

光学矩阵运算 (I)

陈岩松 张东生

(中国科学院物理研究所)

朱伟利

(中央民族学院物理系)

一、概 述

随着现代科学技术和生产力的发展,人们对计算机性能的要求越来越高。一是要求计算机具有极高的运算速度;二是要求计算机不仅能够进行数值运算,而且还要具有某些人工智能。现有的电子计算机的运算速度和储存容量已不能满足这些要求。人们迫切需要研制某种超高速度和超高容量的新型计算机。目前呼声甚高的光学计算机就是这种新型计算机之一。由于光学计算以光速进行,并且具有二维信号并行处理的特点,即对于二维数组的运算可以仅仅通过一次以光速进行的并行运算来完成。因此这种运算的高速度是无以伦比的。

我们知道,在应用计算机求解线性方程组、线性微分方程,以及解卷积运算和线性变换等问题时,都必须将问题化为矩阵运算来处理。因此,光学矩阵运算是光学计算中的基本运算之一。另外,光学运算的并行性特点又尤其适合进行矩阵运算。例如,两个 $n \times n$ 矩阵的相乘共需要 n^2 次乘法和 n^2 次加法运算。如果用电子计算机运算,则需要 $2n^2$ 次计算操作,而利用光学计算去完成,最多只需要 n^2 次计算操作,有的甚至一次就可完成。因此,从七十年代开始,国外许多专家、学者在光学矩阵运算领域内进行了大量的研究工作,发表了许多论文。近年来光学矩阵运算已成为光学领域中的热门课题,但是这项工作的研究在国内所见甚少。为此,本文试图把近年来国外有关光学矩阵运算的研究情况向读者做一介绍,以促进这项工作的研究和发展。

光学矩阵运算主要是指利用光学方法实现下式所示的运算关系:

$$AB = C, \quad (1)$$

其中 A, B 分别为 $m \times L$ 和 $L \times n$ 矩阵, C 是 $m \times n$ 矩阵。在数学上矩阵 C 的运算可通过内乘和外乘两种方法进行。用内乘法计算时,矩阵 C 的 i 行 j 列矩阵元由下式求出:

$$C_{ij} = \sum_K A_{iK} B_{Kj}, \quad (2)$$

其中 $i = 1, 2, \dots, m, j = 1, 2, \dots, n, K = 1, 2, \dots, L$; 用外乘法计算时,应先计算秩为 1 的矩阵 $C^{(K)}$:

$$C^{(K)} = \begin{pmatrix} a_{1K} \\ a_{2K} \\ \vdots \\ a_{mK} \end{pmatrix} (b_{K1} b_{K2} \dots b_{Kn}) \\ = \begin{bmatrix} a_{1K} b_{K1} & a_{1K} b_{K2} & \dots & a_{1K} b_{Kn} \\ a_{2K} b_{K1} & a_{2K} b_{K2} & \dots & a_{2K} b_{Kn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mK} b_{K1} & a_{mK} b_{K2} & \dots & a_{mK} b_{Kn} \end{bmatrix}, \quad (3)$$

然后对 L 个 $C^{(K)}$ 求和,得到矩阵 C ,

$$C = \sum_{K=1}^L C^{(K)}. \quad (4)$$

目前,用内乘法进行矩阵运算的非相干光学系统有两种类型:一类是用掩膜实现数据-信号转换的矩阵运算系统,其数据可以完全并行地输入。例如,由 J. W. Goodman 提出的完全并行的矩阵-矢量相乘系统^[1]和 Hidetoshi Nakano 提出的完全并行的多个矩阵的相乘系统^[2];另一类是由 H. J. Caulfield^[3]发展的用声光调制器实现数据-信号转换的矩阵运算系统,后来由 Speiser^[4]和 Casasent^[5]等人进一步发展。在这种系统中,数据是以脉动和啮合方式

输入的。由于这种系统具有比较好的性能，所以很受重视。用光学方法模拟(3)和(4)式运算过程的外乘光学矩阵运算系统由 Athale 等人提出^[6,7]，这种系统以它简单的光路和较快的运算速度引起人们的注意。另外，还有 Bocker 提出的双光路电光转换矩阵运算系统^[8]。为了应用矩阵运算解决线性方程组问题，Psaltis 首先在矩阵-矢量相乘的掩膜系统中引入了电子反馈回路^[9]。此后，Casasent 又把电子反馈回路引入矩阵运算的声光调制系统^[5,10]，解决了解卷积运算、解线性微分方程和矩阵求逆等问题。

在用光强直接表示矩阵元的矩阵运算系统中，一方面，由于系统的线性区域有限，使得参加运算的各元素的数值变化范围受到限制；另一方面，利用这种元素表示法将无法提高系统的运算精度，最高只能达 8—10bit^[11]。为克服这两方面缺点，尤其是后者，人们引用了二元编码^[12-14]和 2 的补码技术^[11,15-17]，使运算精度提高到 32bit^[18]。

由于上述内容较多，我们将分三篇文章进行介绍，第一篇文章介绍内乘矩阵运算的掩膜光学系统、外乘算法系统和内乘双光路系统。第二篇介绍矩阵运算的声光调制系统和带有反馈回路的矩阵运算系统。第三篇介绍二元编码和 2 的补数编码方法及其各种矩阵运算系统。

二、矩阵运算的掩膜光学系统

矩阵用掩膜输入是光学矩阵运算中最先使用的最基本的方法之一。1974 年，Bocker 利用扫描装置和矩阵的掩膜输入法设计了矩阵-矢量相乘的光学系统^[19]。1975 年，Schneider 等设计了用掩膜输入的矩阵-矩阵并行相乘系统^[20]。由于该系统需要把一个凸透镜横切成几块，并依次相互错位放置，因而使实际制作十分困难。1978 年，Goodman 提出了比较理想的矩阵-矢量相乘系统^[1]，只使用凸透镜和柱透镜。1984 年，Yin-zhong Liang 等人利用多焦点全息透镜给出了完全并行的矩阵-矩阵运算系统^[21]，但其中一个矩阵必须表示成 N^3 形式，

物理

因而增加了额外的空间信道。1985 年，Hidetoshi Nakano 等^[2]克服了上述某些系统的弊端，巧妙地利用矩阵的啮合形式，仅使用一般的凸透镜和柱透镜就可实现任意多个矩阵的相乘运算。因此，该系统是目前比较理想的矩阵运算的掩膜光学系统。

1. 矩阵-矢量相乘系统^[1]

Goodman 实现矩阵-矢量相乘的装置如图 1 所示，其中 LEDS 是发光二极管阵列，透镜 L_1, L_2, L_3, L_4 的焦距都为 F ， L_4 是柱透镜组， L_5 由焦距为 F 的柱透镜和凸透镜组成， D 是垂直排列的探测器阵列， M 是矩阵掩膜。

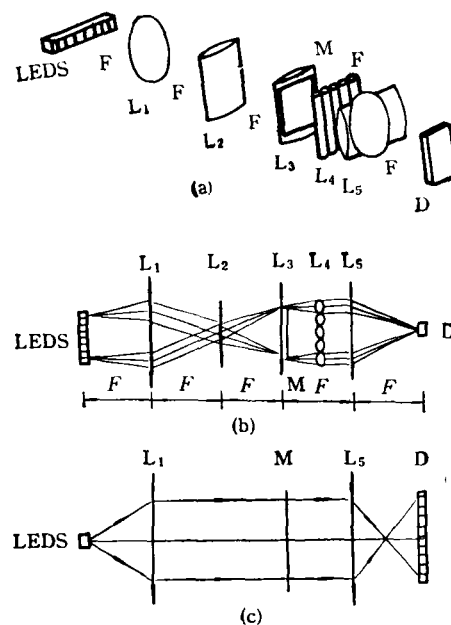


图 1 矩阵-矢量相乘的掩膜系统
(a) 立体图；(b) 俯视图^[1]；(c) 正视图^[1]

当利用该系统实现 $M \times N$ 矩阵 A 与 N 维矢量 B 相乘时，通过控制 N 个发光二极管的强度把矢量 B 的 N 个元素并行输入，透镜 L_1, L_2 使 LED 在水平方向上成象在矩阵掩膜 M 上，而每个 LED 发出的光又在竖直方向散开，均匀地照在掩膜上竖直方向的对应列上；矩阵掩膜 M 由 $M \times N$ 个子单元组成，每个子单元是一个透光量与矩阵 A 的元素 a_{mn} 成正比的光孔； L_3 是场镜； L_4 的作用是用来增强光的效率；透镜 L_5 把从水平方向的某行上的出射光会聚到竖直排

列的某个对应的探测器上, 这样第 m 个探测器接收的光强就成为输出矢量 C 的第 m 个分量值, 它与 $\sum_{n=1}^N a_{mn}b_n$ 成正比, 由此完成了矩阵-矢量相乘运算 $C = AB$.

Goodman 等人以 10 个点的离散傅里叶变换为例, 在此系统上进行实验, 实验结果与理论值基本符合. 由于傅里叶变换是在复数域上进行的, 而上述系统只能进行正实数运算(用光强只能表示正实数矩阵元), 所以应采用复数的正实数表示法^[22]. Goodman 用三个正实数表示一个复数, 这样 10 个点的离散傅里叶变换变成了 30×30 矩阵与 30 维矢量的相乘运算.

2. 多个矩阵相乘系统^[2]

Hidetoshi Nakano 等人提出的完全并行的多个矩阵相乘系统装置见图 2(a), 图 2(b), (c) 分别是光路的俯视图和正视图, 图中给出的是实现三个 3×3 矩阵相乘的示意图. 其中 SL 是凸透镜, CL 和 CL₂ 是柱透镜, SL 和 CL 的焦距都是 f , CL₂ 的焦距为 $f/2$, 箭头方向所指的是透镜的会聚方向, D 是单向漫射体, E 是由三个发光二极管组成的单位矩阵, 矩阵

A, B, C 分别用三个透射率函数按空间点阵形式分布的掩膜来表示, 它们在掩膜上的排列形式如图 3 所示.

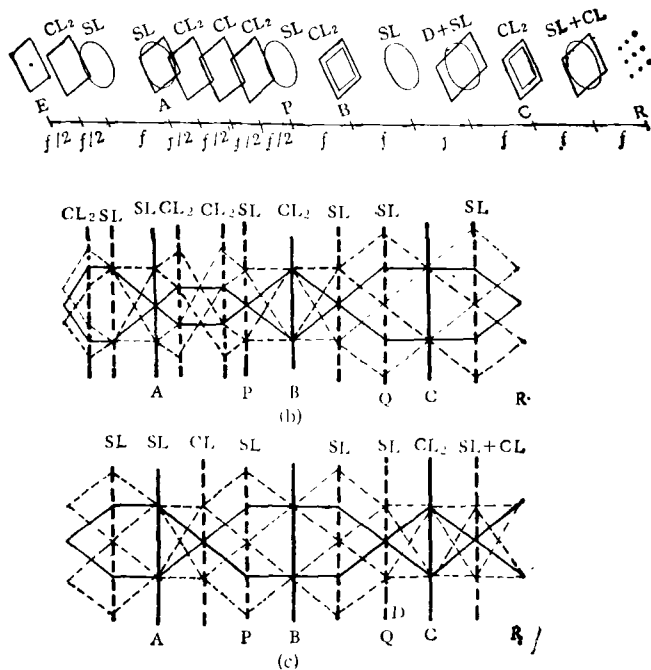


图 2 三个矩阵相乘系统
(a) 立体图; (b) 俯视光路图^[23];
(c) 侧视光路图^[23]

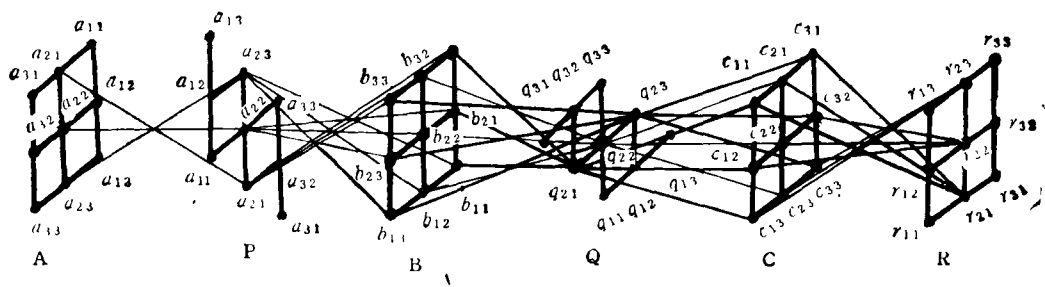


图 3 三个矩阵相乘过程示意简图

为了分析三个矩阵相乘系统的实现过程, 图 3 给出了 A, P, B, Q, C, R 平面上矩阵元素的排列方式和矩阵元的传递方向. 其中 A, B, C 平面分别为三个相乘矩阵的输入平面, 它们都是正常排列形式; P 平面是矩阵 A 经变换后的排列形式, 称为垂直方向的啮合排列形式; Q 平面是 A, B 相乘后的输出矩阵平面, 得到水平方向啮合排列形式的矩阵; R 是 A, B, C

三个矩阵相乘的输出平面, 得到正常排列形式的矩阵. 不妨把此系统分成三部分: 第一部分是从 E 平面到 P 平面, 称为 EP 单元, 它的作用是把正常排列形式的矩阵变成垂直方向啮合排列形式的矩阵; 第二部分是从 P 平面到 Q 平面, 称为 PQ 单元, 它是实现两个矩阵相乘的子系统, 其输入、输出矩阵的排列形式见图 3; 第三部分是从 Q 平面到 R 平面, 称为 QR 单元, 此

表 1

系统类别	性能	操作方式
Bocker 矩阵-矢量相乘系统 ^[19]	相对误差平均值为 5%	部分并行
Schneider 矩阵-矩阵相乘系统 ^[20]	对于两个 50×50 矩阵相乘, 相对误差不超过 1%	完全并行
Goodman 矩阵-矢量相乘系统 ^[21]	运算速度为 3×10 ⁹ 次复值运算/s 实验结果和理论值的符合情况见图 2	完全并行
Ying-zhong Liang 矩阵-矩阵相乘系统 ^[22]	最大相对误差为 0.9%, 平均相对误差为 0.4%	完全并行
Hidetoshi Nakano ^[23] 多矩阵相乘系统	对于三个 3×3 矩阵相乘, 最大相对误差为 9.0%, 平均相对误差为 4.7%; 对于两个 3×3 矩阵相乘, 最大相对误差为 6.0%, 平均相对误差为 1.9%	完全并行

单元也是实现两个矩阵相乘的子系统, 它的输入、输出矩阵的排列形式亦见图 3。从图 2(b) (c) 和图 3 不难看出, 若将 P 平面的矩阵排列形式转 90° , 正是 QR 子系统所要求的矩阵输入形式; 而将 Q 平面的矩阵排列形式转 90° , 则正是 PQ 子系统所要求的矩阵输入形式。由此可见, 如果把 EP 单元与转 90° 的 QR 单元相连, 则是输入、输出矩阵都为正常排列形式的两个矩阵相乘系统; 如果把 $EP + PQ$ 与转 90° 的 $PQ + QR$ 相连, 则是输入、输出矩阵都为正常排列形式的四个矩阵相乘的系统; 同样, 不难用 EP, PQ 和 QR 这三个基本单元组成任意多个矩阵相乘的系统, 这正是本系统的突出特点。

3. 各种掩膜系统的性能比较

各种掩膜系统的性能比较如表 1 所示。

三、外乘系统及其它

1. 外乘光学矩阵运算^[6,7]

用外乘法实现两个矩阵相乘的光学系统由 Athale 等人提出。下面以两个 4×4 矩阵相乘为例, 说明两种外乘法实现矩阵-矩阵相乘的光学系统, 如图 4, 5 所示。图 4 中 P_1, P_2 均为电光调制器, P_3 为探测器点阵。三者实际上紧贴在一起放置, 本图为便于说明而使呈分离状态。图 5 中, P_1 是发光二极管阵列, P_2 是电光调制器, P_3 是探测器阵列。图 4, 5 中, 如果 A

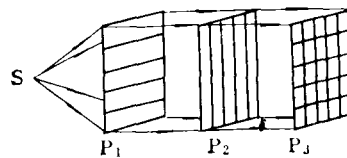


图 4 电光-电光调制矩阵外乘系统示意图^[6]

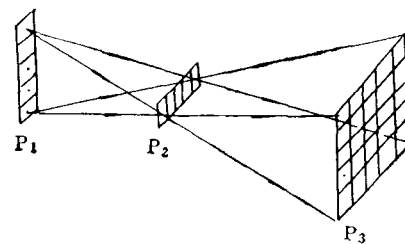


图 5 光源-电光调制矩阵外乘系统示意图^[6]

矩阵的列元素和 B 矩阵对应的行元素分别从 P_1, P_2 上并行输入 (与此同时, 图 4 中需用一均匀分布的光脉冲照在 P_1 调制器上), 则在探测器点阵 P_3 上得到

$$C^{(K)} = \begin{pmatrix} a_{1K} \\ a_{2K} \\ a_{3K} \\ a_{4K} \end{pmatrix} (b_{K1} b_{K2} b_{K3} b_{K4})$$

$$= \begin{pmatrix} a_{1K} b_{K1} & a_{1K} b_{K2} & a_{1K} b_{K3} & a_{1K} b_{K4} \\ a_{2K} b_{K1} & a_{2K} b_{K2} & a_{2K} b_{K3} & a_{2K} b_{K4} \\ a_{3K} b_{K1} & a_{3K} b_{K2} & a_{3K} b_{K3} & a_{3K} b_{K4} \\ a_{4K} b_{K1} & a_{4K} b_{K2} & a_{4K} b_{K3} & a_{4K} b_{K4} \end{pmatrix} \quad (5)$$

将 A 的列元素和 B 的行元素依次从 P_1, P_2 输入, 则在输出探测器点阵上得到四个如(5)式形

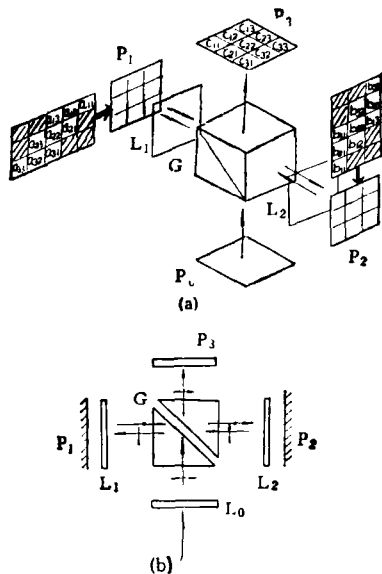


图6 双光路矩阵-矩阵相乘示意图^[4]
(a) 立体图; (b) 侧视图

式的矩阵,这些矩阵累加的结果就是 A, B 之积 C 矩阵,即

$$C = \sum_{K=1}^4 C^{(K)}$$

另外,若把图6中的发光二极管列阵和电光调制器换成两个声光调制器,同样也可实现外乘矩阵运算,原理是类似的,详情可参见文献[6, 7].

2. 双光路电光调制矩阵-矩阵相乘系统^[8]

由 Bocker 提出的双光路电光矩阵-矩阵相乘系统,其特点是光束沿两个相互垂直的方向行进.系统的装置如图6(a)所示.图6(b)是系统光路的侧视图.

图6中 P_1, P_2 是反射光阀(它的作用是使

照射在它上面的光强度被调制后按原路反射回去), L_0 是起偏器, L_1, L_2 是 $1/4$ 波片, P_3 是探测器点阵, G 是格兰棱镜. 准直的均匀平行光经起偏器 L_0 后,产生偏振方向与纸面平行的偏振光,该方向的偏振光不能穿过格兰棱镜而被反射到 P_1 上,经光阀调制后又反射回格兰棱镜. 由于在入射和反射过程中两次通过 $1/4$ 波片 L_1 , 所以从 P_1 射向棱镜的光的偏振方向旋转 90° , 正好垂直于纸面. 这个方向的光能够顺利地通过格兰棱镜,透过 L_2 后照射在光阀 P_2 上,经 P_2 调制后又按照原路返回,第二次通过 $1/4$ 波片 L_2 , 偏振方向又变得与纸面平行. 于是光束被格兰棱镜反射,到达探测器点阵 P_3 上. 如果矩阵 A 元素和矩阵 B 元素分别从 P_1, P_2 以啮合形式输入[见图6(a)], 则在探测器点阵上就可得到输出矩阵 C .

- [1] J. W. Goodman et al, *Opt. Lett.*, 2(1978), 1.
- [2] Hidetoshi Nakano et al, *Appl. Opt.*, 24(1985), 4238.
- [3] H. J. Caulfield et al., *Opt. Comm.*, 40(1981), 86.
- [4] J. M. Spoiser, *Proc. SPIE*, 298(1981), 41.
- [5] D. Casasent et al., *Appl. Opt.*, 22(1983), 115.
- [6] R. A. Athale et al., *Appl. Opt.*, 21(1982), 2089.
- [7] R. A. Athale et al., *Proc. IEEE*, 72(1984), 931.
- [8] R. P. Bocker et al., *Appl. Opt.*, 22(1983), 804.
- [9] D. Psaltis et al., *Opt. Lett.*, 4(1979), 348.
- [10] D. Casasent et al., *Appl. Opt.*, 21(1982), 1859.
- [11] R. P. Bocker et al., *Appl. Opt.*, 22(1983), 2019.
- [12] R. A. Athale et al., *Appl. Opt.*, 22(1983), 368.
- [13] P. S. Guilfoyle, *Opt. Eng.*, 23(1984), 20.
- [14] D. Casasent et al., *Appl. Opt.*, 24(1985) 1477.
- [15] R. P. Bocker, *Opt. Eng.*, 23(1984), 26.
- [16] A. P. Goutzoulis, *Appl. Opt.*, 23(1984), 4095.
- [17] B. K. Taylor et al., *Appl. Opt.*, 25(1986), 956.
- [18] W. T. Rhodes et al., *Proc. IEEE*, 72(1984), 820.
- [19] R. P. Bocker, *Appl. Opt.*, 13(1974), 1671.
- [20] W. Schneider et al., *Opt. Acta*, 22(1975), 879.
- [21] Yin-zhong Liang et al., *Opt. Lett.*, 9(1984), 322.

1988 年第 11 期《物理》内容预告

穆斯堡尔谱学在固体物理中的应用(郑裕芳);非晶半导体超晶格研究中的几个问题(于振瑞等);II-VI 族化合物的光致发光(董志武);应用极化的原子束测量表面的电场梯度——自旋极化的原子核表面谱法(叶慧);日本红外辐射陶瓷材料的最新进展(汤大新等);圆桌会议:美国工业中的物理学研究(陈康琪译);一种新的光谱学技术——光热释电光谱 (PPES) 及其应用(肖定全);半导体超薄层外延技术(彭英才);干涉法测量晶体的压电系数(尹鑫等);MT-1 型真空摩擦磨损试验机简介(张人信);质子和电子对薄膜聚合物改性的研究(王广厚);光学矩阵运算(II)(陈岩松);我国第一部物理学——《远西奇器图说》(王政一)。