

# 光学多比特矩阵乘法器

周少敏 邬敏贤 金国藩  
(清华大学精密系, 北京 100084)

**提要:** 本文提出并实现了一种利用改进的符号数(MSD)算法和多窗口解码光学符号代换法则(MW-OSSR)实现多比特矩阵相乘的光学方法, 它具有精度高和速度快的特点。实验中的计算速度为每秒实现两个  $2 \times 2$  阶 32 比特的 MSD 矩阵相乘。

**关键词:** 光矩阵计算, 光学符号代换

## Optical multi-bit matrix multiplier

Zhou Shaomin, Wu Minxian, Jin Guofan  
(Department of Precision Instruments, Tsinghua University, Beijing)

**Abstract:** A method of optical matrix multiplication using Modified Signed-Digit (MSD) arithmetic and multi-window decoding optical symbolic substitution rule is proposed and realized. It operates at high speed and high accuracy, and multiplies experimentally two 22MSD matrix in 32 bits one second.

**Key words:** optical matrix computation, optical symbolic substitution

光学上的许多变换和信号处理, 均可转化成一种基本的矩阵运算, 其中最引人注目的是光矩阵相乘, 而且已出现了许多具体方法<sup>[1]</sup>。目前所存在的问题是运算的时序性、低精度或者需要很复杂的光电系统等。本文将提出一种改进的符号数算法(MSDA)和多窗口解码光学符号代换法则(MW-OSSR)实现多比特光矩阵计算的方法, 它具有精度高和速度快的特点。

### 一、改进的符号数(MSD)多比特矩阵乘法

任意一个十进制数  $D$  可由下式表示成 MSD 数<sup>[2]</sup>:

$$(D) = \sum_{k=0}^{N-1} a_k Z^k = [a_{N-1} a_{N-2} \cdots a_k \cdots a_1 a_0]_{\text{MSD}} \quad (1)$$

其中,  $a_k \in \{1, 0, \bar{1}\}$ ,  $N$  为 MSD 数的表达精度。一个 MSD 负数是其对应的 MSD 正数的逻辑补(定义:  $1' = \bar{1}$ ,  $\bar{1}' = 1$ ,  $0' = 0$ , “'”表示取补), 例如

$$(13)_D = [10\bar{1}01]_{\text{MSD}} = [01101]_{\text{MSD}}$$

$$(-13)_D = [\bar{1}010\bar{1}]_{\text{MSD}} = [0\bar{1}\bar{1}0\bar{1}]_{\text{MSD}}$$

MSD 多比特矩阵相乘分为求部分积阵列和部分积阵列求和两步来进行。以两个  $2 \times 2$  阶  $N$  位精度的 MSD 矩阵相乘为例:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}^{N-1} a_{11}^{N-2} \dots a_{11}^1 a_{11}^0 & a_{12}^{N-1} a_{12}^{N-2} \dots a_{12}^1 a_{12}^0 \\ a_{21}^{N-1} a_{21}^{N-2} \dots a_{21}^1 a_{21}^0 & a_{22}^{N-1} a_{22}^{N-2} \dots a_{22}^1 a_{22}^0 \end{pmatrix} \quad (2)$$

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11}^{N-1} b_{11}^{N-2} \dots b_{11}^1 b_{11}^0 & b_{12}^{N-1} b_{12}^{N-2} \dots b_{12}^1 b_{12}^0 \\ b_{21}^{N-1} b_{21}^{N-2} \dots b_{21}^1 b_{21}^0 & b_{22}^{N-1} b_{22}^{N-2} \dots b_{22}^1 b_{22}^0 \end{pmatrix} \quad (3)$$

其积矩阵为

$$O = O_1 + O_2 \quad (4)$$

其中,

$$O_1 = A_1 \cdot B_1 = (a_{11}^{N-1} a_{11}^{N-2} \dots a_{11}^1 a_{11}^0 a_{21}^{N-1} a_{21}^{N-2} \dots a_{21}^1 a_{21}^0)^T \cdot (b_{11}^{N-1} b_{11}^{N-2} \dots b_{11}^1 b_{11}^0 b_{12}^{N-1} b_{12}^{N-2} \dots b_{12}^1 b_{12}^0) \quad (5)$$

$$O_2 = A_2 \cdot B_2 = (a_{12}^{N-1} a_{12}^{N-2} \dots a_{12}^1 a_{12}^0 a_{22}^{N-1} a_{22}^{N-2} \dots a_{22}^1 a_{22}^0)^T \cdot (b_{21}^{N-1} b_{21}^{N-2} \dots b_{21}^1 b_{21}^0 b_{22}^{N-1} b_{22}^{N-2} \dots b_{22}^1 b_{22}^0) \quad (6)$$

我们称  $O_1$ 、 $O_2$  为部分积,  $(\dots)^T$  表示转置向量。式 (5)、(6) 相乘结果中的任一元素  $a_i^k b_{lm}^k$  ( $K_1, K_2 = 0, 1, 2, \dots, N-1; i, j, l, m = 1, 2$ ) 按图 1(a) 所示规则计算, 称之为  $M$  变换<sup>[3]</sup>。

部分积阵列的求和是通过设置“树结构 (tree structure)”完成的。对部分积阵列依次每两行为一组进行 MSD 移位加法 (当  $N$  为奇数时, 剩下最后一行为单独一组), 经过  $N/2$  次 (或  $N/2+1$  次) 循环后, 将得到最终结果。两个 MSD 数  $X, Y$  的加法分三步进行:

步骤一: 对  $X, Y$  作  $T, W$  变换; 步骤二: 将  $T$  变换结果左移一位再与  $W$  变换结果作  $T', W'$  变换;

步骤三: 将  $T'$  变换结果左移一位再与  $W'$  变换结果作  $T$  变换。 $T, T', W, W'$  变换规则由图 1(b) 给出, 它们具有一个隐含的进位方式, 使得两个多位 MSD 数的加法可并行完成<sup>[3]</sup>。

## 二、多窗口解码光学符号代换法则 (MW-OSSR) 用于 MSD 多比特矩阵外积计算

光学符号代换法则以二维模式的相关 (识别) 和传递 (代换) 为基础, 具有并行处理和扇入扇出 (fan-in fan-out) 功能。为了实现多通道处理, 作者曾提出过一种所谓的多窗口解码光学符号代换法则<sup>[4]</sup>, 它可以把  $M^2$  个识别通道减少到  $P$  个通道:

$$P = 2(M-2) + 1 \quad (7)$$

式中,  $M \geq 2$  为输入数据中具有不同值的数据数。图 2 示出了二通道的 MW-OSSR, 其对应的光学系统如图 3 所示。输入数据  $A, B$  经  $M_1, M_2, BS_2, BS_3$  移位合成并由  $L_1, BS_1$  成像到 BSO-PROM 上, 进行 NOR 运算 (用  $Ar^+$  激光束写入), 然后, 用 He-Ne 激光束对 PROM

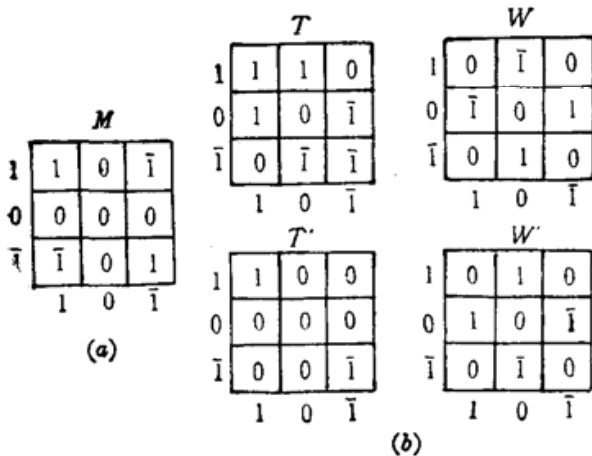


Fig. 1 Mathematical operation rule of MSD

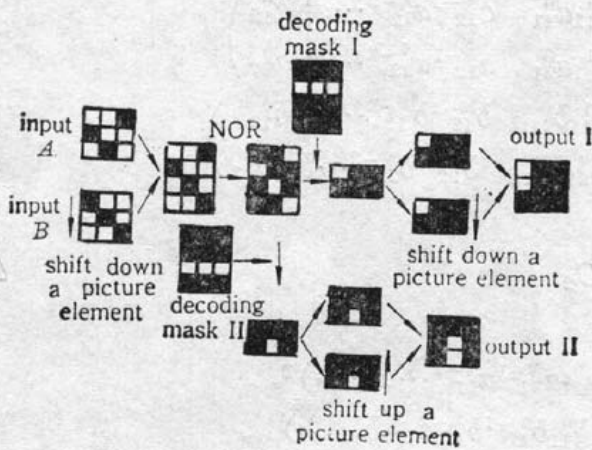


Fig. 2 Principle of two channel MW-OSSR

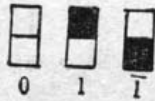


Fig. 4 Dual rail coding of operator

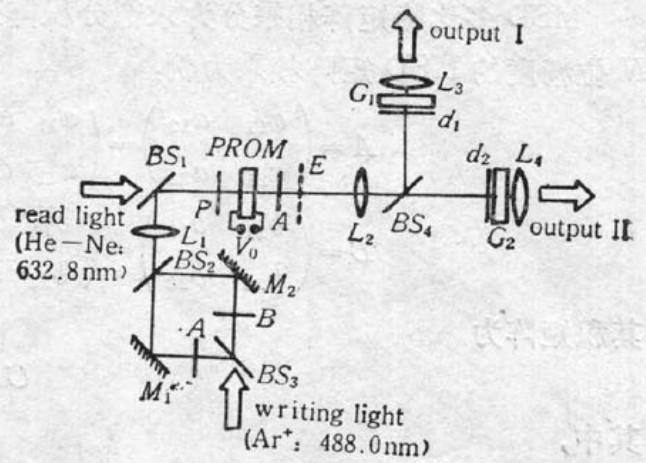


Fig. 3 Optical system of two channel MW-OSSR

$BS_1 \sim BS_4$ —beam splitter;  $L_1 \sim L_4$ —lenses;  $M_1, M_2$ —mirrors;  $A, B$ —input patterns;  $G_1, G_2$ —diffraction gratings;  $d_1, d_2$ —decoding masks;  $E$ —equivalent decoding;  $P$ —polarizer;  $A$ —analyzer

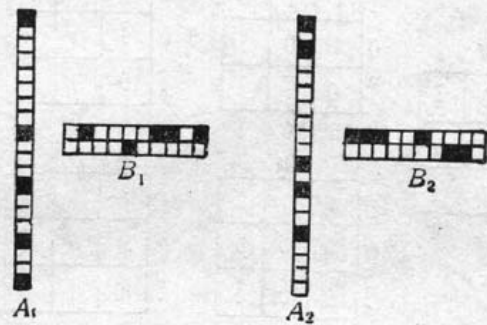


Fig. 5 Dual rail coding results of  $A_1, B_1, A_2$  and  $B_2$

进行读操作，其结果被  $BS_4$  分成两部分，分别在相应的通道解码、代换及输出，代换过程由衍射光栅  $G_1$  和  $G_2$  来完成。

图 1 示出的运算规则实质上也是一种符号代换过程，有 9 种待识别模式：11、10、1 $\bar{1}$ 、01、00、0 $\bar{1}$ 、 $\bar{1}\bar{1}$ 、 $\bar{1}0$  及  $\bar{1}\bar{1}$ ，代换结果为 1、0 或  $\bar{1}$ 。注意到乘积运算时，含“0”模式的代换结果为 0。在光强上对部分积阵列不参加贡献，所以，实际上仅仅需要识别 11、 $\bar{1}\bar{1}$ 、1 $\bar{1}$  及  $\bar{1}1$  等 4 种模式。首先对操作数进行二则编码（图 4），按 MW-OSSR 原理，11 和  $\bar{1}\bar{1}$  模式的识别只需一个通道，其余两个模式各用一个通道。这就是  $M = 3$  (1, 0,  $\bar{1}$  三种)， $P = 3$  的情况。

具体地设  $A, B$  分别为

$$A = \begin{pmatrix} 17 & 24 \\ 9 & 20 \end{pmatrix}$$

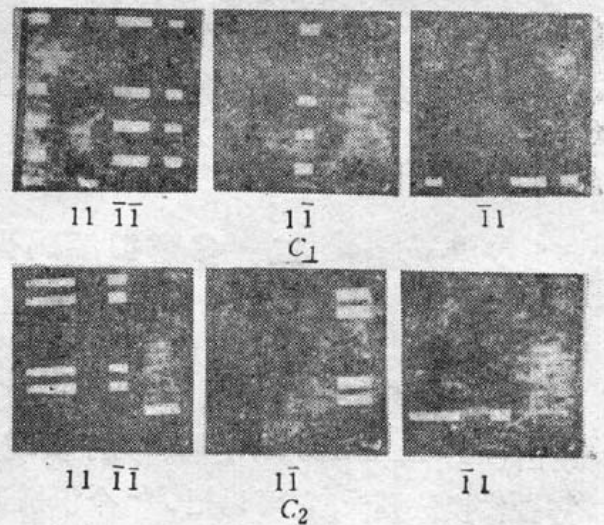


Fig. 6 Experimental results of partial product computation

$$B = \begin{pmatrix} 7 & 13 \\ 28 & 10 \end{pmatrix}$$

取  $N=5$  时,有

$$C_1 = (100010101\bar{1})^x (0100\bar{1}01101)_{\text{MSD}}$$

$$C_2 = (1100011\bar{1}00)^x (1110010\bar{1}10)_{\text{MSD}}$$

其编码结果如图 5 所示,图 6 为实验结果,实验中的代换是用分束镜及反射镜完成的,但当通道数较多时,适合用衍射光栅来完成。

上述矩阵相乘所需的时间  $T_M$  由两部分组成: 每个外积式的计算时间  $T_P$ , 每个部分积阵列移位相加时间  $T_A$ :

$$T_M = 2(T_P + T_A) \approx 2(3T_r \times \log_2 N + T_r) \quad (8)$$

上式中,  $T_r$  表示 BSO-PROM 的响应时间,约等于 5 ms, 忽略了光束在自由空间的传输时间。实验中所用 PROM 器件的有效口径和分辨率分别为  $25 \times 25 \text{ mm}^2$  和 10 线/mm, 像素尺寸为  $1 \times 0.5 \text{ mm}^2$ , 则系统的计算速度为每秒可实现两个  $2 \times 2$  阶的  $N=32$  比特的 MSD 矩阵相乘,  $N$  由下式求得:

$$25^2/2 N \times 2 N = T_M = 2T_r(3 \log_2 N + 1) \quad (9)$$

### 三、结 论

在上述方式中,通过减小像素尺寸和 NOR 门阵列器件的响应时间,可以使计算速度大大提高。此外,MSD 数字表达精度  $N$  决定了最大操作数大小和计算精度,设整数部分需  $N_1$  比特表示,则  $N - N_1$  表示小数部分,当  $N=5$ ,  $N_1=2$  时,计算精度为 0.125, 当  $N=32$ ,  $N_1=16$  时,计算精度可达  $10^{-5}$ 。

### 参 考 文 献

- 1 陈岩松 *et al.*, 物理, **17**(10), 626(1988)
- 2 B. Drake *et al.*, *Opt. Eng.*, **25**(1), 38(1986)
- 3 N. Takagi *et al.*, *IEEE Trans. Comput.*, **C-34**(9), 789 (1985)
- 4 周少敏, 鄢敏贤 *et al.*, 仪器仪表学报, **11**(1), 34(1990)

(上接第 480 页)

为 592.339 nm、592.357 nm 和 592.374 nm, 而其中以  $1s_5 \rightarrow 4s_1''$  谱线的跃迁几率最大, 所以, 我们认为 592.3 nm 附近观察到的光电流信号应属于  $1s_5 - (2p) - 4s_1''$  的双光子跃迁的信号。同理, 可以判断 597.7 nm 附近的光电流信号应为  $1s_4 - (2p) - 4s_1''$  的双光子跃迁信号。因为  $1s_4$  态的原子布居比  $1s_5$  态少, 且中间的虚能级偏离大, 因而观察到的双光子跃迁的光电流信号比 592.3 nm 跃迁的更小。

### 参 考 文 献

- 1 钟旭滨, 刘炳模 *et al.*, 中国激光, **15**(7), 440 (1988)
- 2 K. C. Smyth *et al.*, *Chem. Phys. Lett.*, **55**, 466(1978)
- 3 International Colloquium on Optogalvanic Spectroscopy and its Applications, June 20th-24th, 1983, Aussois (France), 23

(收稿日期: 1989 年 3 月 16 日)