



电子计算机

在情报工作中的应用

科学技术文献出版社

37.216
597
C.3

电子计算机在情报 工作中的应用

曾民族 高崇谦 编译

37.216/597

科学技术文献出版社



电子计算机在情报工作中的应用

编辑者：中国科学技术情报研究所

出版者：科学技术文献出版社

印刷者：北京印刷三厂

新华书店北京发行所发行 各地新华书店经销

开本：850×1168 1/32 印张：9.25 字数：247千字

1980年5月北京第一版第一次印刷

印数：1—16,800册

科技新书目：140—19

统一书号：15176·458 定价：0.95元

出版说明

情报工作中应用电子计算机，已经在国际上广泛发展和普及。目前我国情报部门和图书馆界也正在大力研究采用电子计算机。为了介绍国外电子计算机在情报工作中应用的情况，我们编译出版了《电子计算机在情报工作中的应用》一书。这本书基本上出自日本《情报管理》期刊1973年连载的“计算机和它的应用”讲座，同时收集了日本科学技术情报中心电子计算机系统概况等有关文章。书中内容既有电子计算机在情报工作中应用的一般知识，又有情报系统设计的原理和实例，可供情报部门和图书馆界从事电子计算机设计和应用的同志们参考。

本书由曾民族、高崇谦编译。

科学技术文献出版社

1979年6月

35295

目 录

一、电子计算机的一般知识	(1)
二、电子计算机在情报管理中的应用	(22)
三、情报系统设计概论	(49)
四、资料档设计实例	(66)
五、市售资料档应用设计实例	(79)
六、程序设计概论	(99)
七、情报检索程序	(114)
八、电子计算机的操作	(162)
九、输入输出的管理	(174)
十、日语文献的电子计算机处理	(184)
十一、日本科技情报中心文献检索磁带的特征及其 使用方法	(205)
十二、日本科技情报中心文献检索磁带用的计算机 程序	(231)
十三、日本科技情报中心电子计算机系统概况	(241)
十四、日本科技情报中心联机情报检索系统	(256)
十五、日本科技情报中心汉字联机情报检索系统 (JOIS-K)	(275)

一、电子计算机的一般知识

(一) 电子计算机能完成的工作

人类在历史发展过程中制造了各种机器和工具，大致可分两类：

第一类是完成人本身不能做的工作的。木锯就是一个例子，因为人没有具备进行这种工作的器官。另一类是完成人能够做的工作的，但效率比人高得多。例如，一把铁锤就比人手更容易把对手打倒。

电子计算机看起来像前者，实际却属于后者。从本质上说，人不能做的工作，电子计算机也不能做。但是，人们需要几年才能解出的计算题等，电子计算机能以极高的速度而且非常准确地解出。总之，人能做的事，电子计算机也能做，而且效率比人高得多，尽管不是所有的事都如此。

电子计算机能做的工作，基本上可归纳为以下五类：读、写、算、记、存。

“读”、“写”就是输入、输出，主要是通过机械实现的，处理速度以每字千分之一秒（毫秒）为单位计算。

“算”就是运算，即进行加减乘除的算术运算和大小同异等比较的逻辑运算。各种运算是通过电子技术实现的，处理速度以每字百万分之一秒（微秒）为单位计算。

“记”、“存”就是内存和外存，都是通过电磁手段实现的。但是，前者是通过电路检索存储区的，处理速度以每字百万分之一秒为单位计算，而后者是通过机械实现的，处理速度以每字千分之一秒为单位计算。

(二) 输入装置

由于电子计算机目前还不能直接识别人类语言或任意书写的文字，所以在输入前必须把数据转换成机器可读的形式。主要使用键盘穿孔机将字母、数字或符号转换成穿孔的卡片（图1）或穿孔的纸带（图2）。在读入之前，穿孔卡片或穿孔纸带必须利用键盘验孔机等根据原始资料进行检验。

穿孔卡片通过卡片读出器，穿孔纸带通过纸带读出器，根据孔的有无转换成电流的有无，即转换为电脉冲列（图3），这就实现了计算机的读入。

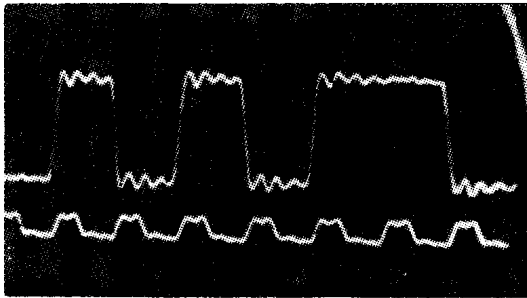


图3 电脉冲列

表1 主要输入装置

载体	数据表示方式	设备名称	速度
卡片	孔眼组合	卡片读出器	600~1,500张/分
纸带	孔眼组合	纸带读出器	200~1,000字/分
传票	特定字体的文字、 数字、符号和标志	光学符号读出器	200~400张/分

最近研制出的光学符号读出器 (OCR), 可以读出 特定 字体的文字或固定位置的标志, 很有利于节省从原始传票转换为卡片或纸带的手工作业时间。

此外, 还可利用打字机或联机系统, 通过通信线路从不同的终端装置采用相应的方法进行输入。各种主要输入装置如表 1 所示。

(三) 数据的表示

数据通过输入装置转换成电脉冲列。所谓电脉冲, 是指对电压或电流的大小并无意义、而只是对某一规定值在某一时刻超过或等于零时才有意义的电的流动。这种在离散上才有意义的表示, 称为数字表示。

一个脉冲按其有或无只能表示两种状态。将此作为信息量的最小单位, 称为 1 位。两个脉冲即 2 位, 可以表示 4 种状态, 三个脉冲即 3 位, 可以表示 8 种状态, 一般地说, n 个脉冲即 n 位, 可表示 2^n 种状态。

0 ~ 9 的 10 个数字, 由于 $2^3 < 10 < 2^4$, 故用 4 位就能表示。如用 6 位, 则可表示 $2^6 = 64$ 种字母、数字和符号。

目前采用的所谓 EBCDIC, 以 8 位表示 1 个字母、数字和符号, 尤其是 1 位数值数字可用 4 位表示 (表 2)。就是说, 一个字母可用两位数值数字表示。此时, 表示一字母的 8 位称为 1 位组。数值既可采用 8 位或 4 位表示一位数, 列出任何位数, 也可采用二进制方法表示。所谓二进制方法, 即 0, 1, 2, 3, ……变成 0, 1, 10, 11, ……。换句话说, 十进制方法是逢 10 进一位, 而二进制方法则是逢 2 进一位 (见表 2)。二进制方法本来不很普及, 但为了简化运算装置的结构, 1940 年根据维诺的提议, 在电子计算机中广泛采用。二进制方法是电子计算机内部问题, 用户不详细了解也无妨。

表 2

EBCDIC 实 例

英 文 字 母		日 语 假 名		符 号		文 字 数 字	
A	11000001	ア	10 00 00 01	•	01 00 10 11	0	11 11 00 00
B	11000010	イ	10 00 00 10	*	01 01 11 00	1	11 11 00 01
Z	11101001	ン	10 11 11 01	Y	01 01 10 11	9	11 11 10 01
数 值 数 字		二 进 制 表 示 的 数 字 (16 位 长)					
0	0 0 0 0	0	0000000000000000	10	0000000000001010		
1	0 0 0 1	1	0000000000000001	11	000000000001011		
9	1 0 0 1	9	0000000000001001	100	0000000001100100		

(四) 存 储 器

输入的数据必须存储在某一规定的位置上, 一旦需要时能够很方便地使用。因此, 需要设有存储器。

存储器由数千万字节的存储单位组成。为了识别这些存储单位, 每 1 字节应编好地址, 例如 0 地址, 第 1 地址, 第 2 地址…。根据机种不同, 有的以几个字节或几位组成 1 字, 每字设有地址。总之, 数据按地址存储在规定的位置上, 并从规定的位置取出。存入和取出所需时间, 称为存取时间。

将数据存储在给定的存储位置时, 该位置原来存储的内容即自行清除而存上新的数据。但是如果只是再生数据, 则该存储位置中的内容不会消失。

目前所用的存储器是磁芯存储器。如图 4 所示, 磁芯存储器把外径不足 1 毫米的铁氧体环借助导线穿成网状, 所需存储位置选在交叉的纵横线 (写入线), 一旦输入脉冲, 处在交点上的磁芯便被磁化。也就是说, 通过输入 1 个脉冲使 1 个磁芯磁化来实现存储。此外, 如图 5 所示, 磁芯斜向还通有导线 (读出线), 现在反过来使交点上的磁芯磁化方向相反, 结果在读出线上便激励起一个脉

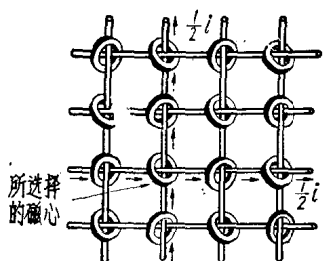


图4 选择与存储结构

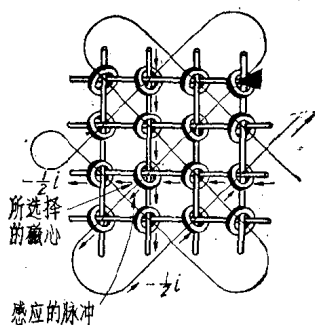


图5 再生结构

冲，这就是实现再生。此时，由于记忆消失，实际上等于把同样信息再次写入。将8个磁芯网重叠，则在连结磁芯的垂直方向上可存储1字节表示的字母、数字或符号(图6)。每一字节的存取时间以微秒计算。

(五) 运算器

存储器的数据，根据处理的需要再生(取出)后，要在运算器内进行运算加工。如上所述，电子计算机进行的运算是算术运算和逻辑运算，积分等运算不能直接进行，而是通过乘法和加法组合运算方法进行的。

在算术运算中，乘除运算可通过反复加减并‘移位’的办法实现(图7)。而且，减法运算使用补数，也可变为加法运算。例如，72减48，先加上使48的各位都成为9的数51(51就是48各位9的补数)，即 $72 + 51 = 123$ ，去掉最前面的1，加在最后一位，即 $23 + 1$

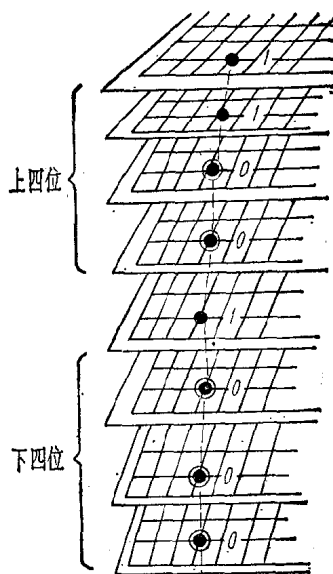


图6 字母H的存储状态

0		332		
+ 12	} 首先加 4 次	第 1 次..... - 13		
12		20		
+ 12		第②次..... - 13		
24		(商的最高位) 72	至此不能再减, 故向 右移一位	
+ 12	} 首先加 4 次	第 1 次..... - 13		
36		59		
+ 12		第 2 次..... - 13		
48		46		
+ 12	} 首先加 4 次	第 3 次..... - 13		
168		33		
+ 12		第 4 次..... - 13		
288		20		
+ 12	} 然后移 1 位, 再加三次	第⑤次..... - 13		
408		(商的最低位) 7.....	余数	
12 × 34 = 408 的计算		332 ÷ 13 = 25 余 7		

图 7 通过加减法完成乘除运算的方法

=24。从上述可见，只要采用加法运算方式，便可以进行任何算术运算。

在电子计算机中，这种运算是借助电子电路（加法电路）完成的，就是说，向输入端输入表示两个数值的脉冲，而在输出端输出表示其和的脉冲。此时，逐位输入采用10进制方法时，就必须对 $10 \times 10 = 100$ 种组合考虑其和，但采用2进制方法时，则 $2 \times 2 = 4$ 种即够，运算机构就可以简单得多。正因为如此，在电子计算机中均采用二进制运算方法。

大小同异的比较即逻辑运算，根据同样道理也可以通过电子电路（比较电路）完成。例如，在输入端输入表示两个数据的脉冲，那么，如果相同，则在输出端有脉冲，如果不同，则在输出端无脉冲，通过这种方法实现同异的比较。

最初，这种电路是用继电器组成的，由于是机械动作，速度较慢。后来，采用了相当于电子开关作用的真空管，但是由于耗电

高、发热量大、寿命短等原因，又改用了晶体管或半导体二极管。其后，为了提高速度或可靠性，又采用了集成电路。在只有1.5平方毫米的基板上，可以制成20个晶体管和10个电阻的集成电路。尤其最近，已研制出在一个基板上可装两台电视机电路的大规模集成电路，并已开始实用（图8）。

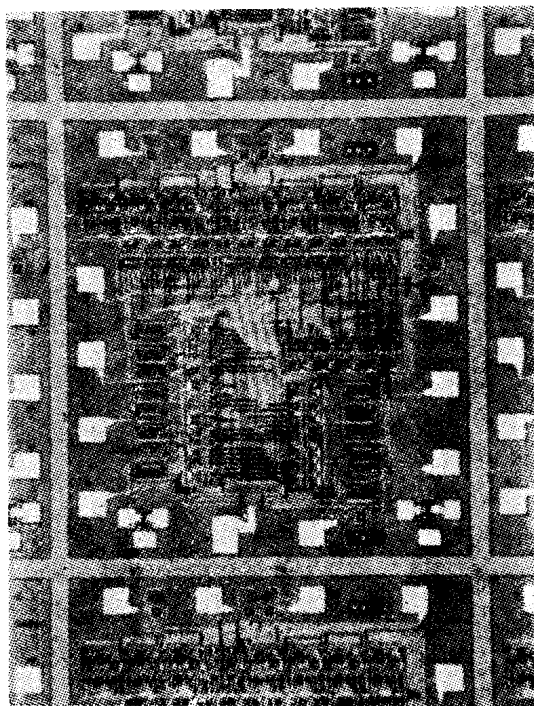


图8 大规模集成电路

第1台电子计算机 ENIAC，使用了18,800个真空管，1946年在宾夕法尼亚大学首次运行。该机的运算速度是：10位数的加减法为1/5000秒，乘法为1/300秒，除法为1/50秒。由人作9位数乘法计算的时间是5分钟，验算12分钟，合计需17分钟（约1000秒）。这就是说，由计算机 ENIAC 作乘法计算，其运算速度比人工高30

万倍。

如果采用 HITAC 8700之类最新电子计算机对相当于10进制的16位数进行计算时,其加减法的运算速度为0.84微秒,乘法为3.99微秒,除法为7.25微秒。这样,电子计算机的运算速度(包括判断)可从微秒单位提高到毫微秒单位(十亿分之一秒)。

(六) 输出装置

在电子计算机内所求得的处理结果,可通过输出装置记录在所需载体上。

各种处理结果可用卡片穿孔机打在卡片上,或用纸带穿孔机打在纸带上,作为输入数据再次使用。

输出的一般形式是印字。输入、输出打字机是逐字打印,行式打印机是逐行打印(每行120~160字)。

此外,还有以下几种输出装置:利用电子计算机输出信号驱动笔描绘曲线或图纸的X-Y坐标绘图仪和类似电视机的阴极射线管显示文字或图像的显示装置(图9)等。在联机系统中还可以通过通信线路向各种终端装置输出。现用的主要输出装置如表3所示。

表3 主要输出装置

载体	数据的表示方法	装置名称	速度
卡片	孔眼组合	卡片穿孔机	100~300张/分
纸带	孔眼组合	纸带穿孔机	50~300字/秒
纸	字母、数字、符号	打字机	10~15字/秒
纸	字母、数字、符号	行式打印机	300~1200行/分
纸	字母、数字、符号、图形	X-Y坐标绘图仪	350~900步/秒
萤光屏	字母、数字、符号、图形	显示装置	100~300字/秒 2000~5000米/秒

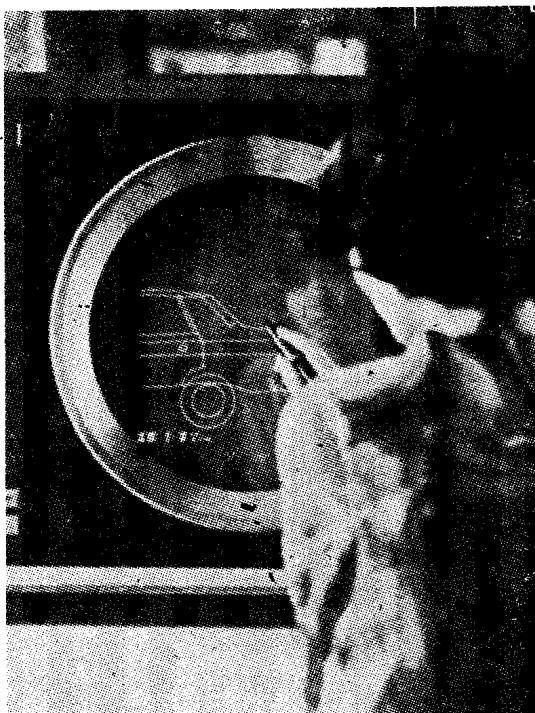


图9 显示装置

(七) 外部存储器

由于磁芯存储器无法（也无必要）容纳账目之类的大量数据，因此，这类数据应存储在费用低廉而容量很大的外部存储器内，必要时才移到磁芯存储器进行处理。这种存储器有两种：磁带存储器和随机存取存储器。

磁带存储器是按照磁带录音机的原理进行存储再生的。存储载体是一卷宽约1.3公分、长约700米、直径约25公分的磁带，可以更换。存储密度为30~60字节/毫米，每卷磁带的存储容量约为2,000万~4,000万字节。但是，由于磁带的起动和制动需要一定的数据间隔，所以实际的有效容量为一半左右。

磁带存储器存取（存储再生）数据时，必须从头按顺序检索到所需位置。因此，存取时间因存储位置不同而相差很大。为了克服这一缺点，还研制出随机存取存储器。这种存储器是根据地址进行查找的，所以无论存储位置如何，其存取时间大致相等。这种存储器的载体是磁鼓和磁盘。

磁鼓存储器是一个转动的涂有磁层的圆柱体，在转轴平行方向装有几个磁头，靠磁头与转角选择所需存储位置，以实现存储再生（图10）。

磁盘存储器是几个涂有磁层的同轴旋转的圆盘，将磁头像拾音器那样伸到磁盘表面上，找出所需的存储位置，以实现存储再生（图11）。

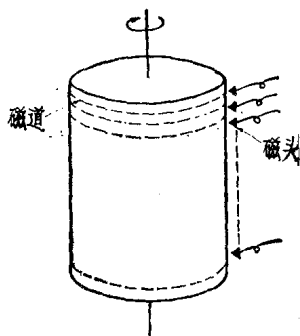


图10 磁鼓存储器的结构

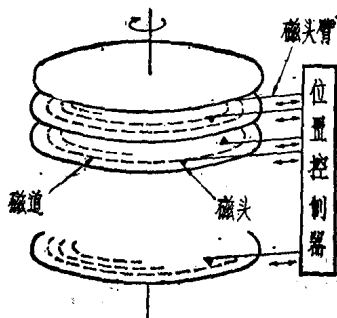


图11 磁盘存储器的结构

磁盘部分可更换的磁盘存储器称为磁盘组件。最近已研制出一种在一台驱动装置上可安装9个磁盘组件而且每个磁盘组件可同时实现存储再生的组合磁盘存储器。现用主要外部存储器如表4所示。

（八）程序存储方式

下面谈谈上节所述各种输入输出装置、运算器和存储器按图12所示接通开关后求出两个数据之和的过程。

表 4 主 要 的 外 部 存 储 器

设 备 名 称		容 量	存取时间	旋 转 速 度
磁 带 存 储 器		20—40兆字节/台	不 定	10~120千字节/秒
随 存 机 储 存 取 器	磁 鼓 存 储 器	1—4兆字节/台	10毫秒	100~200千字节/秒
	磁 盘 组 件 存 储 器	7 兆 字 节 / 台	100毫秒	150千字节/秒
	组 合 磁 盘 存 储 器	230兆字节/台	100毫秒	300千字节/秒

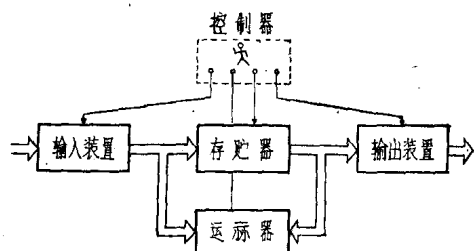


图12 人工控制的计算机

两个数据分别在两张穿孔卡片穿孔，然后装在输入装置（穿孔卡片读出器）上。按表 5 所示顺序，按下开关，两数据之和便由输出装置（行式打印机）输出。

按开关的顺序可以预先确定，如果编好工作顺序单（如表 5 的“开关”栏），则操作员

表 5 求出两个数据之和的处理次序

顺 序	开 关 名 称	处 理 内 容
1	A	读出第 1 张卡片，存储第一个数据
2	A	读出第 2 张卡片，存储第二个数据
3	B	两个数据在运算器内相加求其和
4	C	和数由运算器转入存储器
5	D	打印存储器的和数