Matrix–matrix multiplication by using anisotropic self-diffraction in $BaTiO_3$

Ching-Cherng Sun, Ming-Wen Chang, and Ken Yuh Hsu

We propose a matrix–matrix multiplication by using anisotropic self-diffraction in $BaTiO_3$. The input matrices are carried by the two incident beams with special Bragg-matched incident angles. The output matrices are produced by anisotropic self-diffraction with the polarization orthogonal to those of the incident matrices. By thresholding the output this architecture is particularly suitable for the applications of optical interconnects and an optical switch.

Introduction

Optical computing has been extensively studied and is getting more attraction because of the inherent parallelism of optics. Optical vector–matrix multiplication is useful for applications of signal processing including neural networks and large-scale interconnections.[1] Recently, real-time dynamic recording properties of photorefractive materials have been used to perform image processing and multiwave mixing.[2] Several architectures for parallel vector–matrix multiplications and programmable interconnections have been proposed by using photorefractive crystals.[3] In general we achieve matrix–matrix multiplication by decomposing it into a series of vector–matrix multiplications.[4] Thus, for an $\mathcal{N} \times \mathcal{N}$ matrix, the vector–matrix multiplication needs N-time operations to complete the matrix–matrix multiplication. In the past several years, direct matrix–matrix multiplications have been proposed by using degenerate or nondegenerate four-wave mixing in bulk photorefractive crystals.[3,5,6] In this paper we propose a new method for performing matrix–matrix multiplication by using photorefractive anisotropic self-diffraction in $BaTiO_3$. The characteristics of this architecture are that the whole system needs only two incident beams for carrying the input matrices. The multiplication output is obtained by self-diffraction from a $BaTiO_3$ crystal, and no extra reading beam is needed. Because of the nonlinear property of wave mixing in photorefractive crystals, the best results are obtained when the outputs are thresholded to become binary types. This architecture is useful in particular for performing two-dimensional to two-dimensional interconnections.

K. Y. Hsu is with the Institute of Electro-Optical Engineering, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu, Taiwan 30050, China. The other authors are with the Institute of Optical Sciences, National Central University, Chung-li Taiwan, China.

Received 18 March 1993; revised manuscript received 29 November 1993.

0003-6935/94/204501-07$06.00/0.

Anisotropic Self-Diffraction in $BaTiO_3$

Photorefractive crystal is one kind of real-time dynamic holographic recording material. When the photogenerated mobile charges are redistribued inside the crystal corresponding to the light fringes, a space-charge field is built and the local refractive index is changed through the Pockels effect. As a result, a phase volume hologram is formed.[7] In anisotropic bulk crystal, the Bragg condition determines the diffraction conditions. As shown in Fig. 1, the optic axis is in the direction perpendicular to the plane of incidence, $\mathbf{k}_{e1}$ and $\mathbf{k}_{e2}$ are the two incident wave vectors inside the crystal, and $\mathbf{K}_g$ is the grating vector defined as

$$\mathbf{K}_g = \mathbf{k}_{e2} - \mathbf{k}_{e1}. \tag{1}$$

If $\mathbf{k}_{o1}$ and $\mathbf{k}_{o2}$ are the diffracted waves, then the Bragg conditions for anisotropic self-diffraction are

$$\mathbf{k}_{o2} = \mathbf{k}_{e2} + \mathbf{K}_g, \tag{2}$$

$$\mathbf{k}_{o1} = \mathbf{k}_{e1} - \mathbf{K}_g. \tag{3}$$

In addition to the Bragg condition, the coupling strength of the diffraction depends on the magnitude of the effective coupling coefficient Γ, which is defined as[8]

$$\Gamma = \hat{e}_d^* \cdot \tilde{\chi} \cdot \hat{e}_i, \tag{4}$$

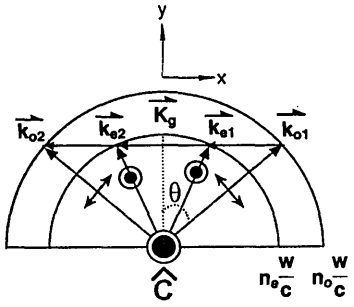

Fig. 1. Wave-vector diagram of anisotropic self-diffraction in $BaTiO_3$. The writing beams are both in extraordinary polarization and the self-diffracted beams are in ordinary polarization, which is orthogonal to the optic axis.

where $\tilde{\chi}$ is the coupling tensor that is described by incident light polarization and crystal orientation,

$$\tilde{\chi} = (\tilde{\epsilon} \cdot \tilde{\gamma} \cdot \hat{e}_{sc} \cdot \tilde{\epsilon}), \tag{5}$$

where $\hat{e}_i$, $\hat{e}_d$, and $\hat{e}_{sc}$ are unit vectors of incident light, diffracted light, and the space-charge field, respectively, $\tilde{\epsilon}$ is the second-rank optical dielectric tensor, and $\tilde{\gamma}$ is the electro-optic tensor.

In $BaTiO_3$ with the incident condition shown in Fig. 1, the refractive indices for ordinary and extraordinary polarizations are both isotropic. If we define the direction of the grating vector as $\hat{k}_g$,

$$\hat{k}_g = (1, 0, 0), \tag{6}$$

then the ordinary and extraordinary polarizations are

$$\hat{e}_o = (\cos\theta, \sin\theta, 0), \tag{7}$$

$$\hat{e}_e = (0, 0, 1), \tag{8}$$

where θ is the angle of ordinary light with respect to the y axis. From Eqs. (4)–(6), the effective coupling coefficient in $BaTiO_3$ can be obtained as

$$\Gamma = \hat{e}_d^* \cdot \begin{pmatrix} 0 & 0 & n_o^2 n_e^2 \gamma_{42} \\ 0 & 0 & 0 \\ n_o^2 n_e^2 \gamma_{42} & 0 & 0 \end{pmatrix} \hat{e}_i. \tag{9}$$

Because of the zero diagonal elements of the matrix in Eq. (9), the coupling between two beams with the same polarization is inhibited. On the other hand, when the incident beams are extraordinarily polarized and the diffracted beams are ordinarily polarized, from Eqs. (7)–(9) the effective coupling coefficient is

$$\Gamma = n_o^2 n_e^2 \gamma_{42} \cos\theta. \tag{10}$$

It is seen that a large effective coupling coefficient exists between different polarized beams because of the direct contribution from the largest electro-optic coefficient γ_{42} of $BaTiO_3$. In addition, under the Bragg conditions shown in Fig. 1, the incident lights $\mathbf{k}_{e1}$ and $\mathbf{k}_{e2}$ not only construct the grating $\mathbf{K}_g$ but are also coupled to $\mathbf{k}_{o1}$ and $\mathbf{k}_{o2}$, respectively. This is the case of anisotropic self-diffraction in $BaTiO_3$.[8–10]

The coupling efficiency in anisotropic self-diffraction[9] can be expressed as

$$\frac{I_o(l)}{I_e(0)} = \frac{1}{1 + \left(\frac{2}{mK\Delta n l}\right)^2}, \tag{11}$$

where $I_o(l)$ is the intensity of diffracted light with ordinary polarization, $I_e(0)$ is the intensity of incident light with extraordinary polarization at position $z = 0$, m is the modulation depth of the grating, K is the wave number in vacuum, l is the interaction length, Δn is the local change of the refractive index,[8,10] and

$$\Delta n = \frac{1}{2}\left[\frac{K_B T}{q} \frac{K_g}{1 + \left(\frac{K_g}{K_o}\right)^2}\right] n_o n_e^2 \gamma_{42}, \tag{12}$$

where

$$K_0 = \left(\frac{N_A q^2}{\epsilon\epsilon_0 K_B T}\right)^{1/2}, \tag{13}$$

$K_B T$ is the thermal energy, q is the electronic charge, N_A is the trap density, and ϵ is the effective dielectric constant in the direction of the grating vector. It can be seen that the diffraction efficiency in Eq. (11) is different from that of Kogelnik's formula.[11] In this case the incident beams act as both writing and reading beams, and the grating dynamics is different from that of a general case.

Matrix–Matrix Multiplication

An $\mathcal{N} \times \mathcal{N}$ matrix–matrix multiplication can be written as

$$\mathscr{C} = \mathscr{A}\mathscr{B}, \tag{14}$$

where the element of the output matrix is

$$c_{ij} = \Sigma\, a_{ik} b_{kj}. \tag{15}$$

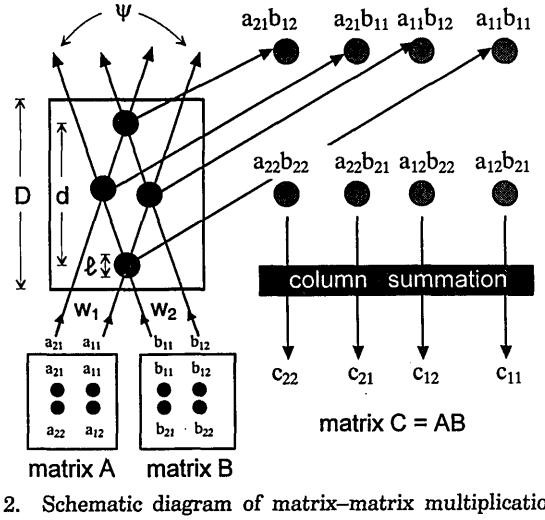

Fig. 2. Schematic diagram of matrix–matrix multiplication by using anisotropic self-diffraction.

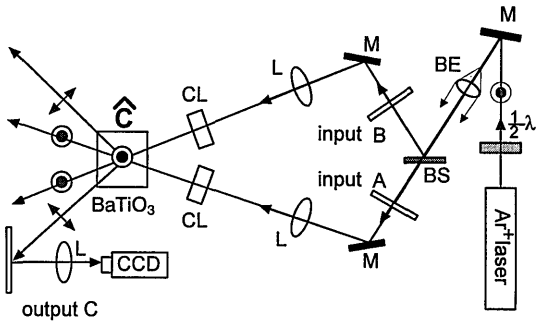

Fig. 3. Experimental setup: M, mirror; L, lens; CL, cylindrical lens; BS, beam splitter; BE, beam expander; $1/2\lambda$, half-wave plate.

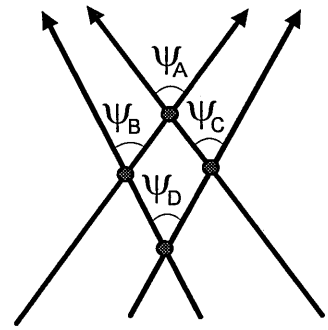

Fig. 5. Converging incident beams induce different incident angles in different interaction units.

Equation (15) shows that the matrix–matrix multiplication can be obtained in two steps: parallel multiplication and summation. The parallel multiplication operation of matrix–matrix multiplication can be performed by using anisotropic self-diffraction in $BaTiO_3$. In this method only two incident beams are needed. As an example, Fig. 2 shows the signals of matrices $\mathscr{A}^T$ and $\mathscr{B}$, which are carried by the two incident beams, and the signals of the corresponding elements of the output matrix are carried by the diffracted beams. For 2×2 matrics, there are two layers of interaction regions in the vertical direction. Each layer is confined in a horizontal plane. And each layer consists of the interaction units of the corresponding row of the first input matrix with the same row of the second input matrix. In this example, four interaction spots in the interaction layer for (a_{21}, a_{11}) and (b_{11}, b_{12}) are shown. For $\mathscr{N} \times \mathscr{N}$ input matrices, there are N interaction layers inside the crystal and each layer has N^2 interaction spots. Totally there are N^3 interaction regions inside the crystal. The right-hand part of Fig. 2 shows the second step of the matrix–matrix multplication. Summation of the diffracted beams along the vertical column produces the element of the output matrix. As an example, c_{11} is the sum of the diffracted beams $a_{11}b_{11}$ from the upper layer and $a_{12}b_{21}$ from the lower layer.

In practical situations the maximal scale of the matrix that can be operated is limited by the crystal dimensions because we need a crystal with enough size along the light propagation direction to accommodate all interaction regions of the corresponding rows. In the vertical direction we only need the crystal size to be larger than the size of the matrix column. By using a simple geometric method, it can be shown that the interaction length d and crystal length D must satisfy

$$d = \frac{1}{\sin \psi}(w_1^2 + 2w_1w_2 \cos \psi + w_2^2)^{1/2} \leq D, \quad (16)$$

where w_1 and w_2 are the widths of the input beams inside the crystal, respectively, d is the whole interac-

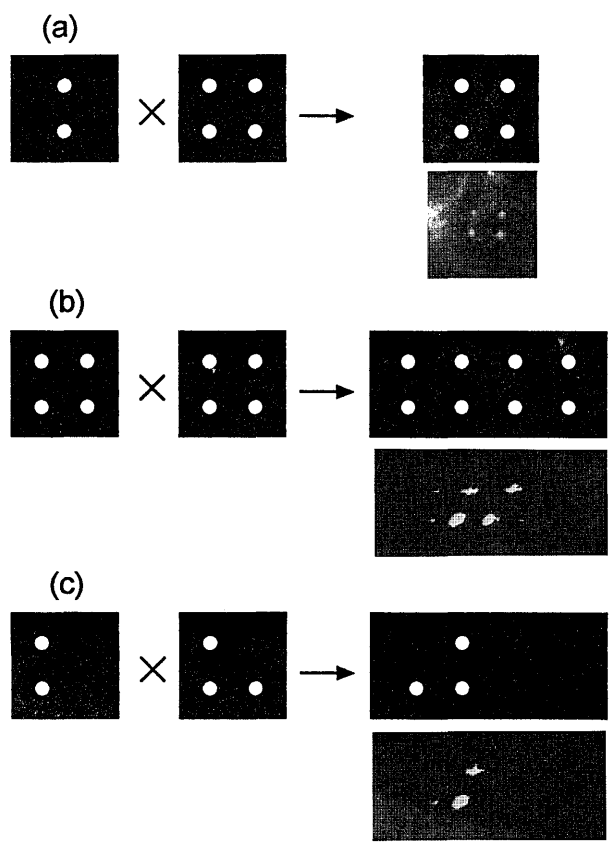

Fig. 4. Inputs and experimental results. The slant outputs of (b) and (c) were caused by the inclination of the output plate.

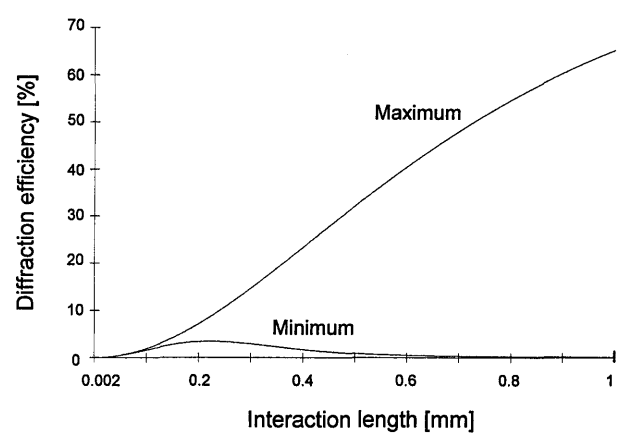

Fig. 6 Theoretical calculations of maximum and minimum diffraction efficiencies as functions of interaction length.

tion length, and ψ is the incident angle inside the crystal that satisfies the Bragg condition[9]:

$$\psi = \sin^{-1}\frac{1}{n_e}\left[\frac{1}{8}(n_o^2 - n_2^2)\right]^{1/2}. \quad (17)$$

For an incident wavelength at 488 nm, n_o is 2.52 and n_e is 2.46.[12] By applying Eq. (17), ψ is 4.5° inside the crystal. If $w_1 = w_2 \equiv w$ and we assume that the spacing between the elements of column and row are equal, the requirement of the crystal length is

$$\frac{D}{w} \geq \frac{2\cos\frac{\psi}{2}}{\sin\psi} \approx 25. \quad (18)$$

In equality (18) shows that a long crystal is required for large-scale operations.

Experimental Results

The experimental setup is shown in Fig. 3. We used a single longitudinal mode TEM_{00} argon-ion laser at 488 nm as the light source. The writing beams are extraordinarily polarized and are expanded and collimated. The input signals are two 2 × 2 matrices. The crystal dimensions are 5.7 mm × 5.2 mm × 5 mm ($a \times b \times c$). From inequality (18) the effective input beamwidth must be less than 0.23 mm in order that all the interaction regions can be covered inside the crystal. However, in the experiment, the effective input beamwidth is 3 mm. In order to cover all the interaction regions in the crystal, we used a spherical lens with 400-mm focal length to focus the input beam to be required size of 0.23 mm. A cylindrical lens was added to recollimate the input beam in the vertical direction because in this direction the crystal dimension must only be larger than the width of the input beam. Figure 4 shows outputs of the multplication for three different input matrices. In each of the three cases, the upper drawing shows the elements of the input matrices and the corresponding diffracted patterns, and the lower photograph shows the experimental results. In the experiments we adjusted the two incident beams around the Bragg angle so that the output intensities are as nearly uniform as possible. In fact, nonuniformity occurs as shown in Figs 4(b) and 4(c). This nonuniformity is due to the Bragg mismatch produced by horizontal focusing of the input beams. As shown in Fig. 5, if the input matrices are both 2 × 2, then the focusing makes the incident angles different for the four interaction regions. Let the incident angles be ψ_A, ψ_B, ψ_C, and ψ_D, respectively. Then it can be found by geometry that ψ_A is the largest angle and ψ_D is the smallest. If the Bragg angle is located at the angle between ψ_B and ψ_C, the Bragg conditions are better

Table 1. Theoretical Calculations of the Diffraction Intensity when the Interaction Length in Every Interaction Unit is 1 mm[a]

Matrix A — 10 LAYERS

Matrix B

1	1	2	3	4	5	6	7	8	9	10
1	65.24	58.96	45.06	33.60	26.09	21.12	17.66	15.15	13.24	11.75
2	20.50	26.77	30.29	24.95	19.77	16.23	13.75	11.94	10.56	9.48
3	6.42	7.86	16.08	21.37	17.96	14.63	12.31	10.66	9.42	8.46
4	2.63	2.41	4.42	13.09	18.24	15.36	12.62	10.70	9.33	8.30
5	1.36	1.07	1.38	3.52	11.70	16.42	13.85	11.41	9.72	8.49
6	0.81	0.60	0.65	1.11	3.13	10.60	14.95	12.64	10.45	8.92
7	0.54	0.39	0.38	0.53	0.98	2.83	9.69	13.74	11.66	9.66
8	0.38	0.27	0.25	0.31	0.47	0.89	2.58	8.97	12.78	10.86
9	0.28	0.20	0.18	0.21	0.28	0.42	0.81	2.38	8.38	11.99
10	0.22	0.16	0.14	0.15	0.19	0.25	0.38	0.74	2.22	7.90

10

[a]Both inputs are 10 × 10 matrices, and the intensity of every input channel is 100 units. Only one interaction layer is shown.

matched by ψ_B and ψ_c than those of ψ_A and ψ_D. This is why the two central spots are brighter than the side spots in Figs. 4(b) and 4(c). Therefore, in large-scale operations, collimated input beams are required for obtaining uniform output.

Discussion

In the following subsections we discuss the dynamic range of output intensity and system speed of this device.

Dynamic Range of the Output Intensity

From Eq. (11) the output intensity is a function of both the modulation depth and the intensities of the input beams. In the operation of $\mathcal{N} \times \mathcal{N}$ matrix multiplications, each element of one matrix interacts with N elements in the corresponding row of the second matrix. For simplification we assume that N is 10 and the intensity of all the input elements is normalized to be 100 units. Numerical calculations of the maximum and minimum diffraction intensities for interaction length from 0.01 to 1 mm are shown in Fig. 6. In this figure we found that the dynamic range of the output intensity is a function of interaction length. Tables 1 and 2 show the output intensity of each interaction unit for interaction lengths of 1 and 0.08 mm, respectively. Note that there are ten such layers stacked vertically for the total output of the matrix multiplication, and Tables 1 and 2 show only one such layer. We have to sum up these ten layers in order to obtain the final output matrix. The parameters we used for the calculations are $N_A = 2 \times 10^{16}$ cm^{-3}, $\gamma_{42} = 1640$ pm/V, and $\epsilon/\epsilon_o = 3700$.[13] As shown in Table 1, long interaction length provides strong diffraction, but at the same time it leads to large nonuniformity of the output. Therefore, we need to choose a moderate interaction length to obtain large diffraction efficiency and good uniformity. An example is shown in Table 2, where the diffraction efficiency is in the range of $(1.12 \pm 0.11)\%$ and the system efficiency is greater than 10%. Here we want to point out that this architecture is difficult for performing general matrix–matrix multiplications. The reason is that, as shown in Eq. (11), the output intensity is a nonlinear function that depends both on the incident intensity and the modulation depth of each interaction unit. Especially when the matrix is large, the modulation depth of an interaction spot can be different from one spot to another spot. In this case the interaction length should be kept short; otherwise great nonlinearity results, as shown in Table 1. Although an analog matrix–matrix multiplication is not achievable, the output can be thresholded, which makes it more appropriate as an optical interconnect. As an example, input A can be treated as a control of the interaction between input B and output C.

System Speed

In photorefractive devices the operation speed depends on the material, mobility of carriers, operation temperature, and total incident intensity.[14] The to-

Table 2. Theoretical Calculations of Diffraction Outputs when the Interaction Length in Every Interaction Unit is 0.03 mm

Matrix A — 10 LAYERS

Matrix B

1	1	2	3	4	5	6	7	8	9	10
1	1.187	1.187	1.187	1.187	1.186	1.186	1.185	1.185	1.184	1.184
2	1.173	1.173	1.173	1.173	1.172	1.172	1.172	1.171	1.171	1.170
3	1.159	1.159	1.159	1.159	1.159	1.159	1.158	1.158	1.157	1.157
4	1.145	1.145	1.145	1.145	1.145	1.145	1.145	1.145	1.144	1.144
5	1.131	1.131	1.131	1.131	1.132	1.132	1.131	1.131	1.131	1.131
6	1.117	1.117	1.118	1.118	1.118	1.118	1.118	1.118	1.118	1.118
7	1.103	1.104	1.104	1.104	1.105	1.105	1.105	1.105	1.105	1.105
8	1.090	1.090	1.091	1.091	1.091	1.092	1.092	1.092	1.092	1.092
9	1.076	1.077	1.077	1.078	1.078	1.078	1.079	1.079	1.079	1.079
10	1.063	1.064	1.064	1.065	1.065	1.065	1.066	1.066	1.066	1.066

10

Table 3. Theoretical Calculations of the Total Incident Intensity at Each Unit when the Interaction Length in the Unit Is 0.03 mm

Matrix A — 10 LAYERS

Matrix B

1	1	2	3	4	5	6	7	8	9	10
1	200.0	199.5	199.1	198.6	198.1	197.7	197.2	196.7	196.3	195.9
2	199.5	199.1	198.6	198.1	197.7	197.2	196.7	196.3	195.8	195.4
3	199.1	198.6	198.1	197.7	197.2	196.7	196.3	195.8	195.4	194.9
4	198.6	198.1	197.7	197.2	196.7	196.3	195.8	195.4	194.9	194.5
5	198.1	197.7	197.2	196.7	196.3	195.8	195.4	194.9	194.4	194.0
6	197.7	197.2	196.7	196.3	195.8	195.4	194.9	194.4	194.0	193.5
7	197.2	196.7	196.3	195.8	195.4	194.9	194.4	194.0	193.5	193.1
8	196.8	196.3	195.8	195.4	194.9	194.4	194.0	193.5	193.1	192.6
9	196.3	195.8	195.4	194.9	194.4	194.0	193.5	193.1	192.6	192.2
10	195.9	195.4	194.9	194.5	194.0	193.5	193.1	192.6	192.2	191.7

10

tal intensity of each interaction unit corresponding to Table 2 is shown in Table 3. Light intensity of the input matrix elements is assumed to be uniform and is normalized to 100 units, and we obtained the intensity at each interaction unit inside the crystal by summing up the two undiffracted beams that passed through the previous units. Since the response speed of the photorefractive crystal is sublinearly proportional to the incident intensity,[14] the speed of each interaction region can be estimated from these intensities. It is seen in Table 3 that the intensity decays approximately linearly from the upper-left pixel to the lower-right corner along the diagonal direction. By calculating the average intensity of the maximum and minimum intensities, the average speed of each interaction region can be estimated. The matrix multiplication is performed inside the crystal as the light beams propagate from the upper-left corner to the lower-right corner of Table 3. The latter units cannot reach steady state until the former units are in steady state. As a result, they form a $2N - 1$ temporal series of interaction spots for completing an $\mathcal{N} \times \mathcal{N}$ matrix–matrix multiplication operation. This temporally sequential behavior for the output has been observed in our optical experiment. Therefore, operation time of the system will be around $2N - 1$ times the average response time of a unit. This means that the system speed is approximately proportional to $1/(2N - 1)$ of the average speed of a unit. By applying a sublinear relationship between the light intensity and speed, we can estimate the system speed simply by calculating the effective system intensity:

$$I_{\text{sys}} = \frac{I_{\text{av}}}{2N - 1}, \tag{19}$$

where

$$I_{\text{av}} = \frac{I_{\max} + I_{\min}}{2}. \tag{20}$$

Therefore the system speed can be estimated by using Eqs. (19) and (20).

Conclusions

We have proposed a new architecture for performing matrix–matrix multiplication by using anisotropic self-diffraction in $BaTiO_3$. The advantages of this device are that only two input matrices are needed for the operation, and the polarizations of the output lights are orthogonal to the input lights. We have also shown that the diffraction efficiency and the uniformity of the output depend on the length of interaction units inside the crystal. By selecting a proper interaction length a compromise between the diffraction efficiency and output uniformity can be achieved. In general, the diffracted intensity is a nonlinear function of the modulation depth of the input signals. Therefore, this method is not appropriate for ordinary matrix–matrix multiplications. Instead, by thresholding the output, the system is

particularly suitable for applications in optical interconnects and optical switches.

This research is sponsored by the National Science Council of China under contract NSC82-0417-E-008-097. The authors thank Wei Shen, Ming-Trong Lee, and Weichi Chang for their help with the experiments.

References

1. H. J. Caulfield, G. Gheen, and B. J. Thompson, eds., *Selected Papers on Optical Computing,* Proc. Soc. Photo-Opt. Instrum. Eng. **1142** (1989).
2. J.-P. Huignard and P. Gunter, "Optical processing using wave mixing in photorefractive crystals," in *Photorefractive Materials and Their Applications II,* P. Gunter and J.-P. Huignard, eds. (Springer-Verlag, New York, 1989), Chap. 6.
3. P. Yeh, A. E. Choiu, P. Beckwith, T. Chang, and M. Khoshnevisan, "Photorefractive nonlinear optics and optical computing," Opt. Eng. **28,** 328–343 (1989).
4. J. W. Goodman, A. D. Dias, and L. M. Woody, "Fully parallel, high-speed incoherent optical methods for performing discrete Fourier transforms," Opt. Lett. **2,** 1–3 (1978).
5. J. Hong and P. Yeh, "Photorefractive parallel matrix–matrix multiplier," Opt. Lett. **16,** 1343–1345 (1991).
6. C. Gu, S. Campbell, and P. Yeh, "Matrix–matrix multiplication by using grating degeneracy in photorefractive media," Opt. Lett. **18,** 146–148 (1993).
7. N. V. Kukhtarev, V. B. Markov, S. G. Odulov, M. S. Soskin, and V. L. Vinetskii, "Holographic storage in electrooptic crystals. I. Steady state," Ferroelectrics **22,** 949–960 (1979).
8. D. A. Temple and C. Warde, "Anisotropic scattering in photorefractive crystals," J. Opt. Soc. Am. B **3,** 337–341 (1986).
9. N. V. Kukhtarev, E. Kratzig, H. C. Kulich, and R. A. Rupp, "Anisotropic self-diffraction in $BaTiO_3$," Appl. Phys. B **35,** 17–21 (1984).
10. D. A. Temple and C. Warde," High-order anisotropic diffraction in photorefractive crystals," J. Opt. Soc. Am. B **5,** 1800–1805 (1988).
11. H. Kogelnik, "Coupled wave theory for thick hologram gratings," Bell Syst. Tech. J. **48,** 2909–2947 (1969).
12. S. H. Wemple, M. DiDomenico, Jr., and I. Camilibel, "Dielectric and optical properties of melt-grown $BaTiO_3$," J. Phys. Chem. Solids **29,** 1797–1806 (1968).
13. M. B. Klein, "Photorefractive properties in $BaTiO_3$," in *Photorefractive Materials and Their Applications I,* P. G. Gunter and J.-P. Huignard, eds. (Springer-Verlag, New York, 1989), pp. 200–204.
14. P. Yeh, "Fundamental limit of the speed of photorefractive effect and its impact on device applications and material research," Appl. Opt. **26,** 602–604 (1987).