a distinct destination address. Then, simultaneously, each message is routed from its originating processor to its destination processor.

In parallel computing there are several classes of problems. Some problems require no communications or only fixed local communications among processing units; examples are image processing and various matrix operations using systolic algorithms. Some problems require very intensive and sometimes dynamic communications among distant processing units; a fourier transform requires communications among distant processing units, but can be implemented by a fixed connection; problems often found in forecasting and AI, on the other hand, require global and dynamic interconnections.

In a general purpose parallel computer, the full advantage of parallelism will only be realized if each processing unit has a direct communication path to every other processing unit. In such a condition, each processing unit can process its data without having serious communication delays, thus resulting in high overall throughput rates. If this condition is not satisfied, the communication cycle may far exceed the processing cycle, and this will cause a serious bottleneck in the overall system speed. For example, the Connection Machine is a 65,536 processor single instruction multiple data stream (SIMD) computer developed at Thinking Machines [30]. Its data transfer time is at least 1000 times slower than the processor instruction step, introducing a serious bottleneck. Thus, we recognize that message routing is a very crucial problem in designing efficient parallel computers.

Among various message routing networks the highest level of interconnection is a crossbar network which uses N^2 interconnects to connect N source units and N destination units. The number of electrical interconnection wires required by each processing unit to communicate with the other processing unit on- and off-board will limit the feasible size of the network. The property of light beams which we briefly mentioned above may give a great potential for an alternative high-speed optical crossbar type of networks.

There are several optical interconnection networks which have already been proposed. One is optical crossbar network. [31] [32] [33] [34] The optical crossbar network typically uses an $N \times N$ spatial light modulator (SLM) to connect N source processors to N destination processors. Each source processor uses a column of the $N \times N$ SLM to address one of N distinct destination processors. The advantage of this optical crossbar is that once all the entries of the $N \times N$ SLM are set, the message can be transmitted at a very high data rate, namely at an optical pulse modulation rate. (Fig. 2) This matrix-vector multiplier based crossbar network has two drawbacks. One is that at most $1/N$ of the power incident on the SLM will reach the detector. The other is that it takes a long time to electrically set an $N \times N$ SLM.

Another network uses multiple stage optical switching networks. [31] [35] [36] [37] This is an optical implementation of various fixed multiple stage electrical interconnection networks. The multiple stage networks typically require the setting of only $O(N \log N)$ switches to connect N sources to N destinations. However, they require $O(\log N)$ steps to route a message and some topologies do not guarantee a non-blocking routing.

To overcome the drawbacks of previously proposed systems, a network which used fixed multiple-exposure holograms to connect N source units to N destination units was presented.[38] [39] (Fig. 3) Unlike the matrix-vector multiplier based optical crossbar networks, this holographic interconnection network uses fixed holograms to steer spatially encoded light beams transmitted from the source units toward the destination units.

The basic idea of steering each light beam to its destination is to use holographic associative matching. Each processor has a fixed hologram which implements connections to other processors. Each processor can establish its connection to other processors by illuminating its hologram by a reference beam. They demonstrated a 4 processor system with two different spatial encoding schemes. The first scheme used a set of N mutually orthogonal patterns to encode destination addresses. In this scheme, the system crosstalk caused by a false matching was minimized. However, as with the traditional crossbar networks, an $N \times N$ SLM was required. To gain an advantage over the traditional matrix-vector multiplier based crossbar network, the second scheme employed a shorter address to reduce the size of SLM, which further led to a shorter reconfiguration speed. In this scheme, each address was encoded as a non-orthogonal pattern of size $\sqrt{2N}$ bits. The system crosstalk was distributed among different detectors by increasing the dimensionality of the detector plane. As we see, if we can electrooptically create N mutually orthogonal patterns, an efficient crossbar network can be implemented by using them as addressing patterns to the holograms. Our optical expanders can create such N mutually orthogonal optical patterns from input optical patterns of size d.

3 Optical Expanders

The optical expander is a non-linear electrooptical filter which creates a large number N of mutually orthogonal patterns. It uses electorooptical devices with at most d boolean input bits. Here d is no greater than $2 \log N$. There may be various ways to implement an optical expander. Our first approach uses matrix-vector multiplication followed by threshold operation. We call this the Matrix-Vector Multiplier (MVM) Optical Expander. Our second approach is based on an novel idea to use only $\log N$ identical switches to digitally deflect the laser beam. We call this the Digital Beam Deflector (DBD) Optical Expander. The following sections first describe the MVM Optical Expander, and then the DBD Optical Expander.

3.1 MVM Optical Expander

First, we describe two different encoding schemes which can be used in this model.

3.1.1 Encoding of Input

Our optical expander uses an optical matrix-vector multiplier. The vector represents one of N distinct input patterns $p_0, p_1, \ldots, p_{N-1}$ of length d. Each pattern has exactly $d/2$ ONs (denoted by 1) and OFFs (denoted by 0), which keep the total power of each pattern equal. A set of such patterns can be easily obtained.

We present two encoding schemes, each of which has an advantage and a disadvantage. The first scheme uses $d = 2 \log N$ bits. Each pattern p_i consists of two bit strings: one encoding i in binary and the other encoding i in one's complement. This encoding, which is called the dual rail coding [40], is often used in optical computing [41]. If this is considered too large, we can set d smaller than this value. To do this, we can recursively enumerate a set of bit strings which has $d/2$ 1s and 0s. We show that this gives $d \approx \log N$ for large d. This can significantly reduce the number of bits required to produce a large N. The analysis of this encoding can be done by using Stirling's formula $x! \approx \sqrt{2x} x^x e^{-x}$. To find the asymptotic value of c as d becomes large, we have;

$$2^{d/c} = \binom{d}{d/2} = \frac{d!}{(d/2!)^2} \approx \frac{\sqrt{2\pi d} d^d e^{-d}}{2\pi(d/2)(d/2)^d e^{-d}} = \sqrt{\frac{2}{\pi d}} 2^d \tag{2}$$

Thus, c can be written as

$$c \approx \frac{d}{d + \frac{1}{2}\log\frac{2}{\pi d}} \tag{3}$$

¿From this equation, we see that as d approaches infinity, c approaches 1. This encoding produces a significantly larger N as d increases. For example, $d = 14$ yields $N = 3432$ in this encoding, whereas in the dual rail encoding, we have $2^7 = 128$. The advantage of the dual rail encoding comes into play when it is used for the generalized optical expander. We will explain this in the next section.

3.1.2 Matrix-Vector Multiplication

We describe the general idea of how to create an orthogonal boolean pattern of size N from a given input of size d. We use matrix-vector multiplication as follows. The matrix which is of size $N \times d$ consists of N rows, where the i-th row is p_i. We can consider this multiplication as a boolean matrix-vector multiplication by using optical devices. Let p_k be an input to the matrix-vector multiplier. Then, the result of the multiplication is an output vector of length N, where the i-th element of the vector is the inner product of p_k and p_i.

$$\begin{pmatrix} p_0 \\ p_1 \\ \cdots \\ p_{k'} \\ \cdots \\ p_k \\ \cdots \\ p_{N-1} \end{pmatrix} (p_k)^T = \begin{pmatrix} . \\ . \\ . \\ 0 \\ . \\ d/2 \\ . \\ . \end{pmatrix} \tag{4}$$

Here, $(p_k)^T$ represents the transposed vector of pattern p_k.

The output vector has value $d/2$ at the k-th position and has a 0 at the k'-th position, where $p_{k'}$ is the bit wise complement of p_k. We apply a threshold operation at a certain intensity level to produce one of N mutually orthogonal patterns of length N.

In the generalized optical expander, input vector p_k may have more than $d/2$ 1s. Thus, the output after the threshold operations may have more than one 1. A control signal of size $d < N$ bits can not set arbitrary positions to 1. This is obvious, since the number of such distinct patterns is 2^N, thus requiring N bits to choose one of them. If the encoding with $c = 2$ is used, it is easy to set several positions of the output to 1s. We can consider each element of the output to be a leaf of a complete binary tree of hight $d/2$. In the dual rail encoding, each bit in the input has its corresponding complement. Setting a particular bit position to 1 and its complement to 0 corresponds to traversing a node to its right child. Similarly, setting a paricular bit position to 0 and its complement to 1 corresponds to traversing a node to its left child. Thus, only one leaf can be finally reached. In order to reach several leaves, both children must be traversed. Setting both bits to 1s corresponds to traversing both children. In this way, the output can have several 1s.

The following section describes optical matrix-vector multipliers and threshold operations in more detail.

3.1.3 Optical Matrix-Vector Multiplier

In this section first we review the traditional matrix-vector multiplier. After that, we describe the matrix-vector multiplier which we use in our optical expander.

Let X be a vector of length N, and let A be a matrix of size $N \times N$. The optical implementation of a parallel, real-time matrix-vector multiplier is shown in (Fig. 4). [17]

The input vector is represented by an array of N light emitting diodes (LEDs), where the light intensity of each LED represents the value of each element in the vector. The matrix is represented by a transparency matrix mask of size $N \times N$. The transmittance of the (i, j) entry of the mask is proportional to the value of $a_{i,j}$. The light from each LED is spread onto the corresponding column in the transparency mask. Thus, the intensity of the light passing the (i, j) entry is proportional to $x_j a_{i,j}$. The light from a whole row is collected, and focused onto the corresponding position of the output vector Y. The intensity of the light at the i-th position of the output vector is then proportional to the i-th element of the product. For the optical expanders we need to multiply a matrix A of size $N \times d$ with a vector X of length d. We can arrange the input vector, the matrix mask, and the output vector into nearly square shapes. This means the vector of length d is formatted in $\sqrt{d} \times \sqrt{d}$, and the matrix of size $d \times N$ is formatted in $\sqrt{dJN} \times \sqrt{dJN}$. The output vector of length N is formatted in $\sqrt{N} \times \sqrt{N}$. The examples are shown in figure 5.

To clarify the idea of this multiplier, we describe an example with size $d = 4$, $N = 6$. The input vector of length d is represented by a $\sqrt{d} \times \sqrt{d}$ either LED or laser diode (LD) array in row major order. If d is a square of some integer, the array can be square. Otherwise, the array may not be square. The matrix mask is divided into d blocks of equal size. These blocks are arranged in $\sqrt{d} \times \sqrt{d}$ matrix form. The i-th column of the matrix is represented by the i-th block in row major order. Each block has N elements where the transparency of each element represents an element of the column. Again, if N is a square of some integer, the shape of each block becomes square. The elements of each column are arranged in row major order in its corresponding block.

In operation, plane waves are produced from the diode array by lenses. The plane wave from the i-th diode illuminates the i-th block which encodes the i-th column. Thus, the plane wave passing the i-th block represents the vector product of the i-th column of the matrix and the i-th element of the vector. All these plane waves are superimposed at the detector array, and thus the product is obtained. The output vector is obtained in $\sqrt{N} \times \sqrt{N}$ matrix form, where the elements are organized in row major order. We use this idea to construct a multiplier for our optical expander.

We represent the input vector X by an LD array of size d formatted in a square shape. The output vector Y of size N can also be formatted in a square shape. An array of lenses is placed in front of the LD array so that a set of plane waves are obtained by the lenses. Each plain wave then illuminates its corresponding matrix mask. Again, an array of lenses with a focal length f_1 is placed distance f_1 away from the matrix masks. This makes a set of fourier transform of the matrix mask patterns to be produced at the back focal plane of the lens array. A lens with a focal length f_2 is placed distance f_2 away from the fourier plane, so that the matrix mask patterns will be superimposed on the back focal plane of this lens. We note that there will be a phase shift for each mask pattern at the superimposed image. The phase shift is proportional to the lateral shift of the mask pattern from the optical axis of the second lens. It can be ignored when the intensity is processed at the image.

This approach seems to require complex matrix mask construction. However, if the detectors are formatted in a two dimensional plane with encoding scheme using $c = 2$, it will become very regular. This is shown in figure 6. The interconnection patterns from an LD to the threshold array can be grouped into $d/4$ different patterns. Then, each interconnection pattern is one of the $d/4$ patterns with one of four different orientations. As we see from the figure, these basic $d/4$ patterns are regular. Figure 6 compares an example of the two encoding schemes.

3.1.4 Threshold Operation

The purpose of performing a threshold operation is to introduce a non-linearity. A non-linearity is necessary to generate a set of N mutually orthogonal boolean patterns from a set of N distinct d-linear independent boolean patterns. Here, we have $d \leq 2 \log N$. As we recall, the output vector from the matrix-vector multiplier has a value $d/2$ at the k-th position and has value 0 at the k'-th position, where $p_{k'}$ is the bitwise complement of input pattern p_i.

One possibility to make the output vector orthogonal is to set the threshold value at intensity $I = I_1 d/2$, where I_1 is a factor which corresponds the intensity of a one. The thresholded output becomes *High*, if the intensity is $I \geq I_1 d/2$, and becomes *Low* otherwise. This will produce one of the mutually orthogonal patterns, since there is exactly one position in the output vector which has intensity $I = I_1 d/2$. This method will work as long as d is small. However, when d becomes relatively large, there is a practical problem. The problem is caused by the physical limit of how finely we can threshold the intensity. When a threshold operation at intensity $I = I_1 d/2$ is performed, intensity $I = I_1 d/2$, and intensity $I \leq I_1(d/2 - 1)$ must be distinguished. This will become difficult when d becomes large. Our solution to this problem is the following. Since the output vector also has exactly one position which has intensity $I = 0$, we threshold intensity at $I = 0$. This means we must distinguish intensity $I = 0$ from intensity $I \geq I_1$. Then, the complement of the thresholded output becomes orthogonal. In this approach we need not concern the physical limit of the thresholding device mentioned above, but rather the sensitivity and dynamic range of the device.

However, in this approach, two types of intensity noises must still be considered. One is caused by the voltage fluctuation of the driving circuit and the other is a background intensity caused by diffraction from different channels. These noises must be kept low enough so that $I = 0$ and $I \geq I_1$ can be distinguished by the detectors. This can be improved by using differential devices at the threshold devices. The generic model of a differential device consists of two PIN photodiodes connected electrically in series.(Fig. 7)

The logical state of the device is defined by the ratio of the intensity of two light beams falling on the two photodiodes. In our threshold device, one photodiode is used to detect the light beam from the LD array, and the other is used to cancel the noise by biasing the voltage drop caused by the first photodiode.

Several optical or electrooptical switches have been designed and demonstrated. The optical logic etalon (OLE) [9] [10] and the self electrooptic effect device (SEED) [11] are the most common devices. As for a differential device, the Symmetric SEED (S-SEED) has been demonstrated [42] and widely used as a basic component.[43] [44] [45] Both the OLE and the SEED have a contrast ratio of approximately 5:1.

Recently, low-threshold electrically pumped vertical-cavity surface-emitting microlaser diode arrays (SELDA's) have been developed [29]. Each laser can be as small as a few micrometers. The output light from each laser has high contrast, since an OFF state produces no light.

The optical output ports such as laser diodes or modulators are based on GaAs technology. The input ports such as photodiode detectors are based on Si technology. This makes it difficult to fabricate laser diodes and photodetectors on the same single substrate. Several methods of fabricating hybrid systems have been investigated [46] [47] [48], and may lead to a large array of optical threshold devices with a very high contrast.

3.1.5 Overall Architecture and Analysis of MVM Optical Expander

The overall architecture is shown in figure 8. The output beam can be used to either address a hologram, or send a message to a destination.

We now analyze three issues concerning the MVM Optical Expander. The first issue is the diffraction limit imposed by a finite size of the lenses. The second concerns the fan-out factor which depends on the minimum detectable power of the detectors at a given operating condition. The third concerns the dynamic range of the detectors.

Diffraction Limit

Here, we consider the diffraction limit of the system. Figure 9 shows a side view of the system. We examine the image formation of a arbitrary matrix mask on the detector plane.

In this figure, each mask has a dimension of $M \times M$. The first lens L_1 has an aperture function $a_1(x, y)$ and has a focal length of f_1. The second lens L_2 has an aperture function $a_2(x, y)$ and has a focal length f_2. Note there is a lateral shift of the mask from the optical axis of the second lens L_2. Let s_x and s_y represent this shift along the x-axis and the y-axis, respectively. For our analysis, we use rectangular apertures for both lenses. Thus, we have

$$a_1(x, y) = \Pi\left(\frac{x}{D_1}\right)\Pi\left(\frac{y}{D_1}\right) \tag{5}$$

$$a_2(x, y) = \Pi\left(\frac{x}{D_2}\right)\Pi\left(\frac{y}{D_2}\right) \tag{6}$$

Here, Π denotes a rectangular function, D_1 and D_2 are the size of each aperture.

We assume that a plain wave is incident on the mask. We denote the mask pattern by $g(x, y)$. Then, the light distribution $E_{z_1}(x, y)$ just in front of the lens L_1 is given as [49]

$$E_{z_1}(x, y) = [g(x, y) \otimes h_{f_1}(x, y)]a_1(x, y) \tag{7}$$

Here, $\otimes$ represents a convolution operator and $h_{f_1}(x, y)$ is a point-spread function of a distance f_1. This point-spread function is the light distribution of a point source at distance $z = f_1$ away. Thus, we have

$$h_z(x, y) = \frac{1}{j\lambda z}\exp\left[jk\left(z + \frac{x^2 + y^2}{2z}\right)\right] \tag{8}$$

Using the Fresnel approximation, we can write the light distribution at the back focal plane of the lens L_1 as

$$E_{z_2}(x, y) = \frac{1}{j\lambda f_1}\exp\left[jk\left(f_1 + \frac{x^2 + y^2}{2f_1}\right)\right]\mathcal{F}\{E_{z_1}(x, y)\} \tag{9}$$

Here, we have the spatial frequencies $\nu_x = x/\lambda f_1$ and $\nu_y = y/\lambda f_1$.
We use the following relations to expand the fourier transform term.

$$\mathcal{F}\{h_z(x,y)\} = H_z(\nu_x, \nu_y) = \exp\left[jk\left(z - \frac{z\lambda^2(\nu_x^2+\nu_y^2)}{2z}\right)\right] \tag{10}$$

Thus, we have

$$\begin{aligned} E_{z_2}(x,y) &= \frac{1}{j\lambda f_1}\exp\left[jk\left(f_1 + \frac{x^2+y^2}{2f_1}\right)\right] G\left(\frac{x}{\lambda f_1}, \frac{y}{\lambda f_1}\right) H_{f_1}\left(\frac{x}{\lambda f_1}, \frac{y}{\lambda f_1}\right) \otimes A_1\left(\frac{x}{\lambda f_1}, \frac{y}{\lambda f_1}\right) \\ &= \frac{1}{j\lambda f_1}\exp[j2\pi k f_1] G\left(\frac{x}{\lambda f_1}, \frac{y}{\lambda f_1}\right) \otimes A_1\left(\frac{x}{\lambda f_1}, \frac{y}{\lambda f_1}\right) \end{aligned} \tag{11}$$

Again, the light distribution just in front of the second lens L_2 can be written by using a point-spread function as

$$E_{z_3}(x,y) = [E_{z_2}(x - s_x, y - s_y) \otimes h_{f_2}(x,y)]a_2(x,y) \tag{12}$$

Here, note that we shift the optical axis by s_x and s_y.
At the detector plane,

$$E_{z_4}(x,y) = \frac{1}{j\lambda f_2}\exp\left[jk\left(f_2 + \frac{x^2+y^2}{2f_2}\right)\right]\mathcal{F}\{E_{z_3}(x,y)\} \tag{13}$$

Here, the spatial frequencies $\nu_x = x/\lambda f_2$ and $\nu_y = y/\lambda f_2$.
Expanding the fourier transformation term yields the following.

$$\begin{aligned} E_{z_4}(x,y) &= \frac{1}{j\lambda f_2}\exp\left[jk\left(f_2 + \frac{x^2+y^2}{2f_2}\right)\right]\mathcal{F}\{E_{z_2}(x - s_x, y - s_y)\}H_{f_2}\left(\frac{x}{\lambda f_2}, \frac{y}{\lambda f_2}\right) \otimes A_2\left(\frac{x}{\lambda f_2}, \frac{y}{\lambda f_2}\right) \\ &= -\frac{f_1}{f_2}\exp[j2k(f_1+f_2)]\exp\left[j2\pi\left(\frac{s_x x}{\lambda f_2} + \frac{s_y y}{\lambda f_2}\right)\right] \times \\ &\quad g\left(-\frac{f_1}{f_2}x, -\frac{f_1}{f_2}y\right) a_1\left(-\frac{f_1}{f_2}x, -\frac{f_1}{f_2}y\right) \otimes A_2\left(\frac{x}{\lambda f_2}, \frac{y}{\lambda f_2}\right) \end{aligned} \tag{14}$$

Since we have a rectangular aperture function. we have

$$A_2(\nu_x, \nu_y) = D_2^2\mathrm{sinc}(D_2\nu_x)\mathrm{sinc}(D_2\nu_y) \tag{15}$$

$$\begin{aligned} E_{z_4}(x,y) &= -\frac{f_1}{f_2}\exp[j2\pi k(f_1+f_2)]\exp\left[j2\left(\frac{s_x x}{\lambda f_2} + \frac{s_y y}{\lambda f_2}\right)\right] \times \\ &\quad g\left(-\frac{f_1}{f_2}x, -\frac{f_1}{f_2}y\right)\Pi\left(-\frac{f_1 x}{f_2 D_1}\right)\Pi\left(-\frac{f_1 y}{f_2 D_1}\right) \otimes D_2^2\mathrm{sinc}\left(\frac{D_2 x}{\lambda f_2}\right)\mathrm{sinc}\left(\frac{D_2 y}{\lambda f_2}\right) \end{aligned} \tag{16}$$

Thus, the resolution of the mask image on the detector is limited by the convolution with the sinc functions. The aperture at the first lens has to be at least as large as the mask size. Thus, we must have:

$$D_1 \geq M \tag{17}$$

The aperture at the second lens determines the resolution of the image. When D_2 approaches infinity, the convolution term becomes a delta function, thus forming a perfect image on the detector plane. A finite sized D_2 reduces the resolution of the image. We now examine the minimum detector spacing defined as the minimum spacing between the centers of two detectors, the detector size and the total number of detectors N. First, the detector size and the minimum detector spacing are limited by the width of the principal lobe of the sinc function which is expressed as

$$\triangle x = \frac{2\lambda f_2}{D_2} \tag{18}$$

Assuming that the f-number of lens L_2 is $f/\# = 1$ and the wave length is 0.8μm, we get $\triangle x$ as 1.6μm. The detector size and the spacing between the detectors are determined by numerically examining the convolution of the sinc function with an input image. We calculated the convolution term with several input images and the results are shown here.(Fig. 10)

We notice that a very small detector size with a very small spacing may cause a problem for a large N, since the influence of the nonprincipal lobes of the sinc function is not negligible. As we increase the detector size, it will become negligible.

Using this analysis, we can determine the overall size. In order to minimize the signal skew caused by the difference in the propagation delay, we minimize the lateral shift of the masks from the optical axis of the second lens. To minimize the lateral shift, we must use the smallest possible mask, the size of which is determined by the resolution of the mask. The resolution of the mask patterns using electron-beam lithography is about 0.5μm.[50] We assume that we use an array of $5 \times 5\mu$m PIN photodiode type of detectors such as SEEDs. As we look at the plots, we notice the convolution of $\Pi(x/3.6\mu\text{m})$ with the sinc function produces a spot of diameter about 5μm. Further examining the graph, we find that if we take the spacing larger than 5μm, the effect of the second lobe is negligible for N up to 2^{100}. If the mask pattern has a spot expressed as $\Pi(x/0.5\mu\text{m})$, we need the ratio of the focal length as $3.6/0.5 \approx 8$. This becomes also the magnification factor of the mask image to the detector image. If the spots on the mask pattern are separated by 1.5μm, this separation corresponds to 12μm on the detector plane. Each mask can be made in $1.5\sqrt{N} \times 1.5\sqrt{N}\mu$m. The first lenses L_1s have a focal length $1.5\sqrt{N}\mu$m with an aperture $1.5\sqrt{N} \times 1.5\sqrt{N}\mu$m. The second lens has a focal length $12\sqrt{N}\mu$m with an aperture $12\sqrt{N} \times 12\sqrt{N}\mu$m. The detector array has N 5×5 μm detectors fabricated in $12\sqrt{N} \times 12\sqrt{N}\mu$m. Then, the light path from the input lasers to the detectors is no more than $30\sqrt{N}\mu$m. The free space propagation delay of this distance is $0.1\sqrt{N}$ psec. We are using an off axis imaging. The difference in the propagation delay caused by the off axis fourier transform is proportional to the lateral shift of the mask from the optical axis. Thus, the maximum difference is $(s_{x'} + s_{y'})/v_c$. Here, $s_{x'}$ is the largest shift of the x-axis, $s_{y'}$ is of the y-axis, and v_c is the speed of light. In this case, we have $s_{x'} = s_{y'} = 0.75\sqrt{NJ\log N}\mu$m, thus the maximum difference is $5\sqrt{NJ\log N}$ fsec. For example, when $N = 1024$, this becomes 0.5 psec, and the total propagation delay is 3.2 psec. As we further increase N to 65,536, the difference is 5.12 psec, and the total propagation delay is 25.6 psec. Thus, a bit rate around a giga hertz does not cause a signal skew problem.

Detectable Power Limit

The largest N is limited by the minimum detectable power of the photodetectors. The optical signal emitted from one source is distributed to N positions of the mask. Let P_0 denote the total radiation power from the source. Then, ignoring other losses, each spot on the mask receives at most P_0/N optical power. In fact, if each spot is $0.5 \times 0.5\mu$m with a spacing 1.5 μm, it receives only $0.25P_0/N$ power. (Fig. 11) This power must correspond to the intensity I_1 at the detector plane.

In this section, we analyze the relation between P_0, N, and a bit rate B. The bit error rate (BER) of a given optical link is given by the integral [51]

$$P(E) = \frac{1}{\sqrt{2\pi}} \int_Q^\infty \exp\left[-\frac{x^2}{2}\right] dx \tag{19}$$

where Q is a parameter relating to the desired error rate $P(E)$. For the case of a PIN photodiode with a high-impedance FET front-end amplifier, the detector sensitivity can be approximated to

$$\eta\bar{P} = Q\frac{hv_c}{\lambda e}\sqrt{\langle i_n^2\rangle} \tag{20}$$

where h is Planck's constant, v_c is the speed of light, λ is the wave length, e is the fundamental charge, and $\langle i_n^2\rangle$ is the noise current power at a given bit rate. At high bit rates, the noise current power is dominated by the thermal channel noise term. This can be written as

$$\langle i_n^2\rangle \approx 8.26 \times 10^{-43} B^3 \tag{21}$$

Here we used the total input capacitance of 1 pF, the FET transconductance 50 mS, and the parameters for GaAs FET and RZ encoding for calculation. [51] For high speed interconnection networks, the BER of 10^{-17} is desired.[52] This rate guarantees an error free operation of a single link running at 1 GHz for about a year. The corresponding Q is 8. The minimum detectable optical power is then

$$\eta \bar{P} = 1.13 \times 10^{-20} B^{1.5} \tag{22}$$

Power versus bit rate is plotted in Figure 12.

The minimum detectable power at 1 GHz is an order of 1μW. Thus, for a bit rate of 1 GHz with BER of 10^{-17}, the required total radiation from a single laser source P_0 is expressed as

$$P_0 = 4N \; [\mu \mathrm{W}] \tag{23}$$

Laser diodes with radiation power of several hundred milliwatts are available. A laser diode with radiation power 300 mW can be distributed among 65,536 detectors.

Dynamic Range Limitation

The minimum power requirement of a detector is determined by its minimum detectable power. In this case, the detectors have to operate not only at this minimum detectable power, but also at a $\log N$ times the larger optical power. The dynamic range of a detector is defined as the difference between the minimum detectable power and the maximum allowable power. The dynamic range depends on the load resistor R_f of the amplifier. As the resistance decreases, the maximum allowable power increases. However, there is a trade off between the maximum allowable power and the minimum detectable power. The noise current power has a term proportional B/R_f. At a low bit rate, this term is not negligible. Thus at a low bit rate, as the load resistance decreases, the noise current power increases. High impedance PIN FET detectors typically have a dynamic range of 15-20 dB over the bit rates between 0.01-1GHz.[51] In this case, $\log N$ can be as large as 30-100. Thus, the dynamic range limitation does not cause any problems in this case.

3.2 DBD Optical Expander

In this architecture we use $\log N$ switching cells to deflect an input laser beam into one of N directions. The $\log N$ switches can be set simultaneously. Once they are set, data can be transmitted by a pulse modulated optical beam at an extremely high rate. Previously, a multiple stage digital laser beam deflector was developed by Meyer et al[8]. This provided a flexible control of an input laser beam by using multiple staged deflectors.(Fig. 13) However, as we described earlier, it had several disadvantages such as a high drive voltage and a variable switch size.

Another approach uses a tree structured N switching cells in $\log N$ stages.(Fig. 14) Each switch can be a conventional electrooptic directional coupler.[49]

A generic model of directional coupler consists of two closely spaced optical guides embedded in an electrooptic substrate like $LiNbO_3$ or GaAs. Electrodes are used to apply an external field to the guides. The applied electric field changes the index of refraction of the guides by the electrooptic effect. To operate the switch, two optical input channels, one control, and two output channels are used. Depending on the control signal, it establishes straight or cross connections between the input and the output channels. To construct the optical expander, this approach requires the total of N switches configured in a complete binary tree structure of depth $\log N$. Here, unlike the previous approach, the switch size is kept constant for each switching position. However, it requires the total of N switches. Our approach does not inherit the drawbacks of these previously proposed systems. It uses only $\log N$ switches of a constant size to produce the output of size N. Thus, we expect that our optical expander can be advantageous for a large N.

3.2.1 Multistage Beam Deflector

We want to keep the cell size equal through all the $\log N$ stages. Then, we can use a set of $\log N$ identical switching cells for all the stages, and also use the same control mechanism for each of them. If at each stage the laser beam is confined in a constant area, the size of each switch can be constant. At the same time, the output beam must be deflected from the original position by a certain amount according to the control signal. To achieve this, we design a reflection/transmission (R/T) cell as follows(Fig. 15). We use these cells to implement a deflector.(Fig. 16)

R/T Switching Cells

At each stage the laser beam hits the cell with a certain angle. Each possible angle corresponds to a possible deflection angle. Therefore, at the first stage there is a single incident angle. There are 2^i distinct incident angles at the i-th stage. It is desirable that the switch has a reflection mode and a transmission mode at any incident angle. However, in practice, this condition may not be satisfied. At this point, our choice in realizing a R/T cell is limited. We consider several approaches which are currently available.

One possibility is to use an optical switch based on electrochemically generated bubbles.[53] The switching element is a slot which is filled with a fluid. The slot is etched in a substrate whose index of refraction is approximately the same as that of the fluid. If there is no bubble in the slot, the incident light is transmitted. When a bubble is electrochemically introduced to the slot, the incident light experiences a total internal reflection from the slot. It is a novel bistable optical switch which is both polarization and wavelength independent. However, the switching speed is too slow for rapidly switching networks. On the other hand, if the network is static for an order of minutes and does not require fast switching, this may provide an efficient data transfer at a high bit rate.

Another possibility is to use a switch which is based on polarization dependent materials. One can use a nematic liquid crystal (NLC) switch. The NLC switches have been previously demonstrated.[54] The NLC switch has NLC molecules which are aligned in a plane. The orientation of the molecules is changed by an applied voltage. Assuming positive dielectric anisotropy, at a low voltage, the molecules are aligned in the layer normal to the direction of light propagation. This corresponds to the optical axis being parallel to the layer but normal to the light propagation. Thus, an incident light whose electric displacement is parallel to the layer behaves as the extraordinary wave. The index of refraction associated with the extraordinary wave is denoted by n_e. For the NLC at $\lambda = 633$ nm, n_e is 1.63. The layer of NLC molecules is sandwitched between a waveguide whose index of refraction is approximately n_e. Thus, the s-polarized incident light is transmitted through the layer. At a high voltage, the molecules rotate 90°, and align normal to the layer. Thus, the optical axis becomes normal to the layer. At this time, the s-polarized incident light behaves as the ordinary wave whose index of refraction n_o is 1.49. Thus, at the incident angle which is greater than the critical angle $\theta_c = sin^{-1}(1.49/1.63) \approx 66°$, it experiences a total internal reflection.(Fig. 17) The switching speed is limited by that of NLC.

To achieve a higher switching speed, an alternative approach is to keep the NLC as a static polarizing beam splitter (PBS) and to use ferroelectric liquid- crystal (FLC) polarization rotator (PR) as an active switching device. Optical switches which used a combination of PRs and PBSs have been demonstrated.[55] They designed an $N \times N$ permutation network which used $N - 1$ stages of $N/2$ switches. All signals passing through the network must go through $N - 1$ switches before exiting. Another optical switching network using a similar switch is a $1 \times N$ optical switch.[56] In their system, each pair of PR and PBS constitutes one port in the $1 \times N$ switch. All the PRs are set to pass the polarized incident light except at the selected port, where the incident light is rotated and the light is reflected at the PBS. They obtained a switching time of 50μsec. In their architecture, the light beam must sequentially go through N stages of switching cells. Our DBD optical expander uses only $\log N$ switches. In this type of architecture the incident angle must stay between the critical angle for p-polarized light and that for s-polarized light, so that at each cell the s-polarized light is transmitted through the cell, while the p-polarized light is totally reflected from the cell. The DBD optical expander using a combination of FLC PR and NLC PBS is shown in figure 18.

It may be possible for an alternative material to be discovered. Then, a significantly simpler and more efficient beam deflector could be constructed.

3.3 Overall Architecture and Analysis of DBD Optical Expander

In this model, the control signal encoded in $\log N$ bits controls the $\log N$ stages of the cells. There are several advantages in the DBD optical expander. One is that the input laser beam can be modulated at a very high rate. Once all the switches are set, the data transfer rate is not limited by the switching speed of the optical expander. Another advantage is that the energy of the input laser beam is efficiently transmitted as the output. The final advantage is that the propagation of each path can be made equal so that the output is always produced after the same propagation delay. As we can see from figure 19, the total length of the system is proportional to N. There are two limiting factors which determine the largest N. One is insertion loss at the interface of each switch, and the other is crosstalk.

3.3.1 Insertion Loss and Crosstalk

The incident angle to the NLC may vary from 65° to 84°. A false reflection which is caused by a slight mismatch of the indices of refraction at the interface is a major loss and crosstalk reason. Thus, here we consider the loss at the interface of the NLC surfaces. We assume the total internal reflection occurs perfectly when the incident angle is greater than the critical angle which is associated with the incident wave. At transmission, we must consider a loss caused by an amount of false reflection at the interface. If the waveguide and the NLC have exactly the same index of refraction, the incident light passes through the interface without a reflection. However, in general, there is a slight mismatch in the indices of the two media. If the incident angle is fixed, it is possible to significantly reduce the reflection by using an anti-reflection coating. If there is no anti-reflection coating, the transmittivity of the s-polarized light through the crystal at a transmission state can be expressed as [57]

$$\begin{aligned} T &= \left(\frac{2\sin\theta_e\cos\theta_g}{\sin(\theta_g+\theta_e)}\frac{2\sin\theta_g\cos\theta_e}{\sin(\theta_e+\theta_g)}\right)^2 \\ &= \left(\frac{4\sin\theta_e\cos\theta_g\sin\theta_g\cos\theta_e}{\sin^2(\theta_g+\theta_e)}\right)^2 \end{aligned} \tag{24}$$

Here, θ_g is the incident angle to the crystal, θ_e is the transmitted angle in the crystal, n_g is the index of refraction in the waveguide, and n_e is the index of refraction associated with the extraordinary wave in the crystal. The first term in the parenthesis corresponds to the transmission from the waveguide into the crystal, and the second term corresponds to the transmission from the crystal back to the waveguide. Using $\sin\theta_e = (n_g/n_e)\sin\theta_g$, the total transmittivity can be written as a function of θ_g.

$$T(\theta_g) = \left(\frac{4n_g\sin^2\theta_g\cos\theta_g\cos(\sin^{-1}(\frac{n_g}{n_e}\sin\theta_g))}{n_e\sin^2(\theta_g+\sin^{-1}(\frac{n_g}{n_e}\sin\theta_g))}\right)^2 \tag{25}$$

Total transmittivity vs. incident angle is plotted in figure 20.

The total reflection of the s-polarized light starts at $\theta_g \approx 84°$. We find that the transmittivity through the crystal can be kept above 96% at the incident angle from 65° to 80°. A higher transmittivity of 99.5% can be obtained if the angle is kept below 75°. Since the input light goes through $\log N$ switches, the crosstalk power does not accumulate.

Unlike the MVM optical expander, the DBD optical expander still lacks a desired switching material, it would be premature to analyze its size using the current assumptions. Therefore, here we discuss the issue of extending the expander into a two dimensional output model, and the proportional size in terms of the size of output N. In the previous figure, the output is produced in a one dimensional vector format. If we have a reflection/transmission switch which reflects or transmit the incident light with any incident angle, it is easy to extend the system to a two dimensional output model.(Fig. 21)

However, the switches based on the NLC require a fixed plane of incidence in order to keep the electric displacement vector of the s-polarized incident light either parallel or perpendicular to the optical axis of the crystal. If not, the light will be decomposed into the ordinary wave and the extraordinary wave inside the crystal. Therefore, under the current circumstances, we must modify our scheme. One can decompose the process of generating a $\sqrt{N}\times\sqrt{N}$ output into two pieces. First we use $(1/2)\log N$ switches of width 1 to produce the output of size $\sqrt{N}$ in a one dimensional vector format. Then, we use $(1/2)\log N$ switches of width $\sqrt{N}$ to generate the output of size $\sqrt{N}\times\sqrt{N}$. The total number of switches is still $\log N$, but half of the switches have a width $\sqrt{N}$. The length of the system is proportional to N.

4 Conclusion

We investigated a problem of electrooptically creating one of the N mutually orthogonal boolean patterns from a given input pattern encoded in $c\log N$ bits, where c is a constant no greater than 2. Several approaches are considered in this paper. One approach is based on the idea of implementing a large line decoder by using optical interconnections. This is done by using optical matrix-vector multiplication followed by a thresholding operation. We analyzed the approach in terms of size, speed, and power requirement. The other approach uses $\log N$ small identical switches to implement a digital beam deflector. The optimal device to implement this switching functions has not yet been found. We considered several choices which are available with current techonology.

We also discussed the applications of optical expanders and showed the importance of the design and development of our optical expanders.

References

[1] A. A. Sawchuk and T. C. Strand, "Digital Optical Computing," Proceedings of IEEE **72(7),** 758-779 (1984).

[2] T. E. Bell, "Optical Computing: A field in flux," IEEE Spectrum **23(8),** 34-57 (1986).

[3] J. E. Midwinter and M. G. Taylor,"Optoelectronic Interconnects in VLSI: The Reality of Digital Optical Computing?," IEEE LCS 40-46 May (1990).

[4] D. Feitelson, *Optical Computing, A Survey for Computer Scientists* (MIT Press, Cambridge, Mass., 1988).

[5] R. Arrathoon, *Optical Computing : digital and symbolic* R. Arrathoon, ed., (Marcel Dekker, New York, 1989).

[6] R. J. Collier, C. B. Burckhardt, and L. H. Lin, *Optical Holography* (Bell Telephone Laboratories, Murray Hill, New Jersey 1971).

[7] D. A. Pinnow and R. W. Dixon, "Alpha-Iodic Acid: A Solution-Grown Crystal with a High Figure of Merit for Acousto-Optic Device Applications," Appl. Phys. Lett. **13,** 156-158 (1968).

[8] H. Meyer, D. Riekmann, K. P. Schmidt, U. J. Schmidt, M. Rahlff, E. Schröder, and W. Thust , "Design and Performance of a 20-Stage Digital Light Beam Deflector," Appl. Opt. **11,** 1732-1736 (1972).

[9] T. Venkatesan, B. Wilkens, Y. H. Lee, M. Warren, G. Olbright, H. M. Gibbs, N. Peyghambarian, J. S. Smith, and A. Yariv, "Fabrication of arrays of GaAs optical bistable devices," Appl. Phys. Lett. **48,** 145-147 (1986).

[10] J. L. Jewell, A. Scherer, S. L. McCall, A. C. Gossard, and J. H. English, "GaAs-AlAs monolithic microresonator arrays," Appl. Phys. Lett. **51,** 94-96 (1987).

[11] D. A. B. Miller *et al.*, "Novel Hybrid Optically Bistable Switch: The Quantum Well Self-Electro-Optic Effect Device," Appl. Phys. Lett. **45,** 13-15 (1984).

[12] V. J. Fowler and J.Schlafer, "A Survey of Laser Beam Deflection Techniques," Proceedings of the IEEE **54,** 1437-1444 (1966).

[13] K. Hwang and F. A. Briggs, *Computer architecture and parallel processing* (McGraw-Hill, New York 1984).

[14] T. Feng , "A Survey of Interconnection Networks," IEEE Comput. **14,** 12-27 (1981).

[15] J. W. Goodman, F. J. Leonberger, S. Kung, and R. A. Athale, "Optical interconnections for VLSI systems," Proceedings of the IEEE **72,** 850-866 (1984).

[16] L. D. Hutcheson, P.Haugen, and A.Husain, "Optical interconnects replace hardwire," IEEE Spectrum **24(3),** 30-35 (1987).

[17] H. J. Caulfield, J. A. Neff, and W. T. Rhodes , "Optical Computing: The coming revolution in optical signal processing," Laser Focus **19(11),** 100-109 (1983).

[18] D. A. B. Miller, "Optics for low-energy communication inside digital processors: quantum detectors, sources, and modulators as efficient impedance converters," Opt. Lett. **14,** 146-148 (1988).

[19] M. R. Feldman, S. C. Esener, C. C. Guest, and S. H. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations," Appl. Opt. **27,** 1742-1751 (1988).

[20] M. R. Feldman, C. C. Guest, T. J. Drabik, and S. C. Esener, "Comparison between electrical and free space optical interconnects for fine grain processor arrays based on interconnect density capabilities," Appl. Opt. **28,** 3820-3829 (1989).

[21] J. D. Ullman, *Computational Aspects of VLSI* (Computer Science Press, Rockwille, Md., 1984).

[22] R. Barakat and J. Reif, "Lower Bounds on the Computational Efficiency of Optical Computing Systems," Appl. Opt. **26,** 1015-1018 (1987).

[23] J. A. Rajchman, "An Optical Read-Write Mass Memory," Appl. Opt. **9,** 2269-2271 (1970).

[24] W. C. Stewart and L. S. Cosentino, "Optics for a Read-Write Holographic Memory," Appl. Opt. **9,** 2271-2275 (1970).

[25] H. Kiemle, "Considerations on Holographic Memories in the Gigabyte Region," Appl. Opt. **13,** 803-807 (1974).

[26] L. d'Auria, J. P. Huignard, C. Slezak, and E. Spits, "Experimental Holographic Read-Write Memory Using 3-D Storage," Appl. Opt. **13,** 808-818 (1974).

[27] D. Chen and J. D. Zook , "An overview of opticadl data storage technology," Proceedings of the IEEE **63,** 1207-1230 (1975).

[28] S. Redfield and L. Hasselink,"Enhanced nondestructive holographic readout in strontium barium niobate," Opt. Lett. **13,** 880-882 (1988).

[29] E. H. Paek, J. R. Wullert II, M. Jain, A. Von Lehmen, A. Scherer, J. Harbison, L. R. Florenz, H. J. Yoo, J. L. Jewell, and Y. H. Lee, "Compact and ultrafast holographic memory using a surface-emitting microlaser diode array," Opt. Lett. **15,** 341-343 (1990).

[30] W. D. Hills, *The Connection Machine* (MIT Press, Cambridge, Mass., 1985).

[31] A. D. McAulay, "Optical crossbar interconnected digital signal processor with basic algorithms," Opt. Eng. **25,** 82-90 (1986).

[32] A. A. Sawchuk, B. J. Jenkins, C. S. Raghavendra, and A. Varma, "Optical matrix-vector implementation of crossbar interconnection networks," Proceedings Intl. Conf. on Parallel Processing 401-404 Aug. (1986).

[33] A. A. Sawchuk, B. J. Jenkins, C. S. Raghavendra, and A. Varma, "Optical Crossbar Networks," Computer **20(6),** 50-60 (1987).

[34] B. Clymer and S. A. Collins Jr., "Optical computer switching network," Opt. Eng. **24,** 74-81 (1985).

[35] G. Eichmann and Y.Li, "Compact optical generalized perfect shuffle," Appl. Opt. **26,** 1167-1169 (1987).

[36] A. W. Lohmann, W. Stork, and G. Stucke, "Optical perfect shuffle," Appl. Opt. **25,** 1530-1531 (1986).

[37] A. W. Lohmann "What classical optics can do for the digital optical computer," Appl. Opt. **25,** 1543-1549 (1986).

[38] E. S. Maniloff, K. M. Johnson, and J. H. Reif, "Holographic Routing Network for Parallel Processing Machines," EPS/Europtica/SPIE International Congress on Optical Science and Engineering, Paris, France, April (1989).

[39] E. S. Maniloff and K. M. Johnson , "Dynamic holographic interconnects using static holograms," Opt. Eng. **29,** 225-229 (1990).

[40] J. Von Neumann, *The Computer and the Brain* (Yale Univ. Press, New Haven, 1958).

[41] M. Murdocca, *A Digital Design Methodology for Optical Computing* (MIT Press, Cambridge, Mass., 1990).

[42] A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham, and L. M. F. Chirovsky, "Symmetric self-electro-optic effect device: Optical set-reset latch," Appl. Phys. Lett. **52,** 1419-1412 (1988).

[43] F. B. McCormick and M. E. Prise, "Optical circuitry for free-space interconnections," Appl. Opt. **29,** 2013-2018 (1990).

[44] A. L. Lentine, D. A. B. Miller, J. E. Henry, J. E. Cunningham, L. M. F. Chirovsky, and L. A. D'Asaro, "Optical logic using electrically connected quantum well PIN diode modulators and detectors," Appl. Opt. **29,** 2153-2163 (1990).

[45] M. E. Prise, N. C. Craft, R. E. LaMarche, M. M. Downs, S. J. Walker, L. A. D'Asaro, and L. M. F. Chirovsky, "Module for optical logic circuits using symmetric self-electrooptic effect devices," Appl.Opt. **29,** 2164-2169 (1990).

[46] S. Sakai, H. Shiraishi, and M. Umeno, "AlGaAs/GaAs Stripe Laser Diodes Fabricated on Si Substrates by MOCVD," IEEE J. Quantum Electron **QE-23,** 1080-1084 (1987).

[47] S. Sakai, X. W. Hu, and M. Umeno, "AlGaAs/GaAs Transverse Junction Stripe Lasers Fabricated on Si Substrates Using Superlattice Intermediate Layers by MOCVD," IEEE J. Quantum Electron **QE-23,** 1085 (1987).

[48] W. Dobblelaere, D. Huang, M. S. Unlu, and H.Morkoç, "AlGaAs/GaAs Multiple Quantum Well Reflection Modulators Grown on Si Substrates," Appl. Phys. Lett. **55,** 94-96 (1988).

[49] K. Iizuka, *Engineering Optics* 2nd ed. (Springer-Verlag, New York, Berlin, Heidelberg 1987).

[50] M. R. Feldman and C. C. Guest, "Interconnection density capabilities of computer generated holograms for optical interconnection of very large scale integrated circuits," Appl. Opt. **28,** 3134-3137 (1989).

[51] T. V. Muoi, "Receiver Design for High-Speed Optical-Fiber Systems," J. Lightwave Tec. **LT-2(3),** 243-267 (1984).

[52] D. H. Hartman, "Digital high speed interconnects: a study of the optical alternative," Opt. Eng. **25,** 1086-1102 (1986).

[53] J. L. Jackel, J. J. Johnson, and W. J. Tomlinson, "Bistable switching using electrochemically generated bubbles," Opt. Lett. **15,** 1470-1472 (1990).

[54] R. A. Soref and D. H. McMahon, "Total switching of unpolarized fiber light with a four-port electro-optic liquid-crystal device," Opt. Lett. **5,** 147-149 (1980).

[55] L. R. McAdams, R. N. McRuer, and J. W. Goodman, "Liquid crystal optical routing switch," Appl. Opt. **29,** 1304-1307 (1990).

[56] L. R. McAdams and J. W. Goodman, "Liquid crystal $1 \times N$ optical switch," Opt. Lett. **15,** 1150-1152 (1990).

[57] R. D. Guenther, *Modern Optics* (Wiley, New York, 1990).

Decimal	(1) $c = 2$	(2) $c \neq 2$
0	000111	000111
1	001110	001011
2	010101	001101
3	011100	001110
4	100011	010011
5	101010	010101
6	110001	010110
7	111000	011010
8		011100
.		...
18		110100
19		111000

Table 1: Two Encoding Schemes $d = 6$

Figure 1: Basic architecture of holographic memory

Figure 2: Optical Crossbar Network

Figure 3: Holographic Message Routing System

Figure 4: Standard Optical Matrix-Vector Multiplier

Figure 5: Example of Another Matrix-Vector Multplier

Figure 6: Mask Patterns of Two Encoding Schemes

Figure 7: Differential Device

Figure 8: MVM Optical Expander

Figure 9: Side View of Optical Expander

Figure 10: Diffraction Limited Image 1

Figure 10: Diffraction Limited Image 2

Figure 11: Mask Pattern

Figure 12: Power vs. Bit Rate

Figure 13: Electrooptical Multistage Deflector

Figure 14: Directional Coupler Based Network

Figure 15: Reflection/Transmission Cell

Figure 16: DBD Optical Expander

Figure 17: NLC Based R/T switch

Figure 18: FLC PR and NLC PBS Based R/T Switch

Figure 19: Proportional Size of DBD Optical Expander

Figure 20: Transmittivity at Transmission State

Figure 21: 2-D DBD Optical Expander