高等学校计算机教材

汇编语言程序设计教程

罗省贤　洪志全　编著

- 8086/8088
- 80x86
- Win32 ASM
- 64位MMX
- 128位SSE
- 80x87

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

高等学校计算机教材

汇编语言程序设计教程

罗省贤　洪志全　编著

電子工業出版社
Publishing House of Electronics Industry
北京・BEIJING

内 容 简 介

汇编语言程序设计是计算机科学与技术专业学生的必修专业基础课，汇编语言的应用在系统软件开发、实时控制和实时处理领域中有着重要的地位。本书根据微型计算机和 32 位汇编语言程序设计技术的快速发展现状，在系统地介绍 8086/8088 基本结构、指令系统、编程方法、输入/输出和中断程序设计的基础上，进一步介绍 80x86 32 位微处理器的基本结构、指令系统、任务切换以及保护方式下的编程方法、用 Win32 汇编语言编写 Windows 窗口程序的方法、MMX 指令集、SSE 指令集及编程方法，以及 80x387 协处理器结构、数据格式、指令系统及其编程方法。

本书涉及的知识面广，内容丰富，易读性强，可作为高等院校汇编语言程序设计教材，还可作为广大工程技术人员的参考书。

图书在版编目（CIP）数据

汇编语言程序设计教程/罗省贤，洪志全编著. —北京：电子工业出版社，2004.1
ISBN 7-5053-9344-8

Ⅰ. 汇… Ⅱ. ①罗…②洪… Ⅲ. 汇编语言—程序设计—教材 Ⅳ.TP313

中国版本图书馆 CIP 数据核字（2003）第 111991 号

责任编辑：王昌铭
印　　刷：北京东光印刷厂
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　邮编　100036
经　　销：各地新华书店
开　　本：787×1 092　1/16　印张：21.25　字数：541 千字
印　　次：2004 年 1 月第 1 次印刷
印　　数：5 000 册　　定价：28.00 元

凡购买电子工业出版社的图书，如有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系。联系电话：（010）68279077。质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

前　言

汇编语言的显著特点是能够直接控制硬件并充分发挥计算机硬件的功能，对于编写高性能的系统软件和应用软件，汇编语言是最有效的语言之一。对于高等院校计算机硬件专业、软件专业以及计算机应用专业的学生，“汇编语言程序设计”是一门必修的核心课程。通过该课程的学习，能够使学生深入地理解计算机内部完成各种复杂操作和运算的基本原理。该课程对于培养学生掌握程序设计基本技能和调试技术也十分重要。

随着微型计算机技术的迅猛发展，Intel 公司推出的 80x86 和 Pentium 芯片已经成为微机中的主导处理器。80x86 是向上兼容的，功能强大，能寻址 4GB 的物理内存，具有 4 个特权级、多任务切换机制以及片内的内存管理单元等，这些功能都在保护方式下实现，可支持多用户、多任务操作系统。此外，32 位汇编语言程序设计技术也有了新的发展，为了使汇编语言程序设计教学适应当前计算机发展水平。我们新编了这本教材，在介绍 8086/8088 汇编语言的基础上，进一步较全面地介绍 80x86 的 32 位以及 64 位汇编语言程序设计方法。

全书共分 9 章。第 1 章至第 4 章全面介绍 8086/8088 的基本结构、寻址方式、指令系统及编程方法。第 5 章主要介绍输入/输出和中断程序设计方法。这几章都是学习汇编语言程序设计不可缺少的重要基础内容。第 6 章着重介绍 80x86 32 位微处理器的基本结构、指令系统、任务切换以及保护方式下的编程方法。第 7 章介绍用 Win32 汇编语言编写 Windows 窗口程序的方法和汇编语言的高级语法。第 8 章介绍 MMX、SSE 指令集及其功能和编程方法。第 9 章介绍 80x87 协处理器的结构、数据格式、指令系统及其编程方法。每章后均附有习题。本书程序实例丰富，增强了易读性，扩充了新知识。我们希望通过本书的学习，使读者能够掌握 32 位汇编语言程序的基本编程方法，提高在新型微机上的汇编语言程序设计能力。

本书适用范围较广，对于计算机科学与技术专业的专科生或本科生，可以重点学习第 1 章至第 5 章、第 7 章，熟练掌握基本的汇编语言程序设计方法，关于第 6 章和第 8, 9 章的内容，则可以根据课时安排有选择地学习。对于研究生则可重点学习第 6 章至第 9 章的内容，掌握 32 位微处理器和协处理器的汇编语言程序设计方法，熟悉保护方式下的编程方法。本书可以作为高等院校汇编语言程序设计教材，此外还可以作为广大工程技术人员的参考书。

本书由罗省贤主编，其中第 1, 2, 3, 4, 7 和 9 章由罗省贤编写，第 5 章由罗省贤和洪志全合写，第 6，8 章由洪志全编写。由于编者水平有限，书中可能会存在不少缺点与问题，恳请广大读者指正。

编　者

2003 年 10 月

目　录

第 1 章　Intel 8086/8088 的基本结构

汇编语言是计算机软件设计的重要工具，在系统软件开发、实时控制和实时处理领域中有着不可替代的地位。用汇编语言编程可以充分发挥计算机硬件的功能，进行高质量的程序设计，开发出的软件具有内存开销小，运算速度快的特点。因此，在已经有众多高级语言和可视化集成开发环境工具的今天，汇编语言仍是一门不可缺少的有效的程序设计语言。

由于汇编语言密切依赖计算机硬件系统，所以要掌握汇编语言程序设计，必须首先熟悉计算机的内部结构以及与机器结构紧密相关的指令系统。本章将先介绍 Intel 8086/8088 的基本结构，为下章学习指令系统打下基础。Intel 8086 是美国电子器件公司（Intel 公司）1978 年最早推出的 16 位微处理器，其内部结构和数据总线都是 16 位。20 世纪 70 年代末，由于 8 位机系统的外部设备和各种 8 位接口部件已被普遍采用，Intel 公司又推出了 16 位内部结构，8 位外部数据总线的准 16 位微处理器 8088。8086 和 8088 指令系统相同，支持 16 位算术运算逻辑，具有 20 位的地址总线，寻址空间为 1MB。

1.1　8086/8088 的功能结构

中央处理部件 CPU 的任务是执行存放在存储器里的指令序列。CPU 一般由运算器和控制器两部分组成，在微机中也常称为微处理器。8086/8088 CPU 按功能可分为两个独立的部分：总线接口单元 BIU（Bus Interface Unit）和执行单元 EU（Execution Unit），其功能结构如图 1.1 所示。

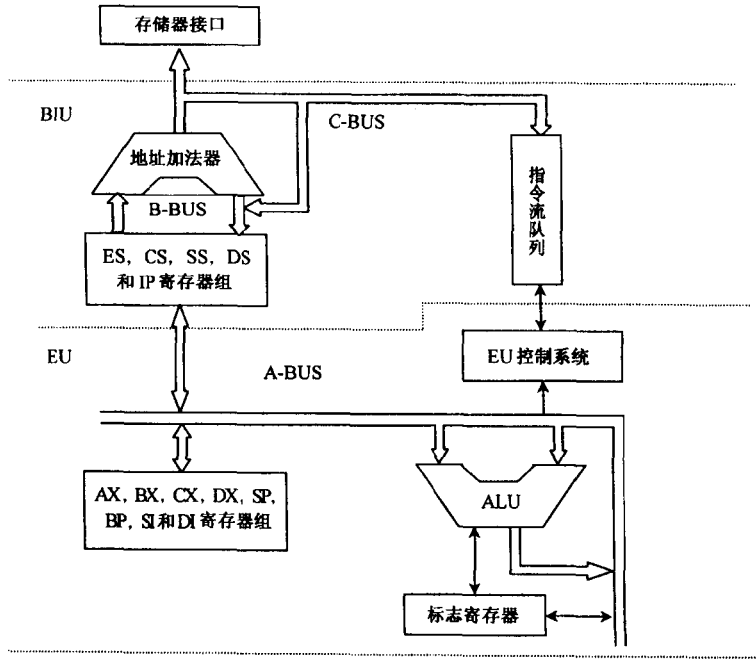

图 1.1　8086/8088 的功能结构

BIU 完成 8086/8088 与存储器之间的信息传送、总线控制和 I/O 数据传送，还包括逻辑地址与物理地址的转换；从存储器中取指令送指令流队列排队，取出执行指令时所需的操作数，并传送给 EU 完成运算和操作。

EU 负责对来自指令流队列中的指令译码并执行，实施算术逻辑运算操作。

由于 BIU 和 EU 是两个相对独立的部件，因此取指令和执行指令可以并行完成，形成指令流水线结构，如图 1.2 所示。指令流水线结构大大减少了 CPU 等待取指令的时间，提高了 CPU 的利用率和系统运行速度。

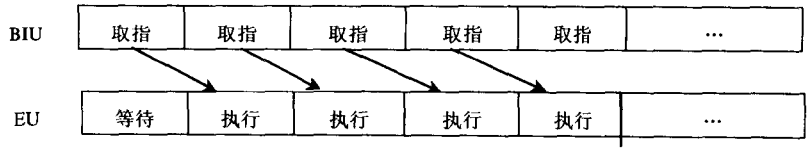

图 1.2　8086/8088 指令流水线结构

1.2　8086/8088 的寄存器结构

8086/8088 共有 14 个 16 位寄存器，一般可分为 4 类：数据寄存器、段寄存器、控制寄存器、指针及变址寄存器，如图 1.3 所示。这些寄存器具有重要的作用，专门用于存放指令执行时需要的各种信息，如操作数、操作数地址以及中间计算结果等。

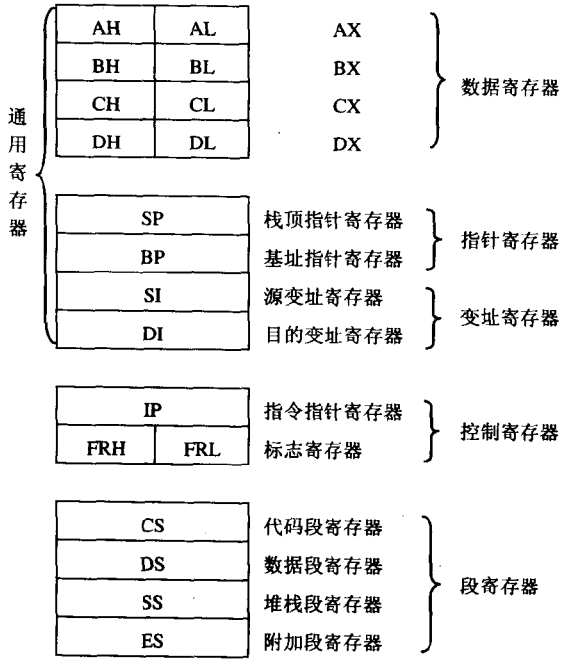

图 1.3　8086/8088 寄存器结构

1．数据寄存器

数据寄存器共有 4 个 16 位寄存器，即 AX、BX、CX 和 DX，通常用于暂存计算过程中的操作数、计算结果或其他信息。数据寄存器既可按字类型（16 位）访问，又可按字节类型（8 位）访问，不仅有上述一般用途，同时还有各自的专门用途。

AX（Accumulator）：累加器，是算术运算的主要寄存器，此外还作为乘、除法运算及I/O操作的专用寄存器。

BX（Base）：基址寄存器，常用于存放存储区的起始地址。

CX（Count）：计数寄存器，常用于循环操作和字串处理的计数控制。

DX（Data）：常与AX共同用于双字长运算，DX存放高位字，AX存放低位字。此外还用于在I/O操作时提供外部设备接口的端口地址。

2．段寄存器

段寄存器共有4个16位寄存器，即CS、DS、SS和ES，专门用于存放段地址。

CS（Code Segment）：代码段地址寄存器，存放代码段的起始地址。代码段用于存放当前正在运行的程序。

DS（Data Segment）：数据段地址寄存器，存放数据段的起始地址。数据段用于存放当前正运行的程序所用的数据。

SS（Stack Segment）：堆栈段地址寄存器，存放堆栈段的起始地址。堆栈是一种以后进先出方式访问的数据结构，堆栈段是定义堆栈的存储区。

ES（Extra Segment）：附加段地址寄存器，存放附加段的起始地址。附加段是附加的数据段，作为辅助数据区存放当前正运行程序所用的数据。

3．指针及变址寄存器

指针及变址寄存器共有4个16位寄存器，即SP、BP、SI和DI，主要用于在访问存储器单元时提供16位偏移地址，其中SI、DI也可以作为数据寄存器使用。（一个存储器单元的物理地址（实际地址）由段地址和偏移地址两部分构成，这一问题将在下节讨论。）

SP（Stack Pointer）：栈顶指针寄存器，提供堆栈栈顶单元的偏移地址，与SS段寄存器联用，以控制数据进栈和出栈。

BP（Base Pointer）：基址指针寄存器，常用于提供堆栈内某个单元的偏移地址，与SS段寄存器联用，可访问堆栈中的任一个存储单元。

SI（Source Index）：源变址寄存器，与DS段寄存器联用，可以访问数据段中的任一个存储单元。

DI（Destination Index）：目的变址寄存器，与ES段寄存器联用，可访问附加段中的任一个存储单元。

SI、DI也常用于在字串操作中提供偏移地址，并具有地址自动增量或减量的功能。

4．控制寄存器

控制寄存器共有2个16位寄存器，即IP和FR。

IP（Instruction Pointer）：指令指针寄存器，存放代码段中指令的偏移地址。在程序执行过程中，始终自动给出下一条要取的指令的偏移地址。IP与CS段寄存器联用，可以确定下一条要取的指令的物理地址，因此IP是很重要的控制寄存器，用于控制程序的执行流程。

FR（Flags Register）：标志寄存器，用于存放反映处理器和运行程序执行结果状态的控制标志和条件码标志。FR中共有9个标志位，其中6个是条件码标志位，3个是控制标志位，如图1.4所示。FR中的内容也可称为程序状态字（PSW, Program Status Word）。

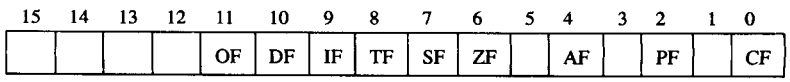

图1.4　标志寄存器

FR 中的 6 个条件码标志位用于记录程序运行结果的状态，一般由系统根据计算结果自动设置。这些标志常用做改变程序执行流程的控制条件，其含义如下。

OF（Overflow Flag）：溢出标志。在运算过程中，如果操作数超过了机器字长所限的数据表示范围，则称为发生溢出，此时 OF 位置 1，否则置为 0。

SF（Sign Flag）：符号标志。运算结果为负时，SF 位置 1，否则置 0。

ZF（Zero Flag）：零标志。运算结果为零时，ZF 位置 1，否则置 0。

CF（Carry Flag）：进位/借位标志。反映运算过程中数据最高有效位是否发生了进位或借位，若发生进位或借位，CF 位置 1，否则置 0。

AF（Auxiliary Carry Flag）：辅助进位/借位标志。运算过程中半字节有进位或借位时，AF 位置 1，否则置 0。

PF（Parity Flag）：奇偶标志。运算结果数的低 8 位中“1”的个数为偶数时，PF 位置 1，否则置 0。

FR 的 3 个控制标志位可在程序中使用指令将其置 1 或置 0，以控制 CPU 的操作方式，其含义如下。

DF（Direction Flag）：方向标志。用于在串操作中控制地址的变化方向。当 DF=1 时，每次操作后 SI、DI 中的地址自减，即串操作从高地址处理到低地址；当 DF=0 时，每次操作后 SI、DI 中的地址自增，串操作从低地址处理到高地址。

IF（Interrupt Flag）：中断标志。当 IF=1 时，允许中断；当 IF=0 时，关闭中断。有关中断系统的原理将在第 5 章介绍。

TF（Trap Flag）：单步标志（也称陷阱标志）。该标志控制 CPU 是否工作在单步操作方式下。TF=1 时，每执行一条指令后，CPU 产生陷阱异常，暂停执行正运行的程序。这种方式为程序员调试程序提供了方便，程序员可在每条指令执行后，查看 CPU 状态及指令执行结果。TF=0 时，CPU 正常工作，不产生陷阱异常。

FR 寄存器的标志位值还有一种符号表示方式。在 DOS 系统提供的汇编语言通信调试工具 DEBUG 中，可以显示 FR 的标志位。为了提高标志位值的可读性，DEBUG 中标志位的值采用了符号表示（见表 1-1），附录二给出了有关 DEBUG 程序使用方法的介绍。

表 1-1　FR 中标志位的符号表示

标志位	标志位名称	符号表示	
		=1	=0
OF	溢出标志（是/否）	OV	NV
DF	方向标志（减/增）	DN	UP
IF	中断标志（开/关）	EI	DI
SF	符号标志（负/正）	NG	PL
ZF	零标志（是/否）	ZR	NZ
AF	辅助进位标志（是/否）	AC	NA
PF	奇偶标志（偶/奇）	PE	PO
CF	进位标志（是/否）	CY	NC

1.3　堆栈与存储器结构

存储器是计算机存储信息的主要部件，程序代码、数据以及程序的运算结果都保存在存

储器中。堆栈则是存储器中的一块专用存储区，在操作系统和应用程序中都具有特殊的重要作用。本节介绍堆栈和存储器的结构及数据存取方式。

1.3.1 堆栈

堆栈是一种数据结构，实际上是在存储器中开辟的一端活动、另一端固定的数据存储区，数据的存取原则是先进后出。

堆栈的固定端称为栈底，活动端称为找顶，数据的存取都在栈顶进行，堆栈栈顶指针寄存器 SP 始终存放栈顶单元的地址，即 SP 始终指向栈顶，跟踪栈顶的变化，如图 1.5 所示。

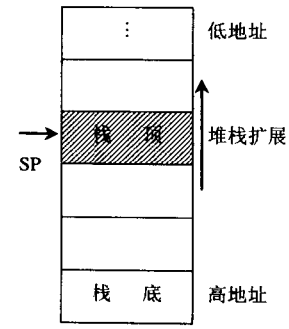

图 1.5 堆栈的结构

堆栈有着极其重要的作用，常用于保存计算和操作过程中的现场数据，保护子程序和中断服务程序返回地址等有关信息。

8086/8088 指令系统有两条指令（PUSH、POP）专门完成数据进栈和出栈操作，具体操作过程将在下一节指令系统中介绍。

注意，堆栈存取必须以字（2 字节）为单位，每次操作（数据进栈或出栈）后，SP 总是指向新的栈顶。随着进栈数据增多，堆栈将不断扩展，栈顶地址（SP）不断减小，因此堆栈是从高地址向低地址方向扩展的。

1.3.2 存储器结构

在存储器中，信息是以字节（8 个二进制位）为最小单位存储的，每一个字节单元分配一个惟一的存储器地址。地址从 0 开始编址，顺序增 1，用无符号二进制数表示，但为了方便书写，常用无符号十六进制数（数后加基数说明符 H 表示十六进制数）表示。例如，64KB 存储器空间的单元编址方法如图 1.6 所示。

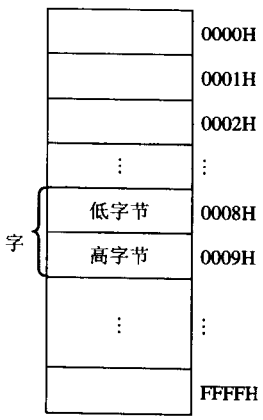

图 1.6 64KB 存储器单元编址

一个字数据在存储器中存于相邻的两个字节单元，数据的低字节存入低地址单元，高字节存入高地址单元，所以按地址递增方式存储一个字数据时，应先存低字节，后存高字节。

要访问一个存储器单元，必须先得到这个存储器单元的物理地址。物理地址要由段地址和偏移地址合成。为什么要段地址和偏移地址呢?这是由于 8086/8088 用于寻址的寄存器都是 16 位，而 16 位寄存器只能寻址 64KB 的地址范围，但 8086/8088 有 20 条地址线，寻址可达 1MB 地址范围，即可在 00000H～FFFFFH 范围内寻址。为了能在 16 位字长的机器中提供 20 位地址，8086/8088 采用了存储器分段管理的方法。

分段的方法是将存储器划分成段，每段最大不超过 64KB，这样段内地址可用 16 位表示，称为偏移地址。段的起始地址称为段基址（常简称为段地址），必须为任一小段的起始地址。小段的划分方法是：从存储器的 0 地址单元开始，每 16 个字节单元为一小段。表 1-2 列出了 1MB 存储器的编址。

表 1-2　1MB 存储器的编址

00000H	00001H	00002H	……	0000EH	0000FH
00010H	00011H	00012H	……	0001EH	0001FH
00020H	00021H	00022H	……	0002EH	0002FH
⋮	⋮	⋮	⋮	⋮	⋮
FFFE0H	FFFE1H	FFFE2H	……	FFFEEH	FFFEFH
FFFF0H	FFFF1H	FFFF2H	……	FFFFEH	FFFFFH

分析表 1-2 中的地址，可以看出表中第 1 列小段起始地址的特点，即小段起始地址最后一位都为 0，在 1M 字节的寻址范围内用做段地址时，最后一位 0 可以不用保存，并且在 1M 字节的寻址范围内正好共有 64KB 个小段起始地址，从而段地址也可以用 16 位来表示。这样用一个 16 位段地址和一个 16 位段内偏移地址表示一个存储器单元，就可以使得每一个存储器单元都有一个惟一的 20 位地址，称为存储器单元的物理地址。因此，根据上述分析，20 位物理地址可按图 1.7 所示的方法形成。

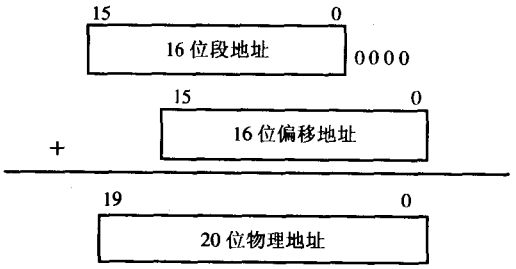

图 1.7　物理地址的形成

20 位物理地址=16 位段地址左移 4 位（即段地址乘以 16）十 16 位偏移地址

例 1-1　已知一个存储器单元的段地址为 3200H，偏移地址为 1210H，则该存储器单元的物理地址为：

$$3200H\times16+1210H=33210H$$

每当 CPU 需要访问一个存储器单元时，就需要产生一个 20 位物理地址，这时会有一个段寄存器（CS、DS、SS 或 ES）被自动选中用于提供段地址，而偏移地址则由指令直接提供或由 16 位可提供地址的寄存器给出。

形成各段内存储单元物理地址时所用的寄存器如图 1.8 所示。

在默认情况下，用于提供数据段或附加段偏移地址的 16 位寄存器可以是基址寄存器或变址寄存器。如果在指令中指定了由段寄存器 DS 或 ES 提供段地址，也可以用 BP 寄存器提供偏移地址。

在存储器中，代码段、数据段、堆栈段和附加段可以各段独立占用 64KB 存储区。分配各段存储空间时，各段之间可以间隔开或邻接，也可以部分重叠或完全重叠共享一段存储区。此外各段长度也可以不同，总之可以根据需要对各段进行不同形式的存储空间分配。各段存储区分配的两种典型方式如图 1.9 和图 1.10 所示。

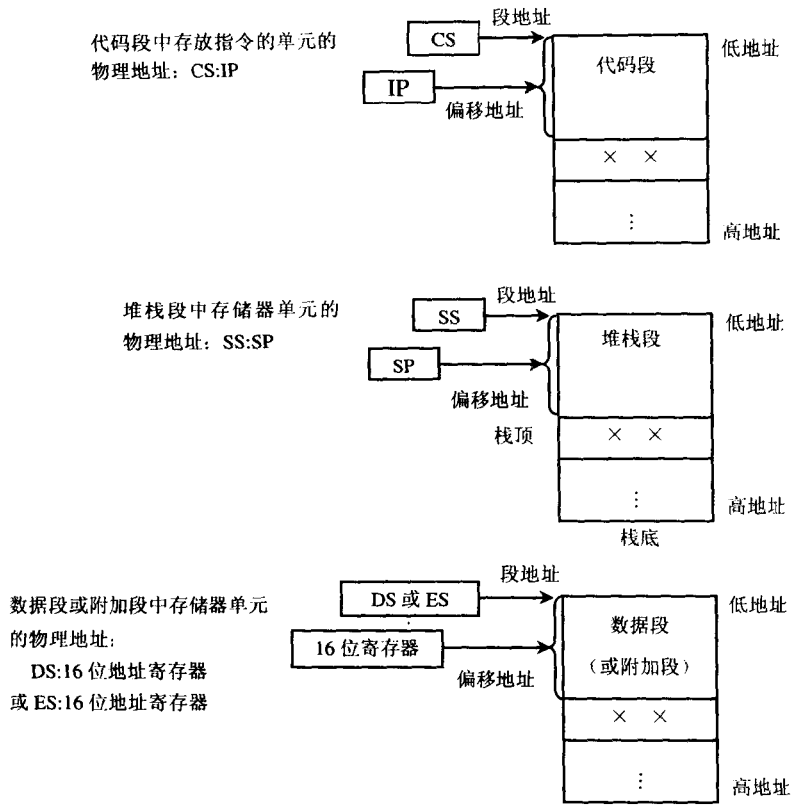

图 1.8　形成各段内存储器单元物理地址所用的寄存器

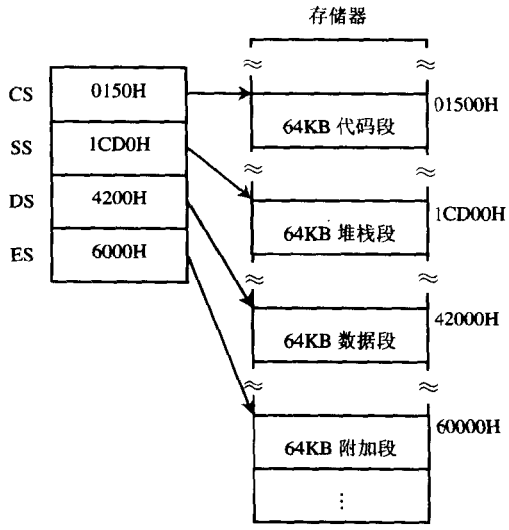

图 1.9　段分配方式之一（各段大小相等）

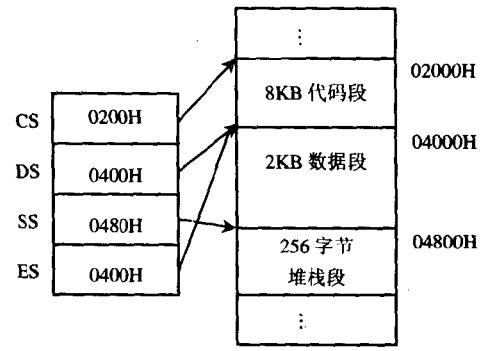

图 1.10　段分配形式之二（各段大小不等）

存储器分段管理法便于将代码区、堆栈区和数据区分配在不同存储区域，这时只需要将各个段寄存器置为相应的段地址就能保证正确地访问各个段。分段法也适应程序装入重定位的要求，只要正确设置 CS 中的段地址，就具备了程序装入不同存储区正常运行的基础。

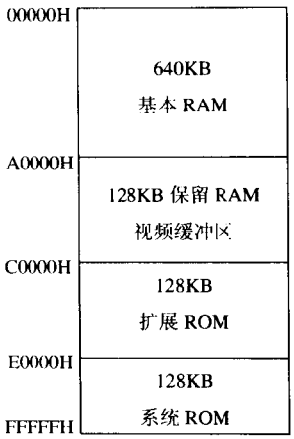

图 1.11　存储器的 1MB 地址空间分配

存储器 1MB 地址空间的分配映像基本上如图 1.11 所示，整个 1MB 地址空间从低地址端到高低址端可分为以下 4 个区段。

1．基本 RAM 区（00000H~9FFFFH）

该区占 640KB，由 DOS 系统管理。操作系统装入后将占用低地址端一小部分空间（存放中断向量表等），其余空间由用户使用。

2．保留 RAM 区（A0000H~BFFFFH）

该区占 128KB，用做视频缓冲区，是显示卡 RAM 的地址映射空间，用于存放屏幕显示信息，这部分地址空间实际上未全用。

3．扩展 ROM 区（C0000H~DFFFFH）

该区占 128KB，由 I/O 接口卡的 ROM 提供支持，用于存放系统不直接支持的外设驱动程序。

4．系统 ROM 区（E0000H~FFFFFH）

该区占 128KB，由系统使用，提供基本输入/输出系统（Basic Input/Output System）ROM-BIOS 程序，用于驱动输入/输出设备、系统加电自检、从磁盘引导操作系统等；此外，还提供了 CMOS 设置程序以及字符/图形符号点阵等信息。

本章介绍了 8086/8088 的功能结构、寄存器结构和用途、堆栈操作的特点以及存储器地址分配和分段管理方法，这些内容是学习下一章要介绍的寻址方式和指令系统的基础。

习题一

1.1　为什么要对存储器进行分段管理？存储器分段管理有什么优点？

1.2　寄存器可以分为哪几组，各组的主要用途是什么？

1.3　堆栈的存取方式有什么特点？

1.4　用连线连接左列概念与右列的解释：

（1）存储器	A. 控制 CPU 操作方式的标志位，共有 3 位：DF、IF、TF。
（2）堆栈	B. 表示存储器字节单元在所在段内位置的偏移量。
（3）段寄存器	C. 记录指令执行结果的标志位，共有 6 位：OF、SF、ZF、AF、PF、CF。
（4）标志寄存器	D. 计算机存储程序和数据等信息的记忆部件。
（5）SP	E. 惟一地代表存储器中每一个字节单元的地址。
（6）IP	F. 以先进后出方式存取数据的一段存储区。
（7）物理地址	G. 保存各段起始地址的寄存器，共有 4 个：CS、DS、ES 和 SS。
（8）偏移地址	H. 存放下一条要取指令的单元偏移地址的寄存器。
（9）条件码标志位	I. 保存堆栈当前栈顶地址的寄存器。
（10）控制标志位	J. 保存条件码标志和控制标志的 16 位寄存器。

1.5　在偏移地址为 00A0H 和 00A4H 的两个单元中，分别存放了两个字数据 1FA0H 和 4B3FH，画图表示出这两个字数据在存储器单元中的存放情况。