А К А Д Е М И Я Н А У К С С С Р
СИБИРСКОЕ ОТДЕЛЕНИЕ
А В Т О М Е Т Р И Я

№ 6 1972

УДК 681.332

Е. С. НЕЖЕВЕНКО, П. Е. ТВЕРДОХЛЕБ
(Новосибирск)

УМНОЖЕНИЕ МАТРИЦ ОПТИЧЕСКИМ МЕТОДОМ

Нами описан прием умножения знакопеременных матриц в оптической системе, который выгодно отличается от известных голографических приемов высоким коэффициентом использования светового потока и возможностями оперативной перестройки значений элементов умножаемых матриц. Указанные преимущества достигнуты при сохранении основного достоинства оптических устройств умножения матриц — одновременности определения значений элементов результирующей матрицы.

Формулу $c_{mq}=\sum\limits_{n=1}^{N} a_{mn}b_{nq}$, $m=1, 2, \ldots, M$, $g=1, 2, \ldots, G$, по которой обычно вычисляют значения элементов результирующей матрицы $C=\|c_{MG}\|$ при известных элементах исходных прямоугольных матриц $A=\|a_{MN}\|$ и $B=\|b_{PG}\|$, $N=P$, представим в эквивалентном виде

$$c_{mq}=c_{mq}^{+}-c_{mq}^{-}, \tag{1}$$

где c_{mq}^{+} — суммарное значение положительных, а c_{mq}^{-} — суммарное значение отрицательных составляющих общей суммы. Это вызвано необходимостью моделирования в оптической системе процесса умножения матриц со знакопеременными элементами.

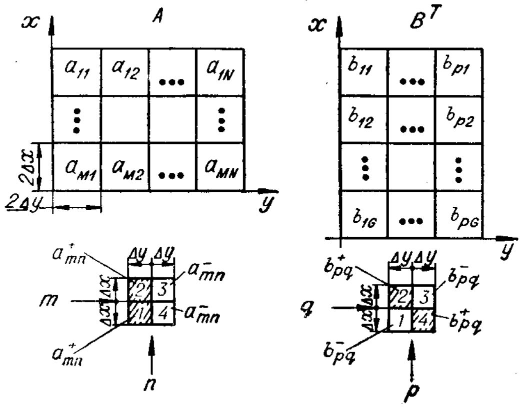

Рис. 1.

Матрицы A и B представим на транспарантах $A(x, y)$ и $B(x, y)$, содержащих соответственно $M\times N$ и $P\times G$ элементов размером $2\Delta x\times 2\Delta y$. Полагаем, что каждый из элементов составлен из четырех элементарных участков размером $\Delta x\times\Delta y$ с номерами *1*, *2*, *3*, *4* (рис. 1). Пропускания (по интенсивности) участков обеспечим пропорциональными значениям элементов матриц. При этом положительные значения a_{mn}^{+} матрицы A будем задавать на участках *1*, *2*, а отрицательные a_{mn}^{-} на участках *3*, *4*. В свою очередь положительные значения b_{pq}^{+} матрицы B, которая в дальнейшем рассматривается в транспонированном виде, представим на участках *2*, *4*, а отрицательные b_{pq}^{-} на участках *1*, *3*. Соотношения между возможными значениями исходных матриц и пропусканиями участков транспарантов приведены в табл. 1 и 2. В этом случае законы изменения пропускания транспарантов можно описать выражениями:

$$A(2m\Delta x, 2n\Delta y) = \sum_{m=1}^{M}\sum_{n=1}^{N}\Big(a_{mn}^{+}\,\mathrm{Rect}\left[\frac{x-(2m-3/2)\Delta x}{\Delta x}\right]\times$$
$$\times\mathrm{Rect}\left[\frac{y-(2n-3/2)\Delta y}{\Delta y}\right]+a_{mn}^{+}\,\mathrm{Rect}\left[\frac{x-(2m-1/2)\Delta x}{\Delta x}\right]\times$$
$$\times\mathrm{Rect}\left[\frac{y-(2n-3/2)\Delta y}{\Delta y}\right]+a_{mn}^{-}\,\mathrm{Rect}\left[\frac{x-(2m-1/2)\Delta x}{\Delta x}\right]\times$$
$$\times\mathrm{Rect}\left[\frac{y-(2n-1/2)\Delta y}{\Delta y}\right]+a_{mn}^{-}\,\mathrm{Rect}\left[\frac{x-(2m-3/2)\Delta x}{\Delta x}\right]\times$$
$$\times\mathrm{Rect}\left[\frac{y-(2n-1/2)\Delta y}{\Delta y}\right]\Big);$$

$$B^{T}(2p\Delta x, 2q\Delta y) = \sum_{p=1}^{P}\sum_{q=1}^{G}\Big(b_{pq}^{-}\,\mathrm{Rect}\left[\frac{x-(2p-3/2)\Delta x}{\Delta x}\right]\times$$
$$\times\mathrm{Rect}\left[\frac{y-(2q-3/2)\Delta y}{\Delta y}\right]+b_{pq}^{+}\,\mathrm{Rect}\left[\frac{x-(2p-1/2)\Delta x}{\Delta x}\right]\times$$
$$\times\mathrm{Rect}\left[\frac{y-(2q-3/2)\Delta y}{\Delta y}\right]+b_{pq}^{-}\,\mathrm{Rect}\left[\frac{x-(2p-1/2)\Delta x}{\Delta x}\right]\times$$
$$\times\mathrm{Rect}\left[\frac{y-(2q-1/2)\Delta y}{\Delta y}\right]+b_{pq}^{+}\,\mathrm{Rect}\left[\frac{x-(2p-3/2)\Delta x}{\Delta x}\right]\times$$
$$\times\mathrm{Rect}\left[\frac{y-(2q-1/2)\Delta y}{\Delta y}\right]\Big), \qquad (2)$$

где

$$\mathrm{Rect}\,[s]=\begin{cases}1, & \text{если } |s|\leqslant 1/2;\\ 0, & \text{если } |s|>1/2,\end{cases}$$

имея, однако, в виду, что под двойной суммой значащими являются либо члены со знаком «+», либо члены со знаком «—», поскольку a_{mn} и b_{pq} могут принимать либо положительные, либо отрицательные значения. Если же $a_{mn}=0$ или $b_{pq}=0$, то члены со знаком «+» и «—» под соответствующими суммами отсутствуют.

Поместим транспаранты с пропусканием (2) соответственно в плоскостях P_1 и P_2 устройства, схема которого приведена на рис. 2. На схеме показаны: *1* — линзовый растр, содержащий M цилиндрических линз с апертурой $2\Delta x\times 2N\Delta y$, образующая которых параллельна оси y; *2*, *4* — ци-

Таблица 1

Матрица A		Транспарант: значение пропускания участков элемента $(2m\,\Delta x,\ 2n\,\Delta y)$ 1	2	3	4	примечание
Значение элемента матрицы	$a_{mn}>0$	a_{mn}^{+}	a_{mn}^{+}	0	0	—
	$a_{mn}<0$	0	0	a_{mn}^{-}	a_{mn}^{-}	$a_{mn}^{-}=\|a_{mn}\|$
	$a_{mn}=0$	0	0	0	0	—

Таблица 2

Матрица B		Транспарант: значение пропускания участков элемента ($2p\,\Delta x$, $2q\,\Delta y$)				
		1	2	3	4	примечание
Значение элемента матрицы	$в_{pq} > 0$	0	$в_{pq}^{+}$	0	$в_{pq}^{+}$	—
	$в_{pq} < 0$	$в_{pq}^{-}$	0	$в_{pq}^{-}$	0	$в_{pq}^{-} = \|в_{pq}\|$
	$в_{pq} = 0$	0	0	0	0	—

линдрические линзы; *3*, *5* — клиновые растры; *6* — сферический объектив. Для простоты фокусные расстояния цилиндрических линз *2*, *4* и объектива *6* будем считать одинаковыми и равными f. Полагаем также, что масштабы плоскостей P_1, P_2, P_3, P_4 одинаковы, т. е. $\Delta x_1 = \ldots = \Delta x_4 = \Delta x$ и $\Delta y_1 = \ldots = \Delta y_4 = \Delta y$.

Если транспарант $A\,(2m\Delta x,\ 2n\Delta y)$ осветить плоской волной света, то совместное действие растра *1* и линзы *2* приведет к тому, что в плоскости P_2 получим совокупность модулированных волн, распространяющихся под разными углами к плоскости yz, т. е.

$$\sum_{m=1}^{M}\sum_{n=1}^{N}\left(2\sqrt{a_{mn}^{+}}\,\mathrm{Rect}\left[\frac{y_2-(2n-3/2)\,\Delta y}{\Delta y}\right]+2\sqrt{a_{mn}^{-}}\,\mathrm{Rect}\left[\frac{y_2-(2n-1/2)\,\Delta y}{\Delta y}\right]\right)\times$$

$$\times\exp\left[-j\frac{2\pi}{\lambda f}\,(2m-1)\,\Delta x x_2\right]. \qquad (3)$$

Допускаем, что линза *2* ограничивающего действия не оказывает. Ограничение и повторная модуляция плоских волн (3) происходят при прохождении их через транспарант с изображением $B^T(2p\Delta x,\ 2q\Delta y)$ и установленный перед ним клиновый растр *3*. Форму клина и способ его установки можно уяснить из рис. 1. Клин сообщает дополнительный линейный фазовый набег волнам, прошедшим только через нижние участки (участки с номерами *1*, *4*) элементов транспаранта, установленного в плоскости P_2.

Поскольку $N = P$, то каждая из N плоских волн (3) освещает соответствующий столбец элементов транспаранта $B^T(2p\Delta x,\ 2q\Delta y)$ полностью. Это обеспечивается выбором отношения $f/f_{\text{р}}$, где f и $f_{\text{р}}$ — фокусные расстояния линзы *2* и линз цилиндрического растра *1*. В результате в задней фокальной плоскости линзы *3* получим распределение световых «штрихов» в виде матрицы размером $M \times N$. Каждый элемент этой матрицы содержит в общем случае четыре «штриха», расположенных на четырех соседних непересекающихся участках и ориентирован-

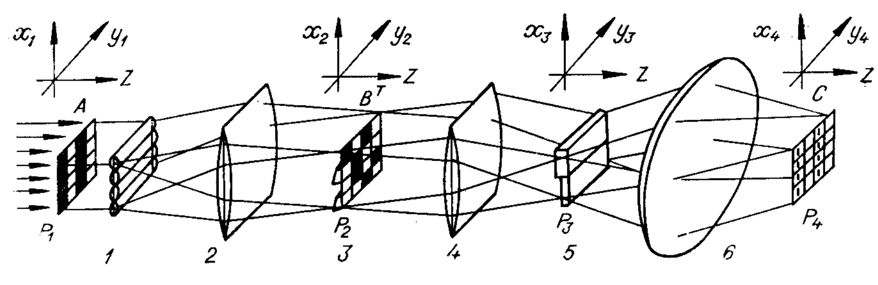

Рис. 2.

ных параллельно оси y_3. Для элемента $(M-m,\ n)$ этой матрицы амплитуды световых «штрихов» распределены по законам:

$$2\sqrt{a^{+}_{mn}}\,\frac{\sin\frac{\pi}{\lambda f}\Delta x\,[x_3-2(M-m)\,\Delta x]}{x_3-2(M-m)\,\Delta x}\,\mathrm{Rect}\left[\frac{y_3-(2n-3/2)\,\Delta y}{\Delta y}\right]\times$$

$$\times\left(\sum_{q=1}^{G}\sqrt{b^{-}_{pq}}\exp\left[-j\frac{2\pi}{\lambda f}\left(2q-\frac{3}{2}\right)\Delta x x_3\right]\right)\text{на участке } 1; \qquad (4)$$

$$2\sqrt{a^{+}_{mn}}\,\frac{\sin\frac{\pi}{\lambda f}\Delta x\,\{x_3-[2(M-m)+1]\,\Delta x\}}{x_3-[2(M-m)+1]\,\Delta x}\,\mathrm{Rect}\left[\frac{y_3-(2n-3/2)\,\Delta y}{\Delta y}\right]\times$$

$$\times\left(\sum_{q=1}^{G}\sqrt{b^{+}_{pq}}\exp\left[-j\frac{2\pi}{\lambda f}(2q-1/2)\,\Delta x x_3\right]\right)\ \text{на участке } 2; \qquad (5)$$

$$2\sqrt{a^{-}_{mn}}\,\frac{\sin\frac{\pi}{\lambda f}\Delta x\,\{x_3-[2(M-m)+1]\,\Delta x\}}{x_3-[2(M-m)+1]\,\Delta x}\,\mathrm{Rect}\left[\frac{y_3-(2n-1/2)\,\Delta y}{\Delta y}\right]\times$$

$$\times\left(\sum_{q=1}^{G}\sqrt{b^{-}_{pq}}\exp\left[-j\frac{2\pi}{\lambda f}(2q-1/2)\,\Delta x x_3\right]\right)\ \text{на участке } 3; \qquad (6)$$

$$2\sqrt{a^{-}_{mn}}\,\frac{\sin\frac{\pi}{\lambda f}\Delta x\,[x_3-2(M-m)\,\Delta x]}{x_3-2(M-m)\,\Delta x}\,\mathrm{Rect}\left[\frac{y_3-(2n-1/2)\,\Delta y}{\Delta y}\right]\times$$

$$\times\left(\sum_{q=1}^{G}\sqrt{b^{+}_{pq}}\exp\left[-j\frac{2\pi}{\lambda f}(2q-3/2)\,\Delta x x_3\right]\right)\ \text{на участке } 4. \qquad (7)$$

Из (4), (7) и (5), (6) следует, что плоские волны вида $\sqrt{a^+}\sqrt{b^-}\exp[\cdot]$ и $\sqrt{a^-}\sqrt{b^+}\exp[\cdot]$ (волны с «отрицательными» амплитудами) суммируются на участках плоскости P_3 с координатами $[2(M-m)\Delta x,\ (2n-1)\Delta y]$ и $[2(M-m)\Delta x,\ 2n\Delta y)]$, а плоские волны вида $\sqrt{a^+}\sqrt{b^+}\exp[\cdot]$ и $\sqrt{a^-}\sqrt{b^-}\exp[\cdot]$ (волны с «положительными» амплитудами) на участках с координатами $\cdot[(2(M-m)+1)\Delta x,\ (2n-1)\Delta y]$ и $[(2(M-m)+1)\Delta x,\ 2n\Delta y]$. Раздельное суммирование компонент (4) — (7) обеспечивается как за счет принятого способа представления на транспарантах элементов матриц, так и за счет действия клинового растра *3*. Параметры последнего выбраны так, что волны, получившие линейный фазовый набег, суммируются на участках *1* и *4* элементов плоскости P_3. Множитель Sinc $[\cdot]$ Rect $[\cdot]$ определяет вид огибающей светового «штриха» и является результатом одномерного преобразования Фурье от апертуры $(\Delta x\times\Delta y)$ значащего участка транспаранта $B^{T}(2p\Delta x,\ 2q\Delta y)$.

Выделим из матрицы, полученной в плоскости P_3, $(M-m)$-ю строку элементов. Амплитуды световых «штрихов» этих элементов будут изменяться в соответствии с (4) — (7) при пробегании индексом $n=p$ значений 1, 2,..., N. Над полученным амплитудным распределением выполним операцию двумерного преобразования Фурье. Интересующий нас результат получим в задней фокальной плоскости объектива *6* в виде амплитудного распределения

$$\left[\sum_{q=1}^{G}\mathrm{Rect}\left[\frac{x_4-(2q-3/2)\,\Delta x}{\Delta x}\right]\left(2\sum_{n=1}^{N}\sqrt{a^{+}_{mn}b^{-}_{pq}}\exp\left[-j\frac{2\pi}{\lambda f}(2n-3/2)\,\Delta y y_4\right]\right)\right]\times$$

$$\times\frac{\sin\frac{\pi}{\lambda f}\Delta y y_4}{y_4}\exp\left[-j\frac{2\pi}{\lambda f}\,2(M-m)\,\Delta x x_4\right]+\left[\sum_{q=1}^{G}\times\right.$$

$$\begin{aligned}
&+\left[\sum_{q=1}^{G}\mathrm{Rect}\left[\frac{x_4-(2q-1/2)\,\Delta x}{\Delta x}\right]\left(2\sum_{n=1}^{N}\sqrt{a_{mn}^{-}b_{nq}^{-}}\times\right.\right.\\
&\left.\left.\times\exp\left[-j\frac{2\pi}{\lambda f}(2n-1/2)\,\Delta y y_4\right]\right)\right]\frac{\sin\frac{\pi}{\lambda f}\,\Delta y y_4}{y_4}\exp\left\{-j\frac{2\pi}{\lambda f}[2(M-m)+1]\times\right.\\
&\left.\times\,\Delta x x_4\right\}+\left[\sum_{q=1}^{G}\mathrm{Rect}\left[\frac{x_4-(2q-3/2)\,\Delta x}{\Delta x}\right]\left(2\sum_{n=1}^{N}\sqrt{a_{mn}^{-}b_{nq}^{+}}\times\right.\right.\\
&\left.\left.\times\exp\left[-j\frac{2\pi}{\lambda f}(2n-1/2)\,\Delta y y_4\right]\right)\right]\frac{\sin\frac{\pi}{\lambda f}\,\Delta y y_4}{y_4}\exp\left[-j\frac{2\pi}{\lambda f}\,2\,(M-m)\times\right.\\
&\left.\times\,\Delta x x_4\right].
\end{aligned}\tag{8}$$

Можно видеть, что «штрихи» $(M-m)$-й строки, ориентированные параллельно оси y_3, преобразуются в столбец «штрихов», ориентированных параллельно оси x_4. Плоские волны с «положительными» амплитудами вида $\sqrt{a^+b^+}$ и $\sqrt{a^-b^-}$ суммируются на участках плоскости P_4 с координатой по оси x $(2q-1)\Delta x$, а волны с «отрицательными» амплитудами вида $\sqrt{a^+b^-}$ и $\sqrt{a^-b^+}$ — на соседних участках с координатой $2q\Delta x$. Вторая координата этих участков зависит от величины линейного набега фаз, сообщаемого одним из клиньев растра «штрихам» $(M-m)$-й строки.

Волны на каждом из рассматриваемых участков интерферируют между собой, поскольку имеют отличающиеся наклонные фазовые фронты. Огибающая результата интерференции имеет вдоль оси y вид функции $\mathrm{Sinc}\,[\cdot]$. Однако, несмотря на явление интерференции, интегральные световые потоки, определяемые на двух соседних участках (площадью $\Delta x\Delta y$) с помощью пары независимых фотоприемников, будут близки к значениям:

$$\begin{aligned}
I_{mq}^{-}\,[(2q-1)\,\Delta x] &= \sum_{n=1}^{N}\left(a_{mn}^{+}b_{nq}^{-}+a_{mn}^{-}b_{nq}^{+}\right);\\
I_{mq}^{+}\,[2q\Delta x] &= \sum_{n=1}^{N}\left(a_{mn}^{+}b_{nq}^{+}+a_{mn}^{-}b_{nq}^{-}\right),
\end{aligned}\tag{9}$$

которые пропорциональны значениям c_{mq}^{-} и c_{mq}^{+} из (1).

Тогда значение элемента c_{mq} матрицы C определяется разностью

$$\Delta I_{mq}=I_{mq}^{+}\,[2q\Delta x]-I_{mq}^{-}\,[(2q-1)\,\Delta x],\tag{10}$$

которая может принимать как положительное, так и отрицательное значения. Предполагается, что вычисление разности (10) будет производиться в выходных электронных блоках устройства умножения матриц.

Остальные столбцы матрицы C воспроизводятся в плоскости P_4 подобным образом в результате выполнения двумерного преобразования Фурье над амплитудным распределением «штрихов» других строк матрицы, полученной в плоскости P_3. Для того чтобы столбцы в пло-

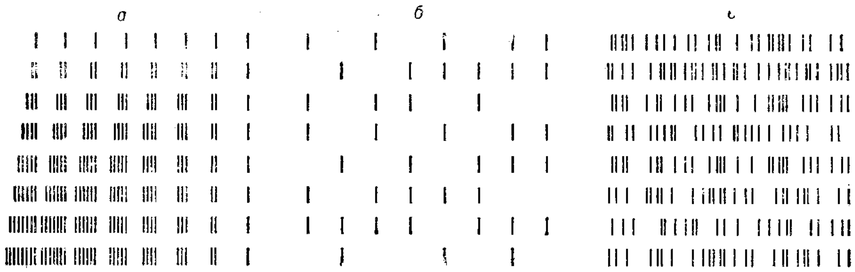

Рис. 3.

скости P_4 не перекрывались и располагались вдоль оси y_4 с постоянным шагом, «штрихи» каждой из строк получают вдоль оси y_3 линейный фазовый набег, изменяющийся на фиксированную величину от строки к строке. Это обеспечивается путем соответствующего выбора параметров клиньев растра. Число клиньев в растре *5* равно числу строк.

Как и в предыдущем случае, значения элементов остальных столбцов матрицы C определяются по формулам (9) и (10). Очевидно, что для обеспечения параллельной работы устройства выходной фотоприемник должен быть многоэлементным с числом элементов, равным $2MG$.

На рис. 3, *а—в* приведены результаты экспериментального исследования описанного устройства для трех ситуаций: а) матрицы A и B — треугольные; б) матрица A — произвольная, B — единичная; в) матрицы A и B — произвольные.

Умножались положительные бинарные матрицы размером 8×8. Растр *3* в устройстве отсутствовал. Для наглядности значения c_{mq} в случаях а) — в) представлены на отдельных участках числом полосок. Такая картина получается в случае небольшого смещения плоскости наблюдения относительно фокальной плоскости объектива *6*. Поскольку полоски воспроизводятся в виде световых «штрихов», то результирующий световой поток на каждом из участков пропорционален значениям c_{mq}.

Таким образом, в описанном устройстве значения знакопеременных матриц A и B задаются непосредственно с помощью управляемых транспарантов и могут оперативно изменяться. Световой поток источника излучения используется почти полностью, за исключением потерь в оптических элементах с конечной апертурой. Применяемые в [1, 2] фильтры-голограммы, требующие реализации оперативного голографического процесса и имеющие сравнительно низкую дифракционную эффективность, здесь не используются. Недостатком устройства является то, что для задания матрицы $A(B)$ с числом элементов $MN(PG)$ на транспаранте необходимо иметь $4MN(4PG)$ участков с изменяемым пропусканием. Однако количество таких участков можно уменьшить до $2MN(4PG)$ путем усложнения разводки световых пучков линзовыми и клиновыми растрами. Этим же путем можно добиться того, что вид результирующей матрицы будет соответствовать виду исходных матриц.

ЛИТЕРАТУРА

1. R. A. Heinz, I. O. Artman, S. M. Lee. Matrix Multiplication by Optical Methods.— Applied Optiics, 1970, v. 9, № 9.
2. D. P. Jablonowski, R. A. Heinz, I. O. Artman. Matrix Multiplication by Optical Methods. Experimental Verification.— Applied Optics, 1972, v. 11, № 1.

*Поступила в редакцию
22 мая 1972 г.*