INCOHERENT OPTICAL MATRIX-MATRIX MULTIPLIER

A.R. Dias
Radar and Optics Division
Environmental Research Institute of Michigan
P.O. Box 8618, Ann Arbor, Michigan 48017

INTRODUCTION

In recent years a growing interest has developed in incoherent optical processing (ref. 1). As the invention of the laser was an important force in coherent optics, the advent of fast and very compact solid state light sources is a major force behind this interest.

Incoherent optical processing has indeed several appealing characteristics, namely higher immunity to optical noise and reduced dynamic range requirements than coherent processing as well as an inherent versatility of input format (serial or parallel).

A very fast fully parallel incoherent optical multiplier was developed at Stanford (refs. 2 and 3) and this device has been shown to be capable of performing complex-valued arithmetic on vector formatted data arriving in parallel to the processor inputs.

There is a considerable number of applications for processors of this sort however there is a need to extend their capability for two-dimensional parallel data. Indeed one can understand the interest in speeding real time processing capabilities in the context of reconnaissance sensors, two-dimensional beamforming, etc.

This work addresses the latter case. We propose an extension of the concepts developed earlier to the new generation of matrix-matrix multipliers using incoherent light. This new technology has the potential of addressing problems involving two-dimensional mathematical transforms, two-dimensional pattern recognition, and high-speed processing of synthetic aperture radar data among others.

SYSTEM ARCHITECTURE

Let us then consider the matrix multiplication

$$C = A B \tag{1}$$

where A is an M x K matrix, and B and C are K x N and M x N matrices respectively.

The element c_{mn} in the m-th row and n-th column is given in terms of the elements g_{mk} of A and b_{kn} of B by

$$c_{mn} = \sum_{k=0}^{K-1} a_{mk} b_{kn} \tag{2}$$

A matrix can, of course, be viewed as an ordered collection of column vectors.

If we define

$$\underline{B} = \sum_{n=0}^{N-1} \underline{B}_n \quad \text{where} \quad \underline{B}_n = \begin{bmatrix} 0 & 0 & \cdots & b_{0n} & \cdots & 0 \\ 0 & 0 & \cdots & b_{1n} & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ 0 & \cdots & & b_{K-1,n} & \cdots & 0 \end{bmatrix} \quad (3a)$$

and

$$\underline{C} = \sum_{n=0}^{N-1} \underline{C}_n \quad \text{where} \quad \underline{C}_n = \begin{bmatrix} 0 & 0 & \cdots & c_{0n} & \cdots & 0 \\ 0 & 0 & \cdots & c_{1n} & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ 0 & \cdots & & c_{M-1,n} & \cdots & 0 \end{bmatrix} \quad (3b)$$

the matrix product $\underline{C} = \underline{A}\,\underline{B}$ becomes the sum of N matrices $\underline{C}_n$ whose only non-zero elements are in its n-th column. In consequence, the matrix product $\underline{C}$ can be seen as an ordered collection of column vectors each of which is the product of $\underline{A}$ by the corresponding column vector of $\underline{B}$. A matrix-matrix multiplication can then be expressed as a series of matrix-vector multiplications. The technology developed for the incoherent optical matrix-vector multiplier (ref. 2) can then be extended to this new application.

The concept here is to combine N matrix-vector multipliers in a single processor. The physical realization of this architecture is depicted in Figure 1.

The elements of matrix $\underline{B}^*$ are entered as irradiance values of a two-dimensional incoherent array that may be realized using narrow-band light-emitting diodes (LED), laser diodes, or even an array of polished optical fiber ends. The elements of matrix $\underline{A}$ are encoded as transmission values of an optical transparency or as reflected values of an optically reflective device. In either case, its implementation can be fixed in time (e.g., a film transparency) or it can vary with time (e.g., a real time optical modulator).

The first block of optics (figure 1) images horizontally the input array $\underline{B}$ onto the optical mask $\underline{A}$ and spreads vertically the light of each light source along the corresponding vertical column. Figure 2 illustrates the illumination produced by the central column of a 3 x 3 input matrix on a 3 x 3 mask. The light transmitted through the mask encodes all the possible inner products of the column vectors of the input matrix by all the row vectors of the stored matrix.

The second block of optics (figure 1) integrates these inner products and focuses the light on the proper output array locations -- imaging vertically and focusing horizontally onto the vertical column corresponding to the particular

---

*At this stage, we assume that the elements of $\underline{A}$ and $\underline{B}$ are real and non-negative. Later, we will remove this restriction.

horizontal row at the input. This is a demanding operation that can be more clearly visualized in the case of a 3 x 3 matrix. We illustrate in figure 3 how we obtain the outputs (1, 1) and (3, 2). The output (3, 2) can readily be seen as the inner product of the 3-rd row of matrix A and the 2-nd column of matrix B as would be expected. Finally, figure 4 shows the light path through a particular cell in the stored mask. The element a_{mk} (the central element in this example) maps the elements of the k-th row of input B onto the elements of the m-th row of output C.

OPTICAL IMPLEMENTATION

The optics block labeled "Optics I" in figure 1 has the same requirements as the mask illumination of the incoherent optical matrix-vector multiplier (ref. 3). In the latter case, a discrete set of classical discrete optical elements in one version and an array of multimode planar waveguides in another version were used. In the present case, a very similar form of the classical discrete optical components approach mentioned could certainly be used.

The second block of optics -- Optics II -- has more demanding requirements. In effect, it must steer the light transmitted by each cell in the stored mask to all N elements in the corresponding row of the output detector array (figure 4). Holographic optical elements can be used as the steering beam devices in this type of processor. An example of the application of this technology was shown by S. Case and his co-workers in the paper "Multifacet Holographic Optical Elements," presented at the Optical Society of America Annual Meeting in October 1980 (Abstract in J.O.S.A., 12, 1980).

In this context, we suggest the geometry depicted in figure 5 where each narrow-band light source and each cell of the stored mask are located behind a holographic optical element (HOE). This is a well established technology with the potential for a highly efficient light management.

COMPLEX-VALUED ARITHMETIC

It is, of course, important for a processor to have the capability to perform complex-valued operations. The incoherent optical processor we report here manipulates light intensities which are real-valued and non-negative in nature. Methods for encoding complex numbers similar to those used in the incoherent optical matrix-vector multiplier (ref. 4) can be applied here.

The complex-valued arithmetic capability built in this processor is an immediate extension of the schemes utilized before, and for completeness, we will briefly elaborate two possible methods.

A. Three Vector Decomposition

Each complex quantity is decomposed uniquely along the three phasors $1 \cdot \exp[j0]$, $1 \cdot \exp[j2\pi/3]$ and $1 \cdot \exp[j4\pi/3]$. There is more than one way to perform this decomposition (ref. 4, Appendix A) but we do not need to dwell on it here.

The three matrices can be decomposed according to

$$\underline{A} = \underline{A_0}\, e^{j0} + \underline{A_1}\, e^{j\frac{2\pi}{3}} + \underline{A_2}\, e^{j\frac{4\pi}{3}}$$

$$\underline{B} = \underline{B_0}\, e^{j0} + \underline{B_1}\, e^{j\frac{2\pi}{3}} + \underline{B_2}\, e^{j\frac{4\pi}{3}} \tag{4}$$

$$\underline{C} = \underline{C_0}\, e^{j0} + \underline{C_1}\, e^{j\frac{2\pi}{3}} + \underline{C_2}\, e^{j\frac{4\pi}{3}}$$

where the elements of the different matrix components are all real-valued non-negative quantities.

The matrix-matrix multiplication $\underline{C} = \underline{A}\,\underline{B}$ can now be expressed in terms of these quantities as

$$\begin{bmatrix} \underline{C_0} \\ \underline{C_1} \\ \underline{C_2} \end{bmatrix} = \begin{bmatrix} \underline{A_0} & \underline{A_2} & \underline{A_1} \\ \underline{A_1} & \underline{A_0} & \underline{A_2} \\ \underline{A_2} & \underline{A_1} & \underline{A_0} \end{bmatrix} \begin{bmatrix} \underline{B_0} \\ \underline{B_1} \\ \underline{B_2} \end{bmatrix} \tag{5}$$

from which the output complex-valued matrix $\underline{C}$ can be obtained after the proper re-combination of its real valued and non-negative components $\underline{C_0}$, $\underline{C_1}$ and $\underline{C_2}$.

The processor's configuration is shown in figure 6, where we have omitted the optics. The size of the input and output arrays increases by a factor of 3 and the stored mask by a factor of 9.

B. Biased Real and Imaginary Parts Decomposition

Another encoding scheme decomposes each complex quantity into its real and imaginary parts and adds to these a bias term.

$\underline{A}$ and $\underline{B}$ can then be decomposed into their real and imaginary parts.

$$\underline{A} = \underline{A_R} + j\underline{A_I}$$

$$\underline{B} = \underline{B_R} + j\underline{B_I} \tag{6}$$

and we can express $\underline{C}$ as

$$\begin{bmatrix} \underline{C_R} \\ \underline{C_I} \end{bmatrix} = \begin{bmatrix} \underline{A_R} & -\underline{A_I} \\ \underline{A_I} & \underline{A_R} \end{bmatrix} \begin{bmatrix} \underline{B_R} \\ \underline{B_I} \end{bmatrix} \tag{7}$$

Assuming known the dynamic range of the elements of the real and imaginary parts of $\underline{A}$ and $\underline{B}$, we assign a bias matrix $\underline{B_A}$ to matrix $\underline{A}$ and a bias matrix $\underline{B_B}$ to matrix $\underline{B}$.

These bias matrices have constant elements. All elements in the same column are identical, but they can vary from column to column.

If now we configure the system as it is shown in figure 7, the detected outputs $\underline{C_R}$ and $\underline{C_I}$ are given by

$$\underline{C_R'} = \underline{A_R}\underline{B_R} - \underline{A_I}\underline{B_I} + \left[\underline{A_R}\underline{B_B} - \underline{A_I}\underline{B_B}\right]$$
$$+ \underline{B_A}\left[\underline{B_R} + \underline{B_B} + \underline{B_I} + \underline{B_B}\right] \tag{8a}$$

$$\underline{C_I'} = \underline{A_I}\underline{B_R} + \underline{A_R}\underline{B_I} + \left[\underline{A_I}\underline{B_B} + \underline{A_R}\underline{B_B}\right]$$
$$+ \underline{B_A}\left[\underline{B_R} + \underline{B_B} + \underline{B_I} + \underline{B_B}\right] \tag{8b}$$

The real and imaginary parts $\underline{C_R}$ and $\underline{C_I}$ of the product $\underline{C}$ (given by equation (4)) can be obtained from equations (8a) and (8b)

$$\underline{C_R} = \underline{C_R'} - \left[\underline{A_R}\underline{B_B} - \underline{A_I}\underline{B_B}\right] - \underline{B_A}\left[\underline{B_R} + \underline{B_B} + \underline{B_I} + \underline{B_B}\right] \tag{9a}$$

$$\underline{C_I} = \underline{C_I'} - \left[\underline{A_I}\underline{B_B} + \underline{A_R}\underline{B_B}\right] - \underline{B_A}\left[\underline{B_R} + \underline{B_B} + \underline{B_I} + \underline{B_B}\right] \tag{9b}$$

In each equation, the first bracket is known a priori and it can be electronically represented by a fixed voltage. The second bracket can be measured, adding an extra transparent row to the stored mask. Hence, with proper and straightforward post-detection electronics, one obtains the desired product real and imaginary parts.

SYSTEM PERFORMANCE ESTIMATE

At this stage of the system development, we address ourselves only to the major performance parameter -- the data throughput.

We can reasonably postulate that it will be feasible to build matrix-matrix multipliers with complex-valued matrices with sizes larger than 100 x 100. Without being overly optimistic, we can envision the electronics side of such processors being driven at clock rates greater than 100 MHz. This implies data throughput rates exceeding 10^{12} complex-valued samples per second. This is at least two degrees of magnitude higher than throughput rates obtained with matrix-vector multipliers.

APPLICATIONS

The parallel processing character of this processor is an essential feature that is extremely useful when a large number of input channels are present simultaneously.

A large number of problems dealing with matrix multiplication can be implemented with this processor. We may refer problems involving Fourier transforms (beam

nulling, beamforming), image processing problems (pattern recognition, filtering), and signal identification (signature reconnaissance), to name just a few.

In this work, we will single out two fundamental mathematical applications that may be implemented using this technique.

A. Two-Dimensional Transforms

Let us consider a two-dimensional array i(k, 1) assuming values in the rectangular range k = 0, 1, . . ., N - 1, 1 = 0.1, . . ., N - 1. A two-dimensional transform over this array produces a second array o(m, n) given by

$$o(m, n) = \sum_{k=0}^{N-1} \sum_{1=0}^{M-1} t(m, n, k, 1)\, i(k, 1) \qquad (10)$$

$$m = 0, 1, \ldots, N - 1$$

$$n = 0, 1, \ldots, M - 1$$

where t(m, n, k, 1) is the transform kernel.

We assume the kernel t(m, n, k, 1) is separable

$$t(m, n, k, 1) = v(m, k)\, h(n, 1). \qquad (11)$$

We can then express o(m, n) as

$$o(m, n) = \sum_{k=0}^{N-1} \sum_{1=0}^{M-1} v(m, k)\, h(n, 1)\, i(k, 1) \qquad (12)$$

and factorizing h(n, 1) and performing first the sum over k, we obtain

$$o(m, n) = \sum_{1=0}^{N-1} h(n, 1) \sum_{k=0}^{N-1} v(m, k) i(k, 1) \qquad (13)$$

The sum over k is a matrix product (see equation (2)) producing the intermediate result

$$s(m, 1) = \sum_{k=0}^{N-1} v(m, k) i(k, 1) \qquad (14)$$

which can be written in matrix notation

$$S = V I \qquad (15)$$

where I is the input matrix, V is the transform kernel matrix along the columns and S is the intermediate matrix. The output can then be expressed as

76

$$o(m, n) = \sum_{1=0}^{M-1} h(n, 1)s(m, 1) \qquad (16)$$

or in matrix notation

$$\underline{O} = \left[\underline{H} \ (\underline{V} \ \underline{I})^T \right]^T \qquad (17)$$

where $\underline{H}$ is the transform kernel matrix along the rows and $\underline{O}$ is the output matrix.

Equation (17) shows that two matrix multiplications in sequence can implement a two-dimensional transform (as given by equation (10)) when the transform kernel is separable.

The immediate application that comes to mind is the two-dimensional discrete Fourier transform (DFT). The discrete Fourier transform kernel is given by

$$t(m, n, k, 1) = \exp \ -j \ \frac{2\pi mk}{N} \ \exp \ -j \ \frac{2\pi n1}{M} \qquad (18)$$

$$v(m, k) = \exp \ -j \ \frac{2\pi mk}{N} \qquad (19a)$$

$$h(n, 1) = \exp \ -j \ \frac{2\pi n1}{M} \qquad (19b)$$

Figure 8 illustrates a basic configuration for a two-dimensional DFT implemented with two matrix-matrix multipliers.

B. Iterative Processing

Another significant application of incoherent optical matrix-matrix multipliers is in the area of iterative processing. Figure 9 shows a basic configuration.

There are several possible uses for iterative techniques. Important examples are the phase retrieval problem related to imaging through turbulence (ref. 5) and the implementation of some numerical methods for matrix inversion.

This work was performed under an ERIM internal research program.

REFERENCES

1. Monahan, M.A., K. Bromley and R.P. Bocker, "Incoherent Optical Correlators, Proc. IEEE, 65, 121 (1977).

2. Goodman, J.W., A.R. Dias and L.M. Woody, "Fully Parallel, High Speed Incoherent Optical Method for Performing Discrete Fourier Transforms," Optics Letters, 1 (1978).

3. Goodman, J.W., A.R. Dias, L.M. Woody and J. Erickson, "Application of Optical Communications Technology to Optical Information Processing," Proc. SPIE 190, LASL Conf. on Optics 79 (1979).

4. Dias, Antonio R., "Incoherent Matrix-Vector Multiplication for High Speed Data Processing" Stanford University Ph.D. Thesis, June 1980, University Microfilms, 80-24641, Ann Arbor, MI. Also published as Stanford Electronics Laboratories, Tech. Report No. L722-4 (June 1980).

5. Fienup, J.R., "Space Object Imaging Through the Turbulent Atmosphere," Optical Eng., 18-5, 529, 1979.
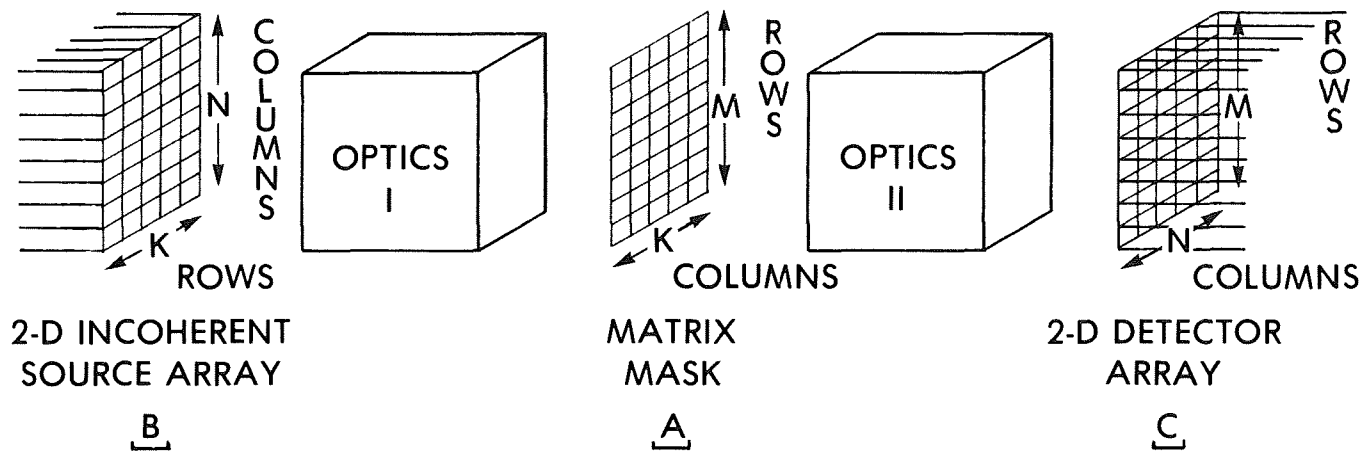
Figure 1.- Basic block form of the incoherent optical matrix-matrix multiplier.
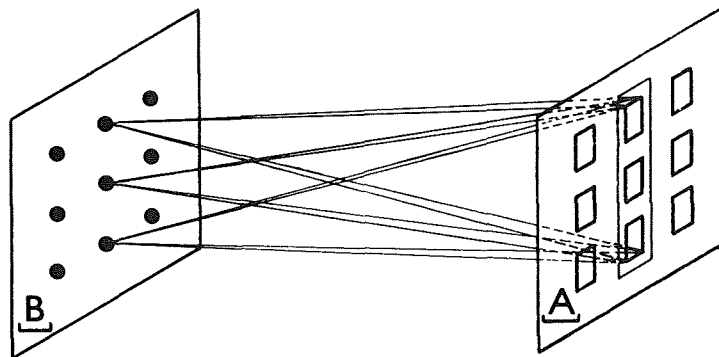
Figure 2.- Example of B to A illumination. For simplicity sake we will refer
to matrix A as the stored mask, though it may be variable with time.
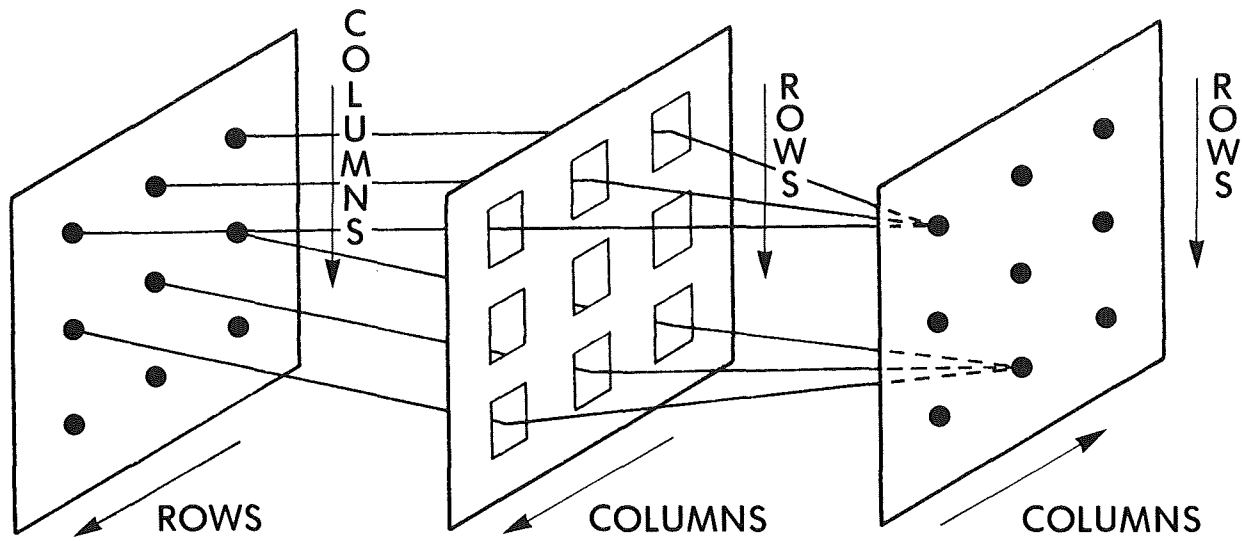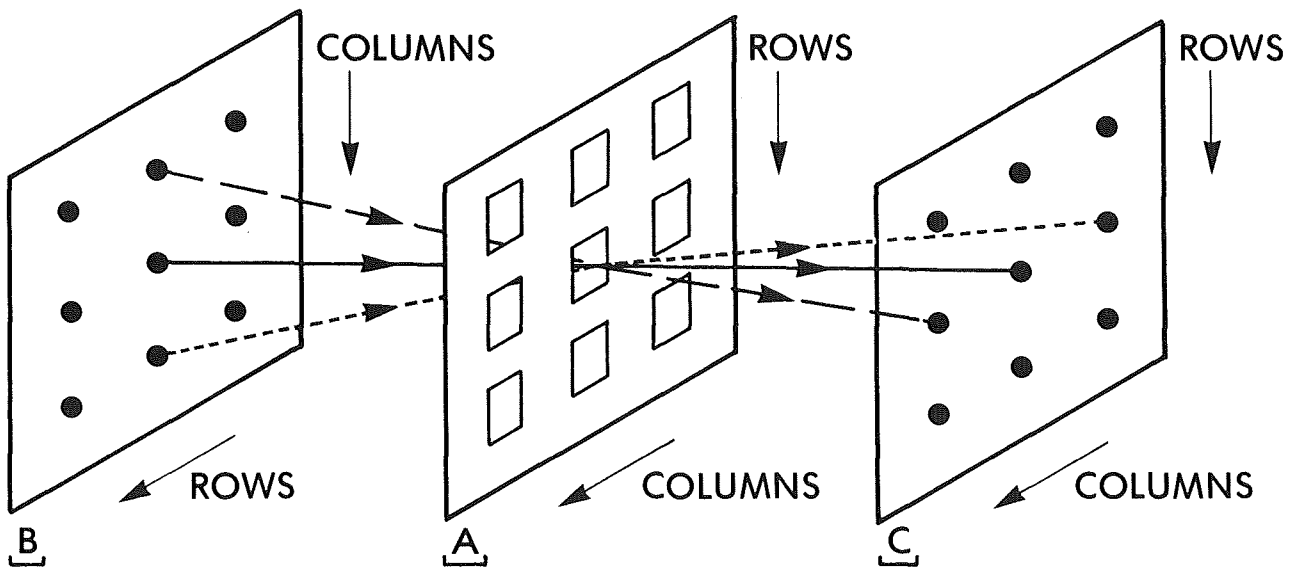
Figure 3.- Example of _A_ to _C_ illumination.

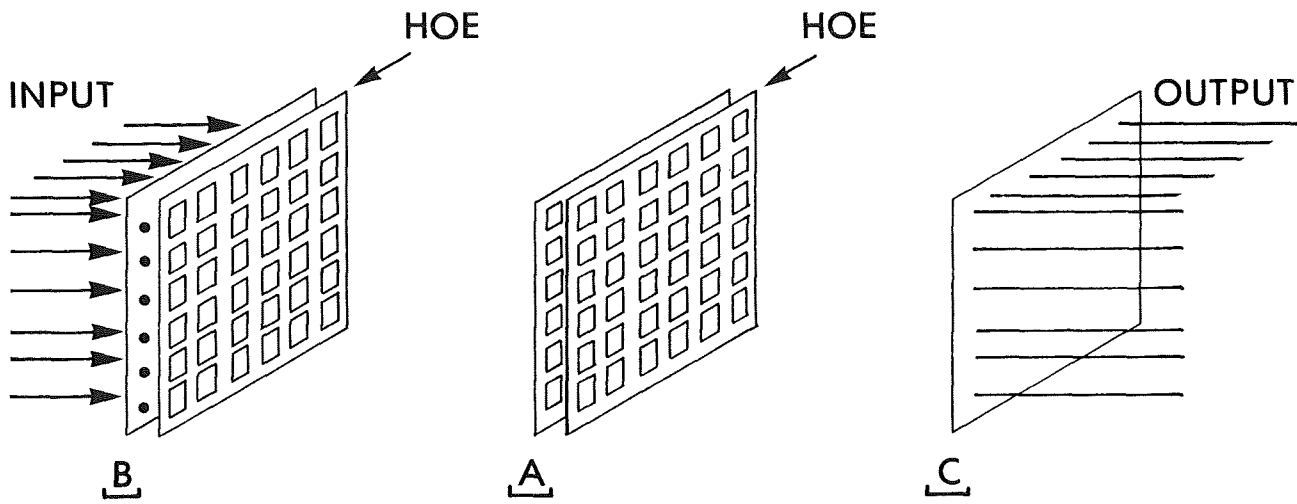Figure 4.- Example of light path through an individual cell in the stored mask.

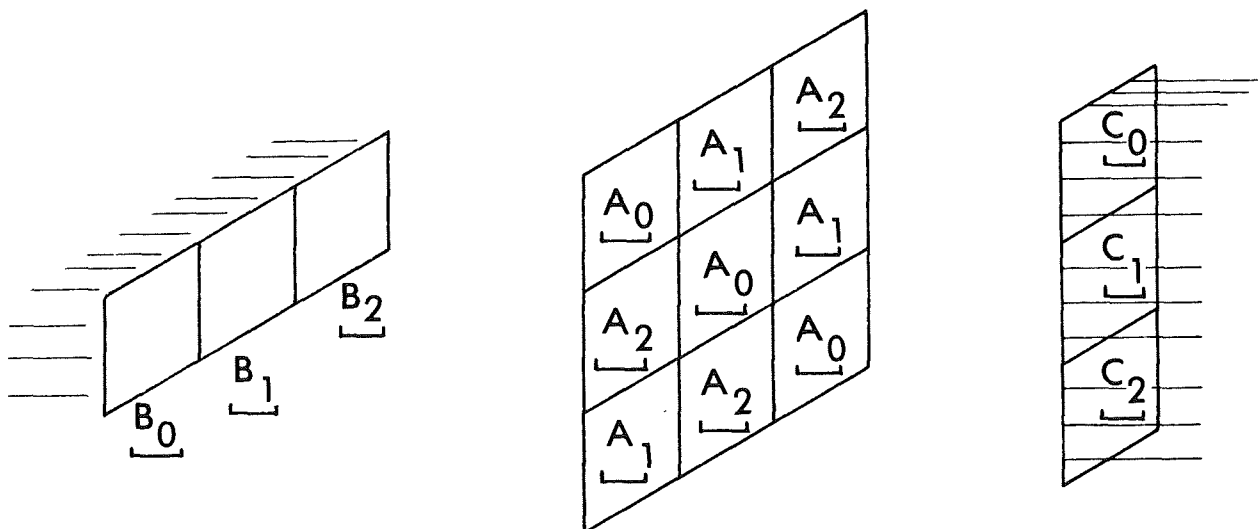Figure 5.- Matrix-matrix multiplier. Implementation with holographic
optical elements.

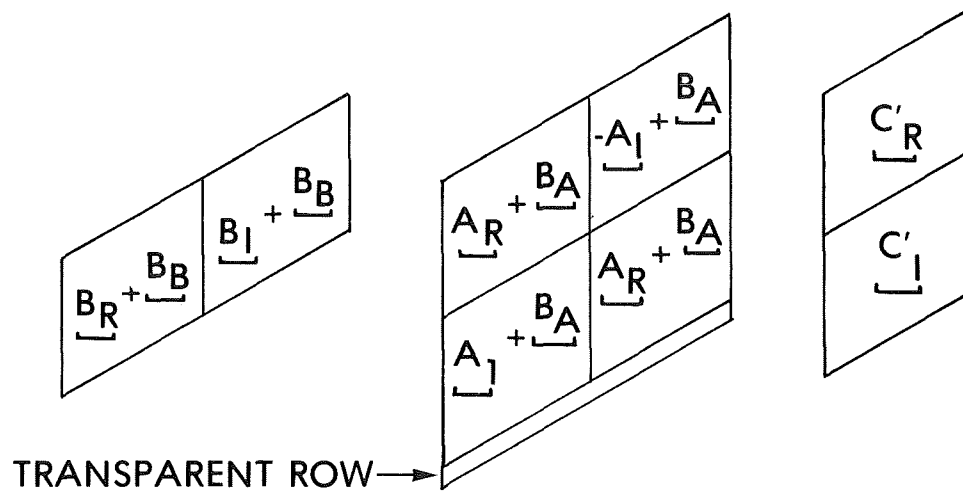Figure 6.- Three-vector decomposition configuration.

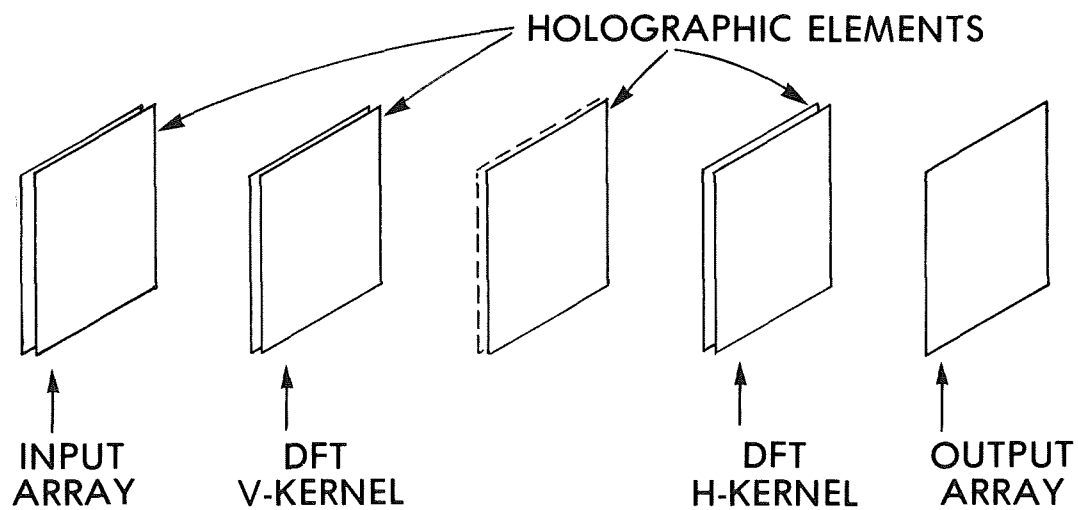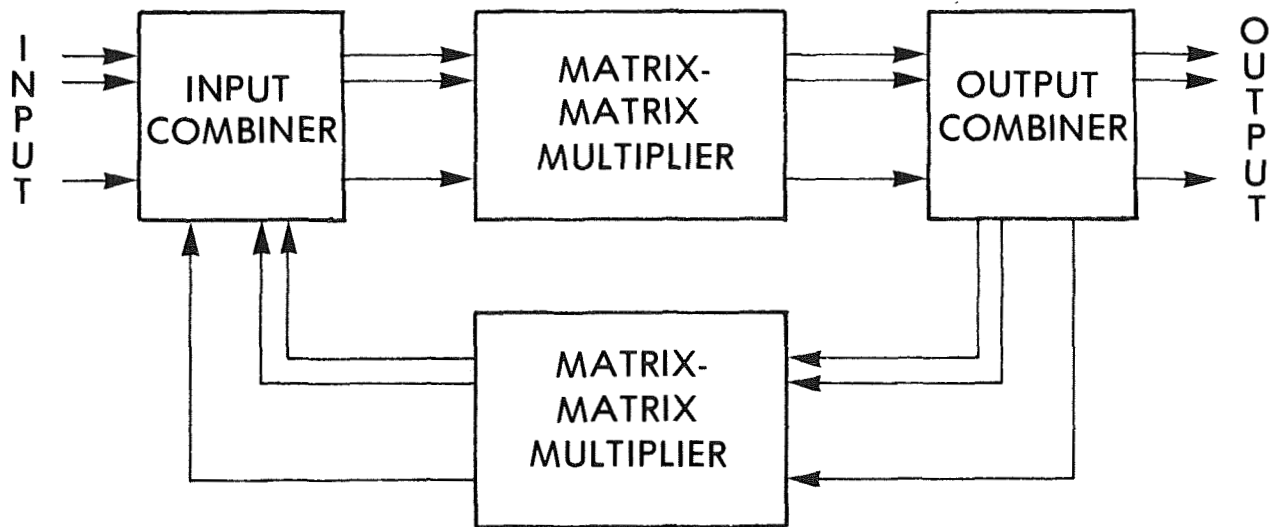Figure 7.- Biased real and imaginary parts decomposition configuration.

Figure 8.- 2D-FFT.

Figure 9.- Iterative processor.