

自动恢复IBM PC机磁盘误删文件的程序设计

李 懋 阳

摘要: 本文详细地介绍了自动恢复IBM PC机磁盘误删文件的程序设计方法

一、引 言

恢复由于某种原因而误删的磁盘文件是许多用户十分关心的问题, 尽管有许多刊物介绍了恢复磁盘误删文件的方法和步骤, 但这些方法都是利用DOS中的DEBUG跟踪程序的功能在人工跟踪的环境下实现的。这种方法虽然可行, 但需操作人员熟悉并掌握磁盘结构、文件存放形式、信息格式及DEBUG程序的使用方法等, 加之DEBUG程序所提供的信息均是以十六进制形式给出的, 读、写、计算都很不方便。特别值得提及的是硬盘最多可有四个分区, 而各分区的位置一般说来各台机器是不相同的。因此, 文件分配表 (FAT) 并不是都放在硬盘物理空间的第三扇段, 它的位置是由硬盘的第一引导模块中的硬盘分区表来确定的, 而第一引导模块所处的扇段非常特殊, 它不属于任何分区, 是独立的一部分, 用DEBUG程序无法读出它的内容, 因此, 寻找各分区文件分配表的位置不但费时费事, 而且不够准确。用上述方法即使对一位计算机专业人员来讲也是一件极其烦琐、细致、易出错的工作, 而IBM PC机是目前国内应用最广泛的机型之一, 其中相当一部分工作人员是非计算机专业人员, 因此, 对他们来说, 恢复误删文件是相当困难的工作。

为此我们编写了自动恢复磁盘误删文件的汇编语言程序。本文着重介绍设计自动恢复误删文件程序的方法, 并将对磁盘空间分配及文件管理方式等做详细讨论。

二、磁盘空间分配及文件管理方式

IBM PC机磁盘操作系统对磁盘文件的管理是以磁盘空间中的引导区、文件分配区、文件目录区、数据占用区为基础的。系统对磁盘格式化后, 这些区被唯一地确定并初始化。对于软盘是整个软盘, 对于硬盘是整个DOS区段。硬盘各区的大小是在格式化时根据DOS区段的大小自行确定的。IBM PC DOS格式化的所有磁盘(硬盘和软盘)的扇段都是以512字节为单位的。

磁盘文件数据信息的记录是由操作系统中的文件管理系统来完成的, 文件管理系统能把有关的数据信息组织成文件并给其命名。文件管理系统通过查找文件的目录项及有关文件分配表便可知道文件的数据信息在磁盘物理空间的位置, 从而调用系统子程序把文件的数据信息从磁盘空间对应的位置读入内存, 供用户使用。

IBMBIOS.COM, 在IBMBIOS.COM 中也有一个格式表, 该表中有些数据(每道内扇段数、目录项数、盘内总扇段数、每簇内扇段数、磁头数等)都是由各格式字生成的。IBM-BIOS.COM 运行时, 由不同的格式字生成不同的格式表, 再由不同的格式表生成不同的设备控制块, 每个磁盘驱动器都有这样的控制器, 这样就解决了不同格式磁盘的兼容性。

软盘上没有第一引导模块, 它的引导区内只有一个引导模块(表3), 即第二引导模块。软盘中引导区的引导模块的信息格式及含意同硬盘第二引导模块相同。

表3 软盘格式化表(相当于硬盘第二引导模块)

0000	EB 2C 90 49 42 4D 20 20 32 2E 30 00 02 02 01 00 ...IBM..2.0.....
0010	02 70 00 D0 02 FD 02 00 09 00 02 00 00 00 00 00 .p.....
0020	0A DF 02 25 02 09 2A FF 50 F6 00 02 CD 19 FA 33 ...%...P.....3
0180	0D 0A 4E 6F 6E 1D 53 79 73 74 65 6D 20 64 69 73 ..Non-System.dia
0190	6B 20 6F 72 20 64 69 73 6B 20 65 72 72 6F 72 0D Non-disk error.
01A0	0A 52 65 70 6C 61 63 65 20 61 6E 64 20 73 74 72 .Replace.and.str
01B0	69 6B 65 20 61 6E 79 20 6B 65 79 20 77 6B 65 6E file.any.ley.when
01C0	20 72 65 61 64 79 0D 0A 00 0D 0A 44 69 73 6B 20 .ready....Disk.
01D0	42 6F 6F 74 20 66 61 69 6C 75 72 65 0D 0A 00 69 Boot.failure...3
01E0	62 6D 62 69 6F 20 20 63 6F 6D 20 69 62 6D 64 6F bmbio..comibmd0
01F0	73 20 20 63 6F 6D 30 00 00 00 00 00 00 00 55 AA s..com0.....U.

无论是第一引导模块, 还是第二引导模块, 无论是硬盘, 还是软盘, 在各引导模块所处的扇段中的最后两个字节均标有“55AAH”“有效签名”, 无“有效签名”为非引导模块或无效引导模块。

文件分配区: 该区是在引导区后, 是存放文件分配表FAT的区域。关于文件分配表的格式、含义及使用方法在参考资料^[1-4]中有详细介绍, 读者可查阅。

文件目录区: 该区在磁盘空间中紧随文件分配表之后, 是为存放每个文件目录项而开辟的磁盘空间, 每个目录项为32字节长, 依次存放在文件目录区之中。每个目录项的格式及含意在硬件手册及DOS手册中均有详细说明。

数据占用区: 任何磁盘文件最终都是以数据的形式出现并存放在这里的, 该区中的各扇段(或几个扇段)与文件分配区中文件分配表内所包含的各文件的簇链有一一对应关系, 一个簇由若干个扇段组成, 每个簇号对应于这组扇段的磁盘空间位置, 操作系统利用这一对应关系寻找某一文件所占的扇段, 并在其各扇段中读、写该文件的信息内容。

文件分配表FAT, 文件目录项, 文件的簇链, 簇链中簇号与扇段的对应关系已有很多资料^[1-4]详细做了介绍, 本文不再赘述。

三、恢复误删文件的可能性

IBM PC 机操作系统删除某一文件是以把该文件在目录区的目录项的文件名的第一个字符改为E5H, 作为被删文件, 并在文件分配表中把该文件的簇链全部改为OOOH, 作为被释放空间来实现的。操作系统在为新文件分配磁盘空间时寻找空簇, 并把空簇对应的扇段分配给新文件。在没有写入新文件的前提下, 被删文件在数据占用区的全部数据信息并没有被清除或覆盖。因此, 只要恢复被删文件目录项中的第一个字符, 即文件名的第一个字符, 并恢复该文件在文件分配表中的簇链即可。当然, 上述可能是在删除该文件后没有写入新文件的前提下才成为可能。一旦删除某一文件后又写入了新文件, 那么被删文件原来占用的磁盘空间, 即被删文件的簇链很有可能重新分配给新写入的文件, 使得被删文件在数据占用区的数据信息被新写入的文件所覆盖, 使得被删文件不能恢复或不能全部恢

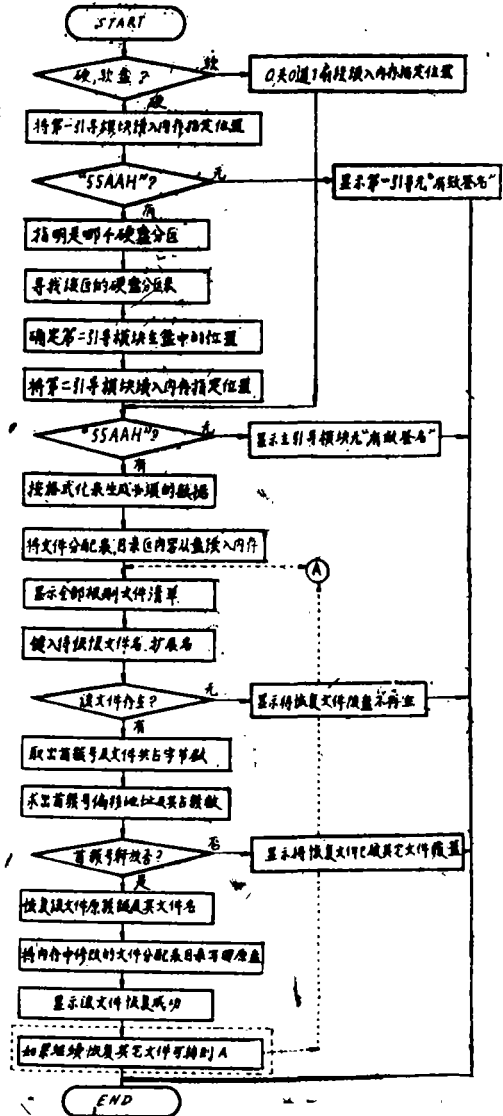
复。因此，当误删了某一文件后，切记不要写入其它文件，以免覆盖原来误删文件的数据及信息内容。

文件分配区中的文件分配表是由各文件占用的簇链组成，各簇链中所包含的簇号在文件分配表中是严格单值递增的，正是由于簇号的这种单值递增性，保证了文件在磁盘空间的独立性和互相不交叉的正确性。因此，能否恢复误删文件，只要检查该文件原来在文件分配表中的簇链是否仍然释放即可。根据簇号的单值递增性不难看出，若首簇号依然释放，则该文件必能恢复。否则有两种可能，其一，若原簇链中尾部一些簇号依然释放，那么这部分簇号对应的信息能够恢复；其二，若原簇链未被释放，那么说明已写入了新文件，而新文件已全部覆盖了误删文件，使得误删文件再没有全部恢复或部分恢复的机会了。

四、编程方法及框图

前文已经提及过硬盘第一引导模块的特殊性，因此，程序中应利用IBM PC机中的ROM BIOS的磁盘I/O模块、键盘输入模块及显示输出模块来实现各种信息的交换。利用磁盘I/O模块来实现磁盘与内存，内存与磁盘的读写；利用显示模块可显示菜单及各种提示；利用键盘输入模块实现文件名的输入及回答各种选择和提示。为了保证原磁盘中信息的正确性和完整性，应尽量减少磁盘的读写次数，误删文件的查找、恢复应在内存中一次完成，若误删文件恢复成功，则把内存中经过修复的信息一次写回原盘原来位置；若误删文件无法恢复，则可直接返回DOS，避免一次写盘机会。下面结合框图做具体说明。

首先，利用显示功能模块显示提问待恢复文件是在硬盘上还是在软盘上，并通过键盘输入回答上述发问。假如待恢复文件是在硬盘上，则利用POM BIOS的磁盘I/O模块将硬盘的0头0柱面，第一扇段（即第一引导模块所处的扇段）的信息读入内存指定位置，并首先检查该扇段最后两个字节是否为“有效签名”55AAH。若无“有效签名”，则此扇段为无签引导模块，应退回DOS并显示“无第一引导模块”或“第一引导模块无效”，倘若有“有效签名”，则按第一引导模块的分区表（表1）确定第二引导模块的位置。第二引导模块的位置确定后，仍利用ROM BIOS的磁盘I/O模块将第二引导模块所处扇段的信息覆盖第一引导模块在内存中的信息，并同样首先检查该扇段最后两个字节是否为“有签签名”55AAH，并以第一引导模块类似的方法处理之。



编程基本框图

其次,根据第二引导模块的格式化表(表2)可确定出每扇段的字节数,每簇的扇段数,BOOT长度,文件分配表FAT长度及个数,目录项数,盘内本区总扇段数,盘格式字,每道内的扇段数、磁头数及DOS起始扇段号。根据上述数据可生成文件分配区及目录区的位置和各区占用的扇段数,每簇字节数,这样就可把文件分配区及目录区的内容用硬盘I/O模块分别从磁盘读入内存的指定位置,为在内存中的查找,恢复做准备。

再次,通过显示模块显示出盘中全部被删文件清单,这对于了解和掌握盘内所有被删文件是十分重要的。如果由于不慎误删了多个文件,乃至全部文件,可根据提供的清单逐个地进行恢复。

最后,利用键盘输入模块输入待恢复文件名及扩展名,根据文件名及扩展名查找在内存中的目录项,根据目录项中的首簇号及该文件占用的总字节数寻找该文件在文件分配表中的簇链(簇链的查找及恢复已有许多资料^[1-4]做了介绍)并检验其簇链是否依然释放,若依然释放,则恢复其簇链,恢复文件名并把在内存中修改过的全部信息一次写回原盘原来位置,显示恢复成功并返回DOS;否则显示不能恢复及其原因并退回DOS。

软盘文件的恢复方法与硬盘完全相同,只是软盘没有第一引导模块,因此软盘文件的恢复可直接读软盘0头0道第一扇段并寻找软盘格式化表(相当于硬盘的第二引导模块),因该方法与硬盘恢复方法相同,这里不另加叙述。

五、程序实例

由于篇幅所限,这里只能提供自动恢复误删文件汇编语言程序的一部分。这部分程序的功能是恢复某一被删文件的簇链和文件名,执行本程序段之前,应将文件分配表从磁盘读入到以FAT为标号的内存地址单元中,把文件目录区的各目录项从磁盘读入到以DIR为标号的地址单元中,被删文件目录项的首地址放入DI寄存器中,被删文件目录项中的首簇号以及由首簇号计算出来的首簇号偏移地址分别压入堆栈,由文件共占字节数计算出的文件共占簇数放入以number为标号的单元中,这样执行本程序段即可恢复某一被删文件的簇链及文件名。最后调用RWFAT和RWDIR子程序把以FAT为标号的文件分配表和以DIR为标号的目录区信息从内存中写回原磁盘原来占用的空间。

这段程序只适用于簇号为12bit(一个半字节)的磁盘(双面九扇段四十道软盘和10M硬盘),对于20M硬盘,因其簇号为16bit(二个字节)而不适用。

六、结束语

根据介绍的方法我们已经成功地编写了自动恢复磁盘误删文件的汇编语言程序,这个程序无论是单簇(小)文件,还是多簇(大)文件;无论是软盘,还是硬盘文件,只要是作为磁盘文件出现的任何文件(只读文件,系统文件,隐性文件等)均可实现自动恢复,不需人工干预。当然,上述可能是在误删某些文件后尚未写入新文件的前提下。能否恢复误删文件,计算机可通过屏幕自动显示出来。通过屏幕显示,还可把盘上全部被删文件清单显示在屏幕上,这对于了解并掌握磁盘所有被删文件是十分重要的。如果不慎误删了多个文件,乃至全部文件,可根据提供的被删文件清单,逐个地进行恢复。

该程序经过近一年时间的使用,证明方法简单,灵活,安全,适用。了解并掌握此方法

不仅仅可以编制恢复磁盘误删文件程序,掌握恢复误删文件的全过程,特别是了解并掌握磁盘空间分配及信息格式(尤其是硬盘第一引导模块,第二引导模块的信息格式),开发新的应用领域是有很多帮助的。

对于前洋同志给予的指导与帮助表示真诚的谢意。

参 考 文 献

- [1] 刘洵,于前洋;《光学机械》1987年第6期P25
- [2] 杨长生;《小型微型计算机系统》浙江大学1986年第1期P61
- [3] 白剑华;《小型微型计算机系统》湖北沙市棉纺厂1986年第12期P56

The Programming of Auto-Recovery Mideleted Files of Disk on IBM PC Computer

Li Maoyang

Adstract

This paper introduces a method for programming of auto-recovery mideleted files of disk on IBM PC computer.