10th International

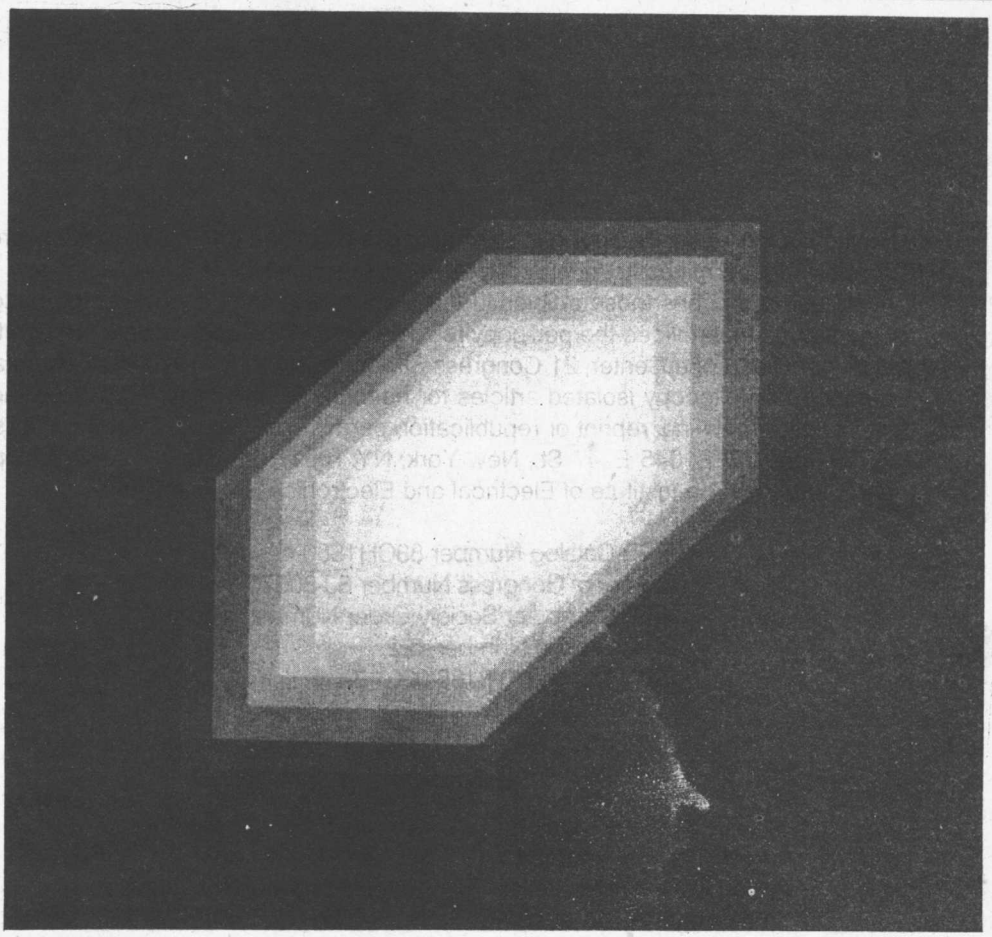
Optical Computing Conference

1983

10th International Optical Computing Conference

April 6 – 8, 1983
Massachusetts Institute of Technology, Cambridge MA

Sponsored by:
IEEE Computer Society
International Commission for Optics
SPIE—The International Society for Optical Engineering



IEEE CATALOG NUMBER 83CH1880-4
LIBRARY OF CONGRESS NUMBER 83-80573
IEEE COMPUTER SOCIETY ORDER NUMBER 465
SPIE VOLUME NUMBER 422
ISBN NUMBER 0-8186-0004-7

IEEE COMPUTER SOCIETY PRESS
IEEE COMPUTER SOCIETY
INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS

Published by IEEE Computer Society Press
1109 Spring Street
Suite 300
Silver Spring, MD 20910

IEEE Catalog Number 83CH1880-4
Library of Congress Number 83-80573
IEEE Computer Society Order Number 465
SPIE Volume Number 422
ISBN Number 0-8186-0004-7

Order from: IEEE Computer Society
Post Office Box 80452
Worldway Postal Center
Los Angeles, CA 90080

IEEE Service Center
445 Hoes Lane
Piscataway, NJ 08854

The Institute of Electrical and Electronics Engineers, Inc.

PREFACE

These are the Proceedings of the 10th International Optical Computing Conference. This, however, is more than just one in a sequence of technical meetings. Instead it marks what is undoubtedly the turning point in the history of optical computing. The words "**optical computing**" seemed foolish to many when the first conference was held. Even in the ninth conference there was progress but no feeling of impending success crowning our efforts. In this, the 10th conference, all of that has changed. We have now learned how to make optical computing not only fast and low-power but also highly accurate. The breakthroughs in digital optical computing have given the participants in this conference the feeling of participation in the beginnings of a revolution. Well financed efforts are under way in the United States, the Soviet Union, Japan, and several other countries to exploit the recently proved ability of optical computing to do special purpose operations that are undreamed of by its electronic counterparts. I hope this Proceedings can go far in conveying to its readers something of the excitement of the meeting.

H.J. Caulfield
Program Chairman

Table of Contents

Late Papers

CHANCES FOR OPTICAL COMPUTING

by A. W. Lohmann

Physikalisches Institut der Universität Erlangen
Erwin-Rommel-Straße 1, 8520 Erlangen, W-Germany

Abstract

Optical computing can mean several things: computing with optical hardware, or/and computing of optical signals. Optical signals, in other words: pictures, often consist of many million bits. If pictures have to be processed for example at TV rate, the optical computer has an edge over the digital computer.

This example is one among several others, that may serve to stabilize the self-confidence of the optical computing community.

1. Optical computing: what is it?

Questions can be answered easier if cut into pieces. My favorite cutting procedure is to divide the issue into a two-by-two matrix.

Categories of signal processing

hardware \ signal	electronic (1D)	optical (2D)
electronic		+ +
optical	+	+ + +

This box identifies three types of optical computing:

processing by means of optical hardware (lower left)
processing of optical signals = pictures (upper right)
picture processing with optical hardware (lower right)

The boundaries are not sharp. Some systems contain sub-systems of different categories. These "hybrid" systems are very interesting. But they are only as good as the transducers are, that connect electronic and optical hardware.

The crosses reflect my own taste and also the popularity of the three types of optical computing on my side of the Atlantic, I guess.

2. Where are our chances ?

There are probably hundred times more people working in digital electronic computing, as compared to us here at IOCC. A direct confrontation is hopeless. What is it that we optics people can do better? We can perform computations at high speed at low cost. For example, the input may consist of 25 pictures per second, each with 4×10^5 pixels. These pictures are projected onto a mask with 4×10^5 pixels. Thereafter the light is collected, or integrated. This constitutes a correlation process with 10^7 MULT/ADD operations per second, for the price of standard TV, a reasonably good lens and a photomultiplier.

Our simple example, which is certainly not the best there is, indicates some features that are favorable for optical computing in general:
(a) the input is in the format of a picture (2D);
(b) the hardware is hybrid (TV + optics);
(c) the algorithm is simple (linear).
Point (c) indicates what is a basic feature of optical computers: they are special purpose computers with a restricted range of algorithms. That is true also for electronic array processors, which are firmly established now in data processing. Hence, "special purpose" is not a forbidding handicap. More severe is the fact that our example constitutes the equivalent of a "hard-wired" processor. The program cannot be modified at high speed. Suitable applications are repetitive routines, mass production of correlations, for example.

Optical computing has a chance, if the data processing project is difficult and/or expensive for the digital computer. But the project has to be within the range of capabilities of optical computing, as outlined above. We should try constantly

CH1880-4/83/0000/0001$01.00 © 1983 IEEE

to improve our chances by enlarging the range of capabilities.

3. Plan of the following chapters

We will begin with the lower left box of our matrix of categories: computing with optical hardware, applied to one-dimensional signals. That chapter will be short only due to my lack of experience in that field. To demonstrate that I am not biased against digital computers, Van Vleck's theorem will be recommended as a tool for the efficient digital correlation of random pictures (speckle pattern). So much for the upper right box. The remaining part of this paper will be devoted to several items of the lower right box, including among others: the stealing of an algorithm, optical logic and triple correlation.

As a "key note lecture" this is more of a pep talk than an ordinary paper. Hence, detailed facts and references are omitted here. A more thorough presentation of the various projects can be found in our regular journal publications and in our annual reports.

4. Processing of one-dimensional data with optical hardware

The hardware in this case consists mainly of integrated optics, fiber optics and acousto-optics.

This field is certainly quite prosperous, mainly because the technology of the components is similar as for IC circuitry. Also the architecture often resembles that of electronic processors. Both features are favorable for the acceptance of such systems.

Well-known examples are optical matrix multipliers and acousto-optical signal correlators. For example, a signal of 10 MHZ bandwidth can be correlated in real time with a time window of 100 microseconds. A digital computer would have to sample at a rate of 10^7 per second and perform 1000 MULT/ADD operations at every 10^{-7} second. The processing speed of 10^{10} operations per second is safely in the exclusive domain of optical computing.

5. Van Vleck's theorem as an efficient tool for digital correlation

Van Vleck's theorem is useful if the autocorrelation of a wild signal (random process) has to be computed. The necessary assumptions (gaussian histogram etc.) are stated in textbooks. You may think of speckle patterns I(x,y) as a "wild signal". Suppose now we want to compute the autocorrelation AC(x,y):

$$AC(x,y) = I(x,y) \otimes I(x,y) \qquad (1)$$

This job is quite elaborate. If I(x,y) contains N samples (say N = 1 million) we have to perform N^2 MULT/ADD operations. A smarter algorithm with only 4N log N operations consists of the following steps: Fourier, modulus square, inverse Fourier.

Van Vleck's theorem offers a faster algorithm. It consists of hardclippling (= binarisation), autocorrelation, arc sine (= $\sin^{-1}$) nonlinearity:

$$I(x,y) \rightarrow H(x,y) \rightarrow H \otimes H = HC$$

$$\rightarrow (2/\pi)\arcsin[HC] = AC(x,y) \qquad (2)$$

The advantage of this algorithm is based on the simplicity of correlating 1-bit signals H(x,y). The N^2 MULT operations of $I \otimes I$ are replaced by N^2 simple AND operations in $H \otimes H$. E. L. O'Neill made me aware of the usefulness of Van Vleck's theorem in the context of speckle interferometry many years ago. Meanwhile Van Vleck's theorem seems to be commonly appreciated by speckle specialists.

6. Stealing a fast algorithm

Many smart people have invented fast algorithms for the digital computer. Very famous is the fast Fourier transform FFT algorithm. It consists of log N layers. From one layer to the next layer one has to perform 2N MULT/ADD operations, all together 2N log N operations. The fast Hadamard algorithm has the same structure. Only the coefficients, that have to be applied when going from one layer to the next layer, are different from the Fourier case.

J. Jahns found a way for stealing the fast Hadamard algorithm. "Stealing" means the conversion from discrete signals to continuous signals, as they are common in analog optics. It may be worthwhile to explain the technique of this act of robbery because it might be possible to do something similar with other algorithms in the digital field.

First, we have to clarify how the analysis of continuous signals and the algebra of discrete signals are related to each other. It is the well-known sampling theorem that ties together these two branches of mathematics:

$$I(x) = \Sigma(n)\ I(n \cdot a)\,\mathrm{sinc}(n - x/a) \qquad (3)$$

Herein is I(x) the subject of continuous analysis and I(n·a) the subject of discrete algebra. This becomes quite apparent, for example, if two functions $I_1(x)$ and $I_2(x)$ are convolved.

$$I(x) = I_1(x) * I_2(x);$$

$$I(na) = \sum_{(m)} I_1(ma) I_2((n-m)a) \qquad (4)$$

Symbolically we may express the relationship between analysis (ALYS) and algebra (ALG) in the following form:

$$ALYS + SAMPLE = ALG \qquad (5)$$

$$ALG + INTERPOL = ALYS \qquad (6)$$

By SAMPLE we mean $I(x) \rightarrow I(na)$ and by INTERPOL $I(na) \rightarrow I(x)$, based on eq. 3.

Now we are prepared to understand J. Jahns' idea. He noticed that the matrix, that describes the step from one layer to the next layer of the fast Hadamard algorithm, is essentially a TOEPLITZ matrix. TOEPLITZ in algebra is equivalent to "shift-invariant" in analysis. Hence, the action of the TOEPLITZ matrix can be performed by a very fast analog computer, the optical spatial filtering processor that implements a convolution. The convolution is the prototype of a linear space-invariant operation. Fortunately, the matrices that connect Hadamard layers are all identical. Hence, one has to apply the optical convolution process log N times sequentially.

This description of J. Jahns' modified Hadamard algorithm is simplified here because I wanted to illustrate how a digital algorithm was converted into an equivalent analog algorithm. The matrix mentioned above is not quite of TOEPLITZ structure. It consists of two interlaced TOEPLITZ matrices that have to be executed in parallel, followed by interlacing and with the inverse of interlacing ("alternate shifting") before convolving.

7. Logic processing, based on theta modulation

In the mid-sixties several people attempted to do truly digital computing by optical means. Laser quenching was one of the basic effects under consideration. I played with a spatial filtering experiment, called "theta modulation" (Appl. Opt. 4, 399, 1965). A few years later Keyes and Armstrong argued on the basis of general laws of physics, that something like an "optical transistor" would probably never be competitive with the electronic transistor. One needs nonlinear interactions in order to implement logic operations. Optics is not non-linear at low power levels - usually a blessing, but not for computing.

Recently, Y. Ichioka (see his contribution at this IOCC conference) revived the interest in optical logic. It is still true, what Keyes and Armstrong said, the individual logic operation is comparatively slow, if performed with photons. But this may be more than offset if many slow operations take place simultaneously. A degree of parallelism of 100 000 is not difficult to do, optically. Such an optical computer would probably be essentially "hard-wired" and of SIMP structure. (SIMP = single instruction, multiple processor).

Stimulated by Y. Ichioka, we took up optical computing again, based on theta modulation. H. Bartelt implemented a theta encoder, based on TV technology. This encoder converts a gray picture into a theta-modulated two-dimensional signal. By means of spatial filtering we can do things like "bit-slicing". When conbining two modulated pictures we can perform the AND operation, for example.

8. Applications of computer generated holograms

A computer hologram (CGH) as spatial filter in an optical filtering setup is like a piece of "firm ware" in a digital computer. Our piece of firm ware is produced by a digital computer, slowly and not cheap. But if the CGH is used as filter to perform convolutions at optical speed on a routine basis, it might be worth its price.

Even today new ways of spatial filtering to be found, as for example in Ch. Thum's projects on "two way code translation" and "modified matched filtering".

9. Phase conjugation as a case of UF computing

UF stands for ultrafast. Suppose we want to transmit an image through a turbulent medium. We first probe the medium by means of a plane wave, coming from an isolated object point. Knowing the distortion caused by the turbulence, we can pre-compensate the object such that the image on the other side of the turbulence will arrive perfectly sharp.

The amount of computation for this task consists of three Fourier transforms, one division and one multiplication, each of these processes done with N pixels (N = 1 million). The number of MULT/ADD operations is about $6N \log N = 36 \times 10^6$. The whole operation has to be repeated about 30 times per second, if the time constant of the turbulence is 1/30 second. That amounts to a processing speed of 10^9

operations per second. In "phase conjugation" this is done, actually. Hence, phase conjugation deserves to be called UF computation.

10. The triple correlation

The (auto-)triple correlation is defined as:

$$\int u(x')u(x'+x)u(x'+y)dx' = T(x,y) \qquad (7)$$

The Fourier transform of T(x,y) is called "Bi-spectrum":

$$\tilde{T}(\nu,\mu) = \iint T(x,y)\exp[-2\pi i(\nu x+\mu y)]dxdy$$

$$\tilde{T}(\nu,\mu) = \tilde{u}(\nu)\tilde{u}(\mu)\tilde{u}(-\nu-\mu) \qquad (8)$$

What I like about the triple correlation is the vast amount of computation that is needed to obtain the two-dimensional T(x,y) from the one-dimensional u(x). If the signal u is a picture u(x,y), the triple correlation will be four-dimensional. Suppose we can compute the triple correlation by optical means, we are probably not endangered by the competition from our digital friends. But is the triple correlation useful for anything?

Yes, the triple correlation is useful, I believe. This insight came to me when I read in "Applied Optics" a paper by T. Sato, a couple of years ago. Sato mentioned that T(x,y) is sometimes quite insensitive to noise, under favorable conditions. To understand this, we replace in eq. 7 the pure signal u(x) by signal plus noise:

$$u(x) = u_o(x) + n(x) \qquad (9)$$

The properties of the noise are assumed to be: signal-independent, stationary, zero-mean. The object per se $u_o(x)$ may also be zero in average:

$$\int u_o(x)dx = 0 \qquad (10)$$

Under these circumstances, the ensemble average of eq. 7 with eq. 9 reduces to:

$$\langle T(x,y)\rangle = T_o(x,y) + \langle T_N(x,y)\rangle \qquad (11)$$

The noise term $\langle T_N\rangle$ will vanish if the probability density function p(n) is symmetrical. In other words, if the skewness of the noise is zero.

When I read T. Sato's paper, I was very happy about the noise-insensitivity of the triple correlation, because I realized that we had been using triple correlations already for several of our projects, without knowing the term "triple correlation" and its noise-insensitivity.

G. Weigelt had invented a modified version of speckle interferometry, called "speckle masking". With that method it is possible to obtain true astronomical images, not only autocorrelations.

C. Thum had measured the conditional probability of pairs of adjacent pixels (pulsed samples) of a halftone image.

G. Weigelt, F. Ernst and J. Übler had measured the conditional probability of bacteria that move north (for example) during (t_n, t_{n+1}) and move east during (t_{n+1}, t_{n+2}). To measure this type of conditional probability one has to evaluate a triplet of movie frames from instants t_n, t_{n+1}, t_{n+2}.

There are more examples of triple correlations in the literature. For example the image contrast of a periodic object, illuminated in partial coherence, is a triple correlation of the source distribution and of two shifted versions of the pupil function. A holographical associative memory (Denisyuk, van Heerden, Gabor) can also be considered as a case of triple correlation. The recording of the memory consists in photographing the joint power spectrum $|\tilde{u}_1(\nu) + \tilde{u}_2(\nu)|^2$. In the read-out process the joint power spectrum is illuminated (= multiplied) by the Fourier spectrum of one of the two signals $\tilde{u}_1$ or $\tilde{u}_2$. The output is a third-order integral:

$$\int|\tilde{u}_1+\tilde{u}_2|^2\cdot\tilde{u}_1\exp(2\pi i\nu x)d\nu = \hat{u}_2(x)+\ldots \qquad (12)$$

The example of the associative memory may appear to be somewhat artificial. But a certain project in oceanography indicates nicely, why a triple correlation is sometimes indispensable. Suppose we want to study the breaking of ocean waves. In deep water the wave profile is cosine-like. It is easy to see why the triple correlation of $\cos(2\pi\nu_o x)$ is zero, since the bi-spectrum (eq. 8) is zero everywhere. Approaching the coast in shallow water a nonlinear component will be added to the profile which is now $\cos(2\pi\nu_o x)+A\cos^2(2\pi\nu_o x)$. A simple graphic calculation shows that the bi-spectrum will be non-zero at 19 points. Thirteen of these points vanish if a bias A/2 is subtracted.

What we may learn from this oceanographic example is the suitability of the triple correlation for testing weak nonlinearities and also for noticing a drift of the bias.

A final question about the triple correlation may be raised, although not answered completely: how much does the triple correlation T(x,y) know about the underlying signal u(x)? Can we reconstruct

u(x) from T(x,y)? What is more useful, the triple correlation or the standard autocorrelation (of second order)? One particular answer is as following: suppose the signal u(x) consists of a signal per se $u_o(x)$ and an isolated peak $\delta(x-x_R)$. If the distance x_R of the peak from the center of $u_o(x)$ is larger than 3/2 of the width of $u_o(x)$, then T(x,y) will consist of several separated islands; some of those islands contain the object per se $u_o(x)$ directly. This rule of recovery happens to be the same as in off-axis Fourier holography, which is a case of second autocorrelation of the total object $u(x) = u_o(x) + (x-x_R)$.

11. Conclusion

I am not pessimistic about the future of optical computing.

ARCHITECTURES FOR A SEQUENTIAL OPTICAL LOGIC PROCESSOR

P. Chavel
Institut d'Optique, Orsay, France

R. Forchheimer
Linkoping University, Linkoping, Sweden

B.K. Jenkins, A.A. Sawchuk, and T.C. Strand
Image Processing Institute-MC 0272
University of Southern California
Los Angeles, California 90089

Abstract

A general technique is described for implementing sequential logic circuits optically. The system consists of a nonlinear transducer which provides a two-dimensional array of gates and one or more computer generated holograms (CGHs) to interconnect the gates. The limitations on the number of gates which can be implemented in an optical system is affected by the interconnection method. We describe three interconnection methods and their respective limitations. One method, which is a hybrid of space-variant and space-invariant CGH elements, provides high gate densities and high gate-utilization rates.

1. Introduction

There has recently been considerable research in optical systems for parallel digital computing with applications in signal processing. The advantages of optical and hybrid optical-electronic systems for high throughput, parallel multi-dimensional processing on signals with large time-bandwidth and space-bandwidth products are well known. Nearly all of these systems to date are basically analog and have severe limitations in accuracy, programmability and flexibility in comparison to electronic digital systems.

Our recent research has concentrated on optical combinatorial and sequential logic systems for parallel digital processing. Some of this work has included parallel A/D conversion [1] and two different implementations of optical combinatorial logic [2], [3]. More recently, we have implemented a parallel optical sequential logic circuit including a clock and a master-slave flip-flop used as a frequency divider [4]. The main components of the sequential logic system are a nonlinear spatial light modulator (SLM) (ideally having a threshold or bistable response function) and a computer generated hologram (CGH) used as a beamsteering element for interconnections. The SLM functions as a two-dimensional array of independent logic gates, and the CGH (or set of them) contains a two-dimensional array of subholograms that interconnect the gates to form a circuit. In the current system the nonlinear element is a Hughes liquid crystal light valve (LCLV) with a 45 degree twisted orientation of the nematic liquid crystal molecules [5]. Although a major limitation of this current SLM is its slow response time (10-100 ms), we feel that recent improvements in both LCLV technology [6] and the exploration of new technologies such as all-optical bistability [7], [8] will significantly improve this. We will not directly consider the question of device speed in this paper.

The main emphasis in this paper is on processor architectures for optical sequential logic. Section 2 of this paper briefly reviews the fundamentals of optical sequential logic. Sections 3 through 5 describe details of CGHs used as interconnection elements. Two basic interconnection methods, space-variant and space-invariant are described. The main limitation on the number of gates is due to space-bandwidth limitations of the CGH and SLM. A hybrid interconnection system having both space-variant and space-invariant elements is described in Section 5, and various types of processors that utilize each type of architecture are described.

2. Fundamentals of 2-D Optical Sequential Logic

In order to implement any logic system, we require two fundamental elements: a nonlinear device to provide the gate function or basic combinatorial operations and an interconnection element (Fig. 1). Furthermore, if we want to provide for sequential logic, the interconnection path must include feedback paths for generating clock signals and for obtaining memory elements. The introduction of feedback and of timing signals makes this work significantly different from previous work with combinatorial logic [2], [3] because the dynamic behavior of the nonlinear device now plays a critical role in the operation of the circuit.

We use the Hughes LCLV as the nonlinear component, although other nonlinear devices could also be used. This device produces a pointwise nonlinear behavior which can to some extent be modified, and in particular can take a shape adequate for our present needs. For example, Fig. 2 depicts a response function for a 45 degree twisted nematic device operated in the backslope

CH1880-4/83/0000/0006$01.00 © 1983 IEEE

mode. We have used the device in this mode to implement the NOR function. If we consider the total input to the device as the sum of two binary inputs, the output will be a binary valued NOR of the inputs. Other binary operations can be performed by altering the characteristic curve of the device [2]. The possible input and output values are indicated in Fig. 2.

The parallelism in the system is evident in the fact that the nonlinearity is applied simultaneously to all points on the device. Thus each resolution element or pixel on the light valve acts as an independent gate. Using resolution figures quoted for current SLMs [5], arrays of 10^5-10^6 pixels can be anticipated.

The remaining problem is how to interconnect the gates. Although several techniques are possible, CGH elements seem to offer the best solution. By using CGH elements in an optical feedback system, the output from any gate can be directed to the input of any other gate or combination of gates. Given that CGH components are to be used for interconnections, there are still a multitude of possible systems for achieving the desired circuit. In the following sections we describe three basic interconnection methods. Naturally, each method offers certain design tradeoffs and limitations. It is the purpose of this paper to examine those tradeoffs and describe how they affect system design.

3. Space-variant Interconnection Method

The most general interconnection system is one in which any gate output can be connected to the input of any gate or combination of gates. If we think of the interconnection scheme as imaging the gate output array plane onto the gate input array plane, this approach represents a space-variant imaging system. The "image" of a gate output consists of a collection of spots (the impulse response of the system for that particular point) which illuminate the appropriate gate inputs, and form the circuit interconnections. Because each object point (gate output) sees a different impulse response (interconnection pattern) this represents a general space-variant system. A space-variant system has been built to demonstrate the concept of sequential optical logic. The demonstration circuit which was implemented comprises a ring oscillator which generates a clock signal and a master-slave flip-flop which is driven by the clock. This system is operational and is described in another paper [4].

A schematic diagram of the optical system used for the space-variant interconnections is shown in Fig. 3. First, the gate outputs are imaged onto the interconnection hologram. This CGH consists of an array of subholograms, one subhologram for each gate. When illuminated by its corresponding gate output, a subhologram will reconstruct an image on the "write" side, or gate input side, of the light valve. The reconstructed images are simple dot patterns, each bright dot illuminating a gate input. Since the

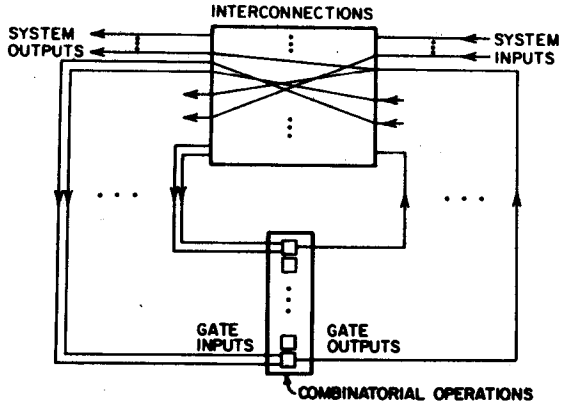


Fig. 1. Functional block diagram of sequential optical logic.

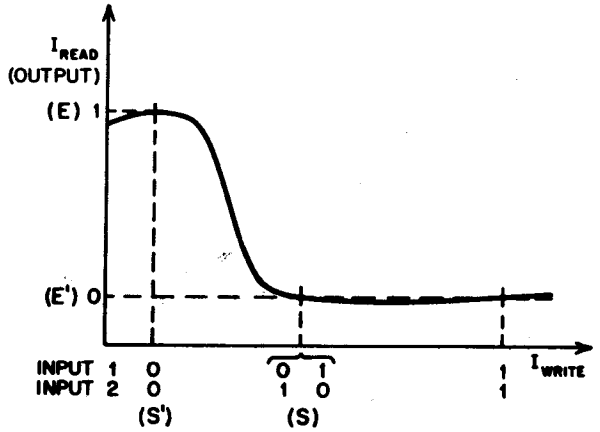


Fig. 2. LCLV input/output characteristic.

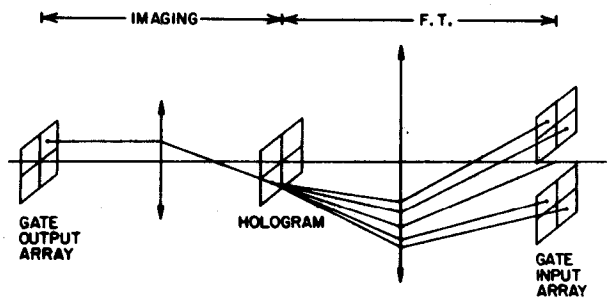


Fig. 3. Space-variant interconnection system. In general the hologram produces multiple diffraction orders, only one of which is used.

reconstructed images can be designed to illuminate any combination of gate inputs, arbitrary interconnections are possible. As shown in Fig. 3, the desired interconnections are formed in one particular diffraction order. Typically, a conjugate image will also be produced, in which case it can be used to probe the system without affecting system operation or to access the system outputs.

While this interconnection scheme allows complete generality, a price is paid in terms of

the space-bandwidth requirements on the CGH. Let there be an NxN array of subholograms on the CGH and an NxN array of gates on the light valve. Each subhologram must have the capability of addressing any of the N^2 gate inputs. The number of addressable points in the reconstruction of a subhologram is equal to the number of complex-valued sample points in the subhologram, assuming a Fourier hologram. Thus the actual space-bandwidth product (SBWP) of each subhologram is

$$S_S = p^2 q^2 N^2 \tag{1}$$

where p^2 is the number of resolution elements in the hologram used to represent one complex-valued sample and q^2 is a factor representing the amount of oversampling in the hologram plane. Generally $p^2>1$ because the complex sample values must be encoded into the hologram, e.g., as real values. Also we generally have $q^2>1$ to avoid crosstalk. These problems are discussed further below.

The entire interconnection hologram consists of N^2 subholograms, one for each gate. Thus the total SBWP of the hologram is S_T where

$$S_T = N^2 S_S = p^2 q^2 N^4. \tag{2}$$

Because $S_T \propto N^4$, we expect that the hologram SBWP, S_T, will quickly become the limiting factor as N increases. We will verify that below, but first we need to study the crosstalk in the gate-input plane to get a feeling for the expected values of q^2.

The crosstalk can be represented by α, the ratio between gate inputs of the worst-case (largest) "zero" value, ℓ_0, and the worst-case (smallest) "one" value, ℓ_1:

$$\alpha = \ell_0/\ell_1. \tag{3}$$

We require $\alpha<1$ in order to distinguish all possible zero and one states. Assume that the intensity profile of a single gate input reconstructed from a subhologram is F(x,y). (The reconstruction of a subhologram can be represented by a set of Dirac delta functions (one for each addressed gate input) convolved with the Fourier transform, W(x,y), of the aperture function of the subhologram. Then $F(x,y) = |W(x,y)|^2$.) Thus the worst-case "one" value is the integral of F over the defined area, a, of the gate input

$$\ell_1 = \iint_a F(x,y)\ dxdy \tag{4}$$

The worst-case zero level occurs when all gates have their maximum input levels except for the gate in question which has a zero input level. If each gate has m inputs (a fan-in of m) and gates are contiguous, then the worst-case zero can be shown to be

$$\ell_0 = m\left[\iint_A F(x,y)\ dxdy - \iint_a F(x,y)dxdy\right] \tag{5}$$

where A is the area of the entire gate array. This assumes that the spatially integrated sum of gate inputs from different subholograms is effectively an incoherent sum. Note that such inputs may actually add coherently in which case they produce interference fringes. Spatial integration over these fringes results in an effective incoherent summation. Similarly, the above equation assumes that the gate inputs reconstructed from a single subhologram also can be modeled as adding incoherently. This will be essentially true if a pseudorandom phase is applied to the m distinct gate inputs. Combining the above equations we get

$$\alpha = m(1/E_0 - 1) \tag{6}$$

where E_0 is the fraction of the single-gate input intensity profile which falls within the defined area of that gate

$$E_0 = \iint_a F(x,y)\ dxdy \Big/ \iint_A F(x,y)\ dxdy \tag{7}$$

If we define a_0 as the minimum gate area determined by the Nyquist theorem and the subhologram area, then

$$a = q^2 a_0 \tag{8}$$

with q^2 being the oversampling factor as defined above. Obviously, as q increases, E_0 approaches unity and α approaches zero.

We now consider an example. If we use a triangle function in x and in y for the aperture (window) function of each subhologram, then its Fourier transform is a two-dimensional $sinc^2$ function, and

$$F(x,y) = sinc^4(x/2)sinc^4(y/2) \tag{9}$$

Choosing q=2 yields a crosstalk $\alpha=(0.11)m$, so that 3-input gates cause a crosstalk of 0.33. Thus for NOR gates the thresholding of the nonlinear device can be performed anywhere between relative input levels of 0.33 and 1.0. In this case the sampling rate is twice the Nyquist rate. Increasing q permits a larger fan-in, e.g., q=3 implies $\alpha=(.0082)m$ and a crosstalk of 0.33 permits 40-input gates to be used. Also note that the use of a more appropriate aperture function could permit smaller values of q.

In order to estimate the SBWP that can be written onto a CGH, we assume the CGH is written using electron-beam lithography, as was the case for the experimental demonstration of the optical logic system [4]. This electron-beam system has written linewidths down to 0.5 μm, and has a maximum file size of 1.024 mm on a side. Files can be stitched together to yield a maximum size of 10 cm on a side. If we minimize the stitch error by making the file boundaries coincident with subhologram boundaries, a SBWP of 4×10^{10} is attainable.

The hologram coding parameter p, defined in Eq. 1, for the case of a Burckhardt hologram [9], has a minimum value of 3, assuming square cells. Having found that q will typically be in the range of 2-3, we conclude that the maximum feasible number of gates corresponds to a value of pq on the order of 10. From Eq. 2 and the above SBWP, we find that the gate array dimension is NxN where

$$N \approx 100\text{-}200 \qquad (10)$$

for space-variant interconnections. Because this is less than the SBWP capabilities of some spatial light modulators, the CGH is the limiting element.

Since the space-variant system allows arbitrary interconnections, the only other possible limitation on the circuits that can be implemented is the requirement that all gates must perform the same binary operation, e.g., NOR in this case. However, since all the Boolean operations may be constructed out of NOR gates, this does not limit the types of processing operations that can be performed. Another feature is that circuits with any degree of inherent parallelism, or lack thereof, can be implemented with approximately equal ease.

4. Space-invariant Interconnection Method

If one is willing to compromise on the arbitrariness of the gate interconnections, a substantial increase in the possible number of gates results. The extreme case is a totally space-invariant interconnection. This is the idea behind the processor suggested by Huang [10]. Here we extend this concept to include sequential circuits. This interconnection method is implemented optically by an imaging system with a space-invariant filter, using one simple hologram for the entire circuit (Fig. 4). The filter has an impulse response consisting of a series of spots which illuminate the appropriate gate inputs as in the space-variant case. However, in this case, the impulse response (interconnection pattern) is the same for every gate output, and the gate inputs are addressed relative to the position of the gate output. The space-variant method worked on the basis of absolute addressing.

An example of a space-invariant interconnection pattern is shown in Fig. 5. Each dot in the figure represents a (NOR) gate, and each arrow represents an interconnection from the output of one gate (dot) to the input of another. Each gate is considered to have one additional, unconnected input for an enable/disable signal. A particular circuit is implemented by disabling the appropriate gates. In the NOR case, a gate is disabled by projecting light onto it (i.e., putting a 1 onto the unconnected input). With the illustrated interconnection pattern it is possible to transfer data in various directions without getting unintended feedback loops. The major limitation of this interconnection method is that the implementation of many circuits will require a large number of gates to be disabled. Obviously, circuits with very regular interconnections can utilize the gates more efficiently than irregular curcuits.

Since the holographic element used in this interconnection system is simple, a very large number of gates can be interconnected. Even allowing the PSF to simultaneously address any set of points in the array, the SBWP required of the hologram is of order $p^2q^2N^2$ (see Eq. 1 above). Thus if the full SBWP available with the CGH could be exploited in this system and if $p^2q^2 \sim 100$ approximately 4×10^8 gates could be interconnected. The hologram for this system could also be recorded optically. In either case, the number of gates with the space-invariant interconnection method is limited by the SBWP of the spatial light modulator.

As mentioned above, the method of disabling gates to implement circuits decreases the number of gates that are actually used, and therefore severely restricts the types of operations that can be performed efficiently. It also adds a degree of complexity to the system. However, this method of optically disabling gates also provides a potential advantage - it provides a means of easily "re-wiring" the system in real time by changing the disable signals. This could offer considerable flexibility in making an adaptive system.

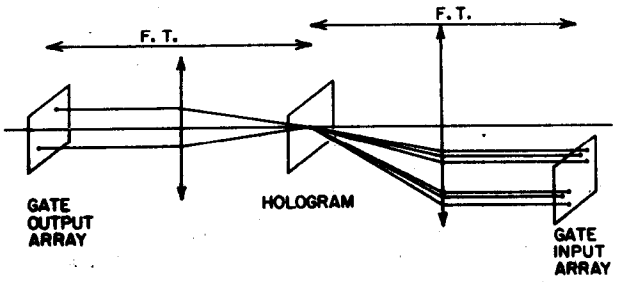


Fig. 4. Space-invariant interconnection system.

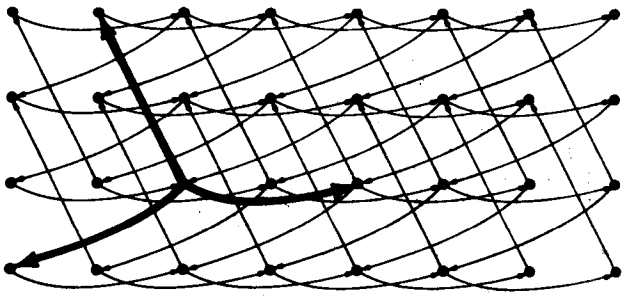


Fig. 5. An example of a space-invariant interconnection network. Nodes represent gates and arrows are the (optical) interconnections.

5. Hybrid Interconnection Method

At this point we have seen two approaches to interconnecting gates. In the space-invariant case there is only one interconnection pattern which is applied to all gates whereas in the

space-variant case the number of distinct interconnection patterns is in general equal to the number of gates. These two approaches represent the extreme cases in terms of the space-bandwidth requirements they place upon the CGH element. The tradeoffs between these two is increased flexibility at the cost of increased hologram complexity. Since the space-invariant case generally suffers from inefficient gate utilization and the space-variant system is limited by the hologram to the number of gates it can address, it is worthwhile considering if there is a combination of techniques which can achieve high gate utilization efficiency and at the same time be limited in gate count only by the space-bandwidth product limitations of the spatial light modulator.

Our approach to this has been to consider a hybrid system which combines space-variant and space-invariant interconnections. The idea is to define a finite number, M, of distinct interconnection patterns. We then assemble our circuit using only these M interconnection patterns. If the total number of gates is N^2 we assume

$$1<<M<<N^2 \tag{11}$$

so that this system is truly intermediate between the space-variant and space-invariant cases. If M is large, we anticipate that we have almost complete flexibility in designing our circuit.

The optical implementation of this system is schematically diagrammed in Fig. 6. Here the gate output array is imaged onto a space-variant filter element as in Fig. 3. The purpose of this element is to deflect the light from each gate output through one of M subholograms in the second CGH element (Fig. 6). These subholograms act as space-invariant filter elements which produce the M different interconnection patterns in the gate input plane.

Although the space-variant element would appear to have the same space-bandwidth limitations as in the simple space-variant case, we note that the SBWP of each subhologram in this plane is now of order M rather than of order N^2. Thus the total SBWP requirement in this element is much less than in the previous space-variant case. The holograms in the space-invariant element generally have a relatively low SBWP.

The SBWP, S_{S_1}, of a subhologram in the first hologram, H_1, is

$$S_{S_1} = p_1^2 q_1^2 M \tag{12}$$

where M is the number of subholograms in the second (space-invariant) hologram, H_2, and p_1^2 and q_1^2 represent coding and oversampling factors, respectively, as in the space-variant interconnection section. Similarly, the SBWP of a subhologram of H_2 is, in the worst case,

$$S_{S_2} = p_2^2 q_2^2 N^2 \tag{13}$$

where N^2=number of gates. This worst case allows the H_2 subhologram to address any n gates in the array. If the gates it addresses are localized, i.e., are all contained in the same portion of the array, then its SBWP can be significantly reduced by the introduction of a carrier frequency [16]. Since H_1 consists of N^2 subholograms and H_2 consists of M subholograms, their total SBWPs are given by

$$S_{T_1} = p_1^2 q_1^2 MN^2 \qquad S_{T_2} = p_2^2 q_2^2 MN^2 \tag{14}$$

and here, again, S_{T_2} is a worst-case estimate. If we assume both holograms are written in the same manner, then $S_{T_1} = S_{T_2} = S$ and $p_1 = p_2 = p$, from which it follows that $q_1 = q_2 = q$, thus

$$S_T = p^2 q^2 MN^2. \tag{15}$$

As in the space-variant case, we need to analyze the crosstalk in order to estimate q.

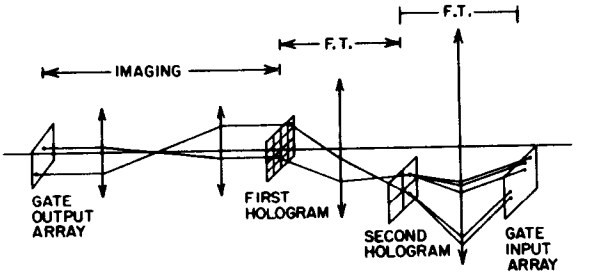


Fig. 6. Hybrid interconnection system. The first holgoram is a space-variant element as in Fig. 3. The second element is an array of space-invariant filters.

For the hybrid interconnection scheme, two sources of crosstalk exist. Inter-pixel crosstalk occurs between pixels in the gate-input plane, and is analogous to the crosstalk treated in the space-variant interconnections section. Inter-hologram crosstalk occurs between subholograms in the second hologram and also contributes to noise in the gate-input plane. (We assume neglible crosstalk at the first hologram because it is in the image plane of the gate output array.)

The inter-pixel crosstalk is completely analogous to the crosstalk in the space-variant interconnection case when applied to H_2 and the gate-input array, and the same equations apply.

In order to analyze the inter-hologram crosstalk, we have to find the effect of this crosstalk in the gate input plane. We assume that each H_1 subhologram addresses only one H_2 subhologram. Through a given H_2 subhologram k, there are only n_k subholograms of H_1 that can address a given gate input ρ, where n_k is the fan-out of subhologram k. Any unintentional illumination of k from one of these H_1 subholograms will contribute crosstalk to gate ρ.