编号：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 实验 | 一 | 二 | 三 | 四 | 五 | 六 | 七 | 八 | 总评 | 教师签名 |
| 成绩 |  |  |  |  |  |  |  |  |  |  |

武汉大学国家网络安全学院

课程实验(设计)报告

题 目： 磁盘格式与数据恢复

专业(班)： 2021级信安6班

学 号： 2021302181156

姓 名： 赵伯俣

课程名称： 软件安全实验

任课教师： 赵磊

2023年 10 月 15日

**目 录**

[实验1 磁盘格式与数据恢复 1](#_Toc22797)

[1.1实验名称 1](#_Toc18370)

[1.2实验目的 1](#_Toc6265)

[1.3实验步骤及内容 1](#_Toc21569)

[1.4实验关键过程、数据及其分析 3](#_Toc18897)

[1.4.1利用WinHex查看物理磁盘和逻辑磁盘。 3](#_Toc5717)

[1.4.2分析主引导扇区的组成部分。 4](#_Toc29022)

[1.4.3分析四个主分区项的内容 5](#_Toc1530)

[1.4.4分析主扩展分区表的结构。 6](#_Toc24753)

[1.4.5分析每个本地逻辑盘的位置大小和逻辑结构及其扇区信息。 7](#_Toc10712)

[1.4.6分析GPT分区的格式。 8](#_Toc6574)

[1.4.7GPT分区结构与MBR的具体差异有哪些？ 8](#_Toc8889)

[1.4.8主分区头所在扇区包括哪些重要内容，验证这些重要内容的有效性。 9](#_Toc3748)

[1.4.9通过分区节点分析自己硬盘的各分区信息。 10](#_Toc17329)

[1.4.10 分析FAT32分区格式的逻辑盘的起始扇区。 11](#_Toc20554)

[1.4.11查看该逻辑盘的FAT1和FAT2的内容和大小。 13](#_Toc9391)

[1.4.12查看该逻辑盘的根目录区。 14](#_Toc166)

[1.4.14新建文件并分析其目录项 14](#_Toc19007)

[1.4.15分析新建文件的簇相关信息 15](#_Toc23077)

[1.4.16删除文件的恢复 16](#_Toc5329)

[1.4.17 课后习题思考 17](#_Toc21608)

[（1）在磁盘分区过程中，用户提供了哪些信息？分析分区工具的工作原理。 17](#_Toc31422)

[（2）高级格式化与低级格式化的具体原理和区别是什么？ 17](#_Toc31898)

[（3）查找资料，对NTFS分区的总体结构进行分析，尝试对NTFS下删除的文件进行手工恢复。 18](#_Toc27612)

[（4）用数据粉碎工具粉碎指定文件，分析其数据粉碎原理。 20](#_Toc1350)

[（5）通过分区表看到的分区字节数为何与资源管理器中看到的分区字节数有差异？ 20](#_Toc29291)

[（6）如果删除的文件是长文件名，如何恢复所有文件名。 20](#_Toc29187)

[1.5实验体会和拓展思考 21](#_Toc1174)

## 实验1 磁盘格式与数据恢复

### 1.1实验名称

磁盘格式与数据恢复

### 1.2实验目的

1. 了解磁盘的物理和逻辑结构
2. 熟悉FAT32文件系统
3. 学会使用磁盘编辑软件
4. 了解文件删除、格式化的基本原理
5. 能够利用工具或者手工恢复被删除的文件

### 1.3实验步骤及内容

**第一阶段：**

* 熟悉WinHex的使用。
  + 熟悉磁盘工具的使用。
  + 利用WinHex查看物理磁盘和逻辑磁盘。
  + 了解WinHex中相关工具的用法。

**第二阶段：**

* 分析本地硬盘的主引导扇区
* 利用磁盘编辑工具查看MBR磁盘分区并分析：
  + 主引导扇区由哪些部分组成？
  + 四个主分区项的内容各代表什么？
  + 分析主扩展分区表的结构。
  + 通过分区项来确定每个本地逻辑盘的位置以及大小，并画出本地硬盘的逻辑结构。
    - 每个本地盘的开始扇区位置，总扇区数，结束扇区位置，各扩展分区表扇区位置，保留空间数量。
* 利用磁盘编辑工具查看GPT磁盘分区并分析
  + GPT分区结构与MBR的具体差异有哪些？
  + 主分区头所在扇区包括哪些重要内容，验证这些重要内容的有效性。
  + 通过分区节点分析自己硬盘的各分区信息。

**第三阶段：**

* 熟悉FAT32文件格式。
  + 用WinHex打开某个FAT32分区格式的逻辑盘。
  + 查看该逻辑盘的起始扇区，分析起始扇区中的相关字段（BPB:BIOS Parameter Block）。
  + 查看FAT1和FAT2的内容和大小。
  + 查看该逻辑盘的根目录区。
  + 查看某个文件的目录项结构和FAT链以及具体存储位置。
    - 在根目录下建立文本文件：test-学号后3位.txt，其中填充60K左右的文本字符保存（注意：先行存储其他数据使得该文件的首簇高位不为0）。
    - 查看该文件的目录项，对其进行分析，并得到该文件所在位置以及大小。
    - 查看首簇位置，并得到簇链表。通过簇链表查看该文件内容。
    - 记录首簇位置（14H-15H,1AH-1BH）和文件大小（1CH-1FH）。

**第四阶段：**

* 手工恢复被删除的文件
  + 删除前面所建立的文件。(del&shift+del)
  + 利用WinHex在该文件所在盘符查找该.txt文件的目录项。
  + 查看目录项的变化。
  + 利用该残余目录项来计算被删除的文件所在的位置。
  + 手工恢复该文件（文件名、首簇高位、簇链表修复）。

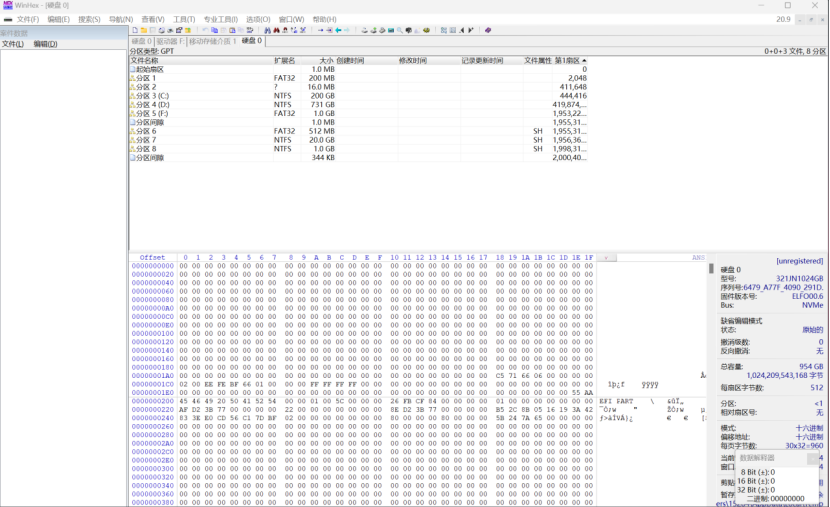
**课后习题思考：**

* 在磁盘分区过程中，用户提供了哪些信息？分析分区工具的工作原理。
* 高级格式化与低级格式化的具体原理和区别是什么？
* 查找资料，对NTFS分区的总体结构进行分析，尝试对NTFS下删除的文件进行手工恢复。
* 用数据粉碎工具（如金山、360、Strongdisk等）粉碎指定文件，分析其数据粉碎原理。
* 通过分区表看到的分区字节数为何与资源管理器中看到的分区字节数有差异？
* 如果删除的文件是长文件名，如何恢复所有文件名。

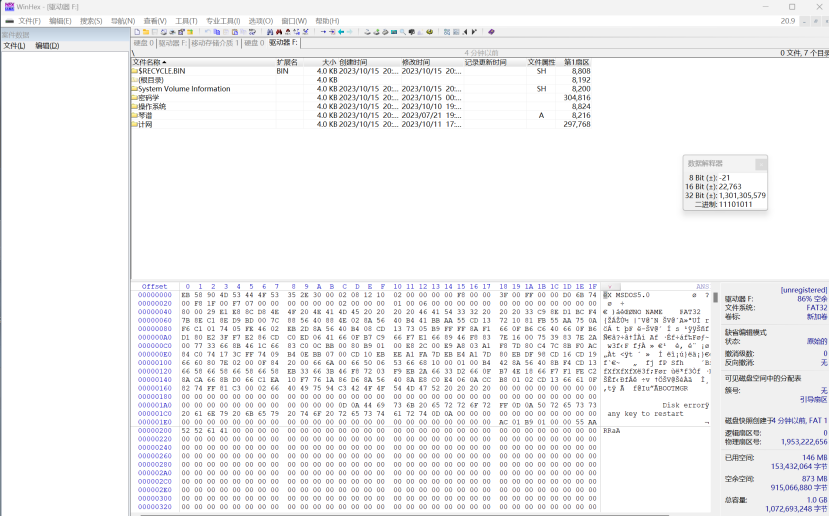
### **1.4实验关键过程、数据及其分析**

#### 1.4.1利用WinHex查看物理磁盘和逻辑磁盘。

利用WinHex打开本地的物理盘的情况如下图所示，存在有8个分区和2个分区间隔。

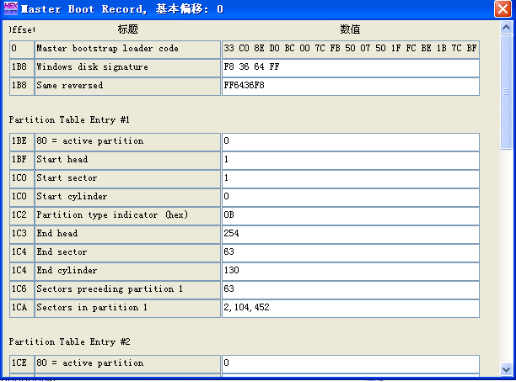


使用WinHex打开本地的逻辑分区后可以看到在逻辑分区中的众多文件目录。

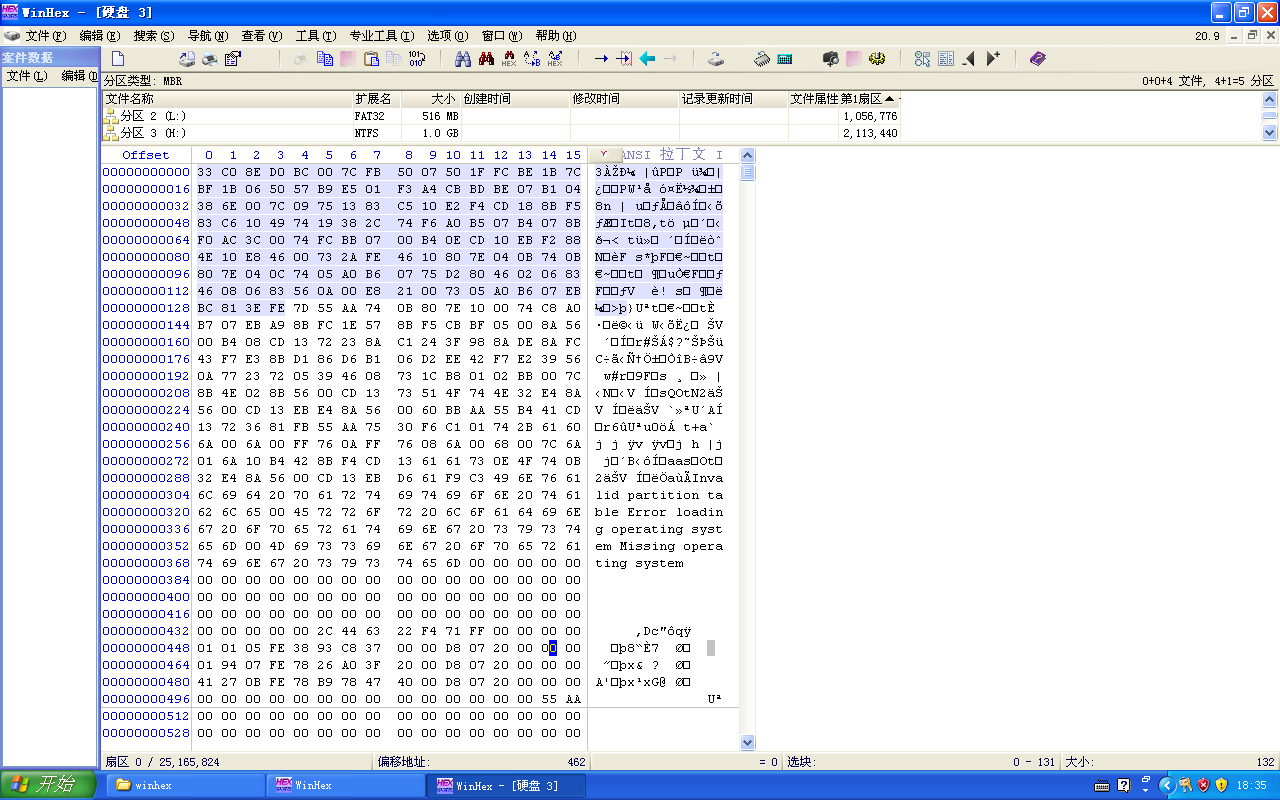


#### 1.4.2分析主引导扇区的组成部分。

将主引导扇区的分区表打开后如下图所示。



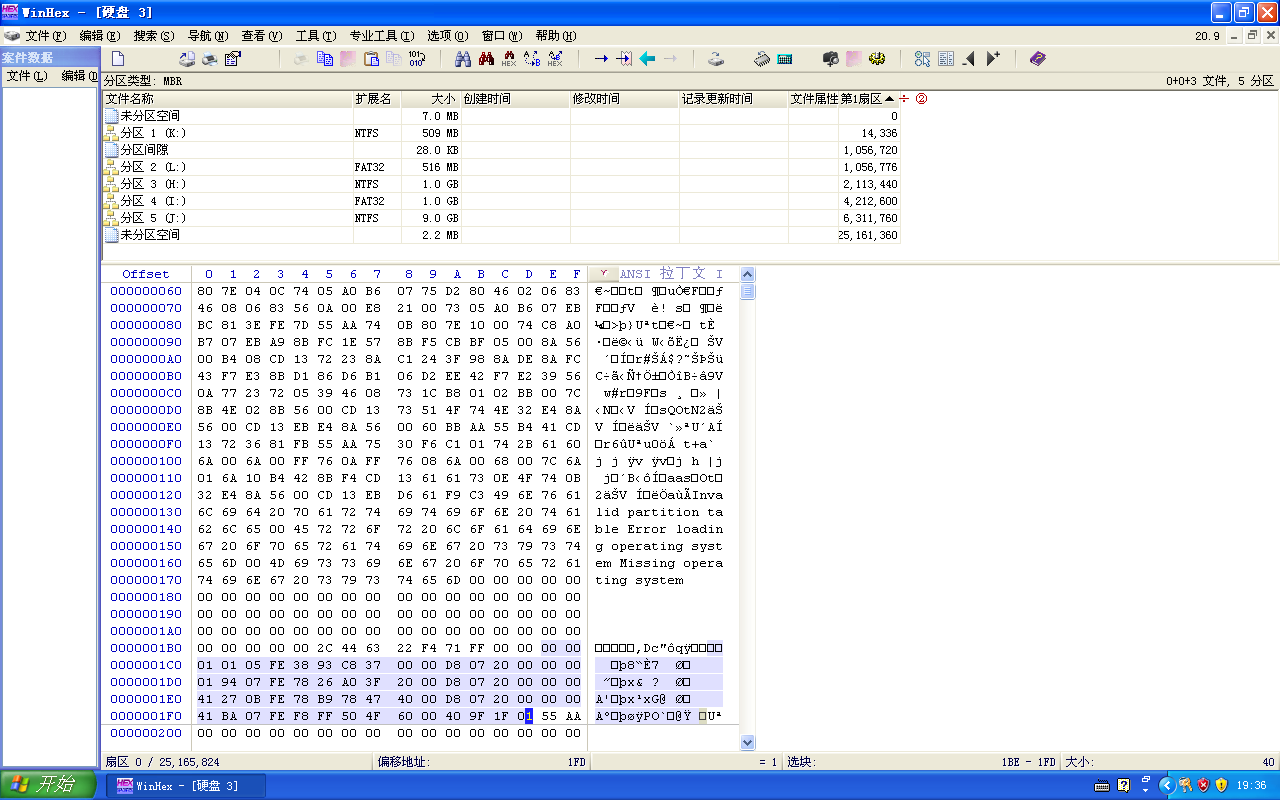
分区表中的具体内容如下图所示



在一个主引导扇区中，前1B8H（440）个字节为主引导扇区的引导代码，从第1BEH（446）个字节开始为主引导扇区的分区表，分区表占用64个字节，最后主引导扇区以55AA结尾。

#### 1.4.3分析四个主分区项的内容

四个主分区项的内容如下图所示，

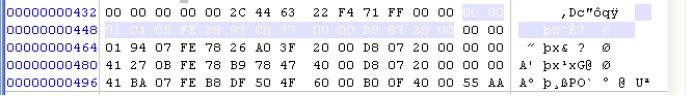


其中分区表由64B组成，分区表一共有四个，其中每个分区表的结构如下表所示。

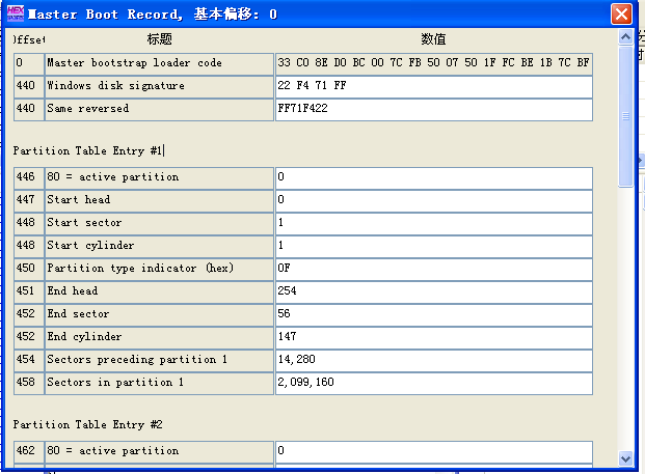
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 存储字节 | 数据内容 | 分区1 | 分区2 | 分区3 | 分区4 |
| 1 | 引导标志 | 00 | 00 | 00 | 00 |
| 2 | 本分区的起始磁道号 | 00 | 00 | 00 | 00 |
| 3 | 本分区的起始扇区号 | 01 | 01 | 41 | 41 |
| 4 | 本分区的起始柱面号 | 01 | 94 | 27 | BA |
| 5 | 分区类型 | 05 | 07 | 0B | 07 |
| 6 | 本分区的结束磁道号 | FE | FE | FE | FE |
| 7 | 本分区的结束扇区号 | 38 | 78 | 78 | F8 |
| 8 | 本分区的结束柱面号 | 93 | 26 | B9 | FF |
| 9-12 | 本分区之前已经占用的扇区数 | 000037C8 | 00203FA0 | 00404778 | 00604F50 |
| 13-16 | 本分区的总扇区数 | 002007D8 | 002007D8 | 002007D8 | 011F9F40 |

#### 1.4.4分析主扩展分区表的结构。

打开一个主拓展分区的分区表如下图所示。

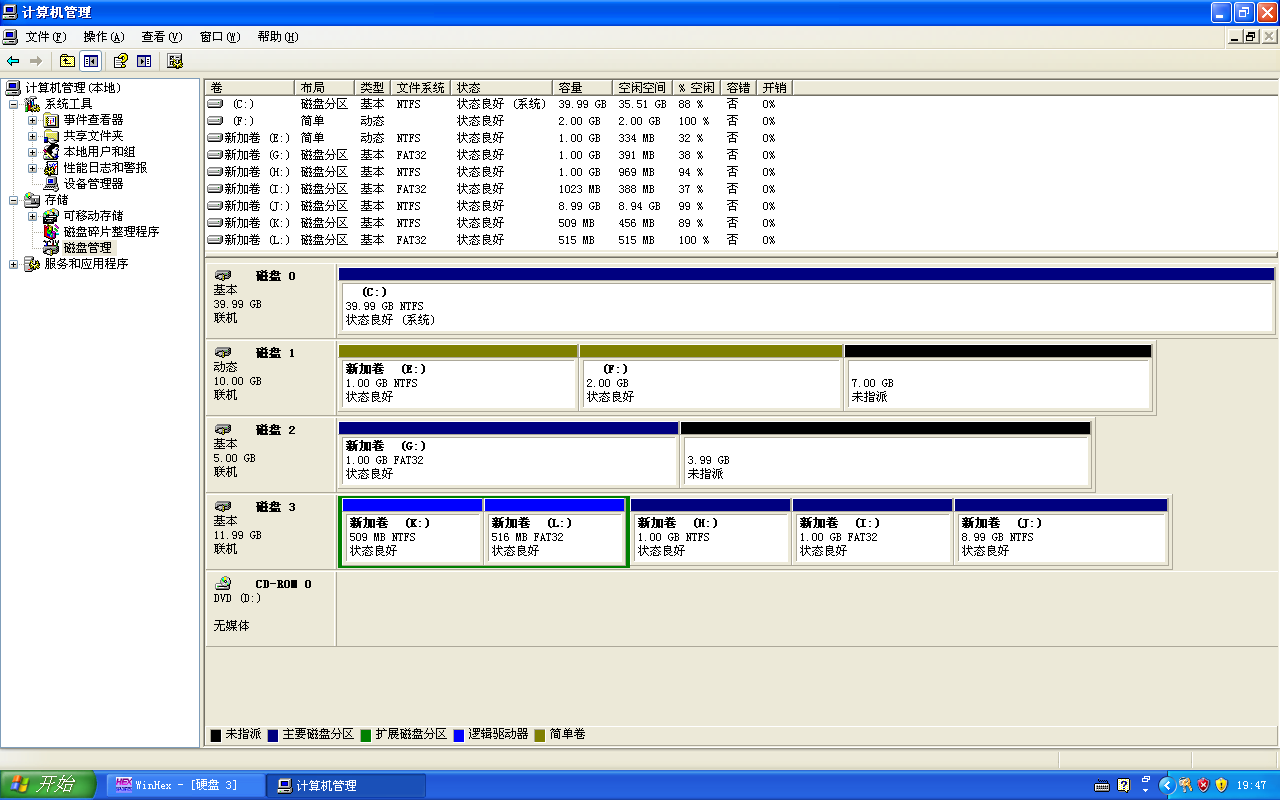


使用Winhex打开其模板如下图所示。

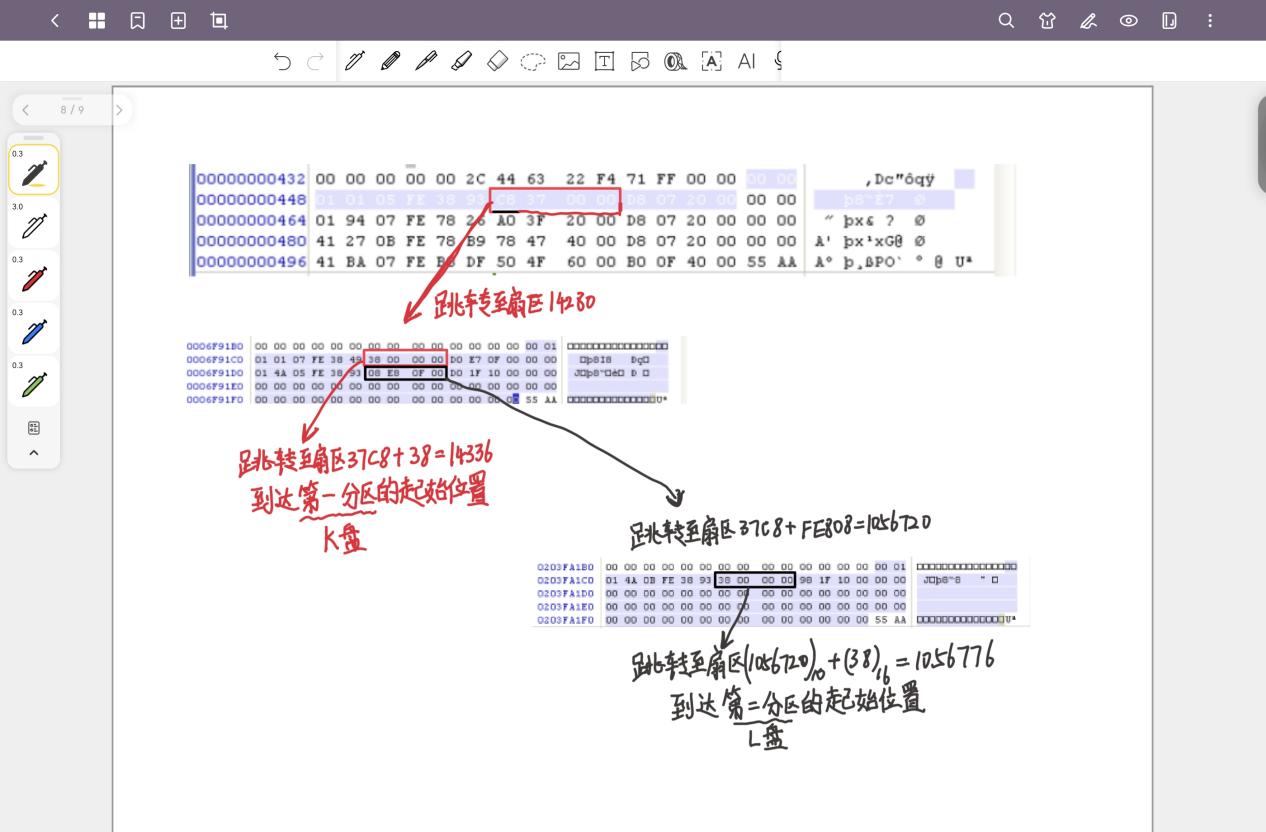


一个拓展分区中可以有多个逻辑驱动器，扩展分区中的每个逻辑驱动器的分区信息都存在一个类似于MBR的扩展引导记录中，如上图所示，扩展引导记录包括分区表和结束标志“55 AA”，没有引导代码部分。

EBR中分区表的第一项描述第一个逻辑驱动器，第二项指向下一个逻辑驱动器的EBR。如果不存在下一个逻辑驱动器，第二项就不需要使用。下图为在磁盘管理器中查看逻辑磁盘状态。绿色框中的部分就是主拓展分区。

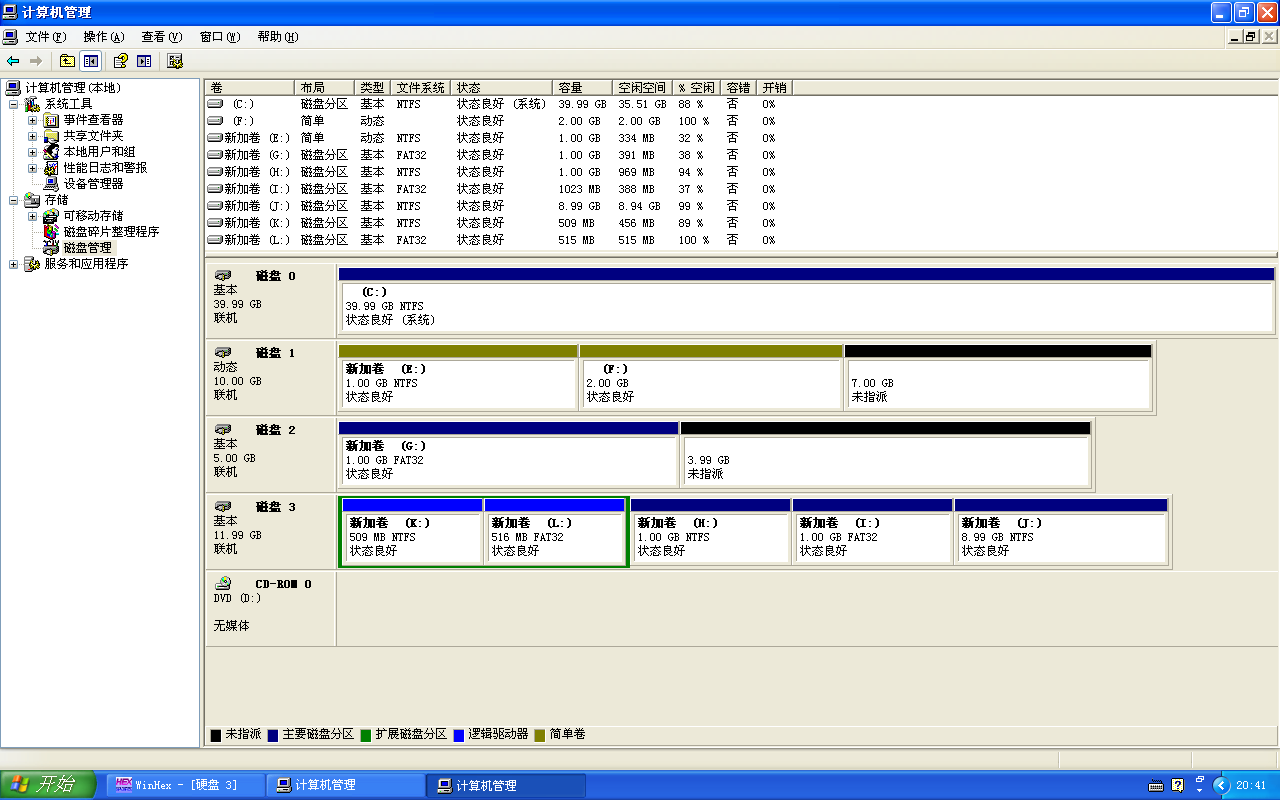


主拓展分区在硬盘中的结构如下图所示。

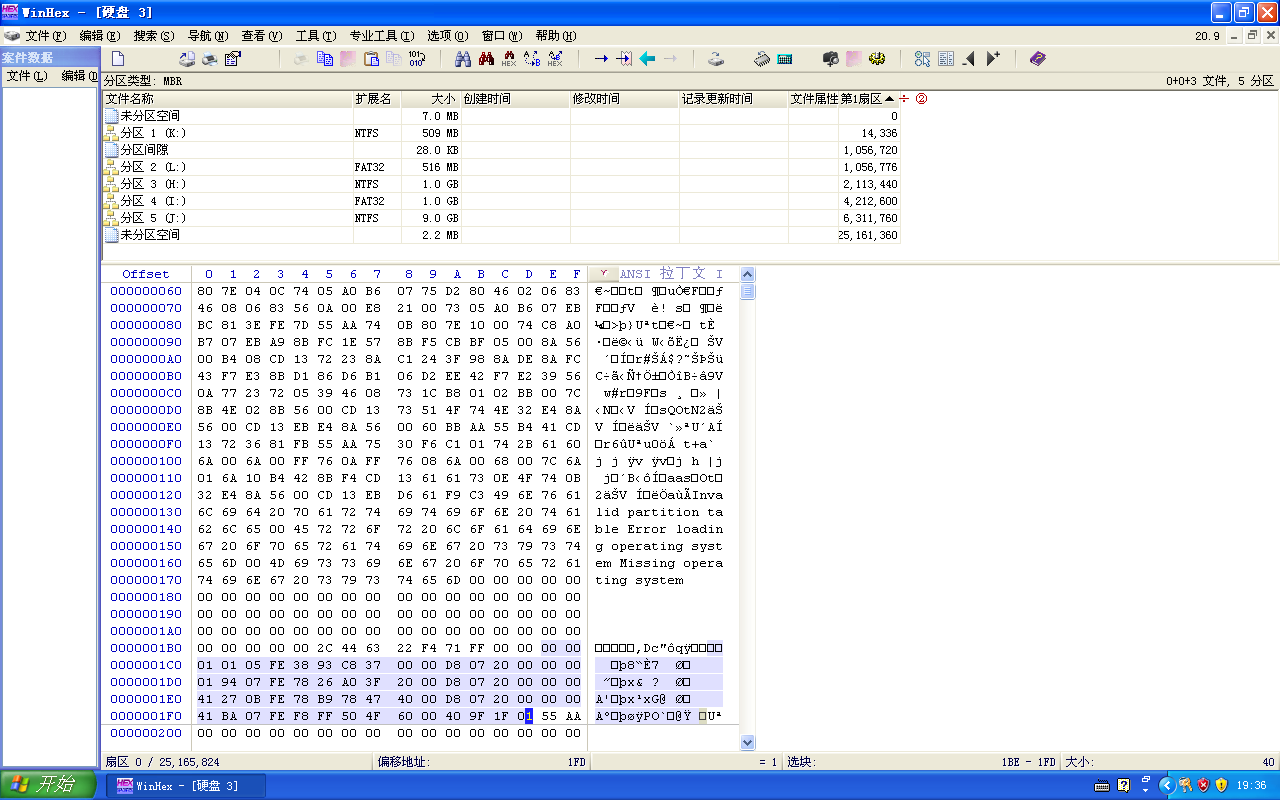


#### 1.4.5分析每个本地逻辑盘的位置大小和逻辑结构及其扇区信息。

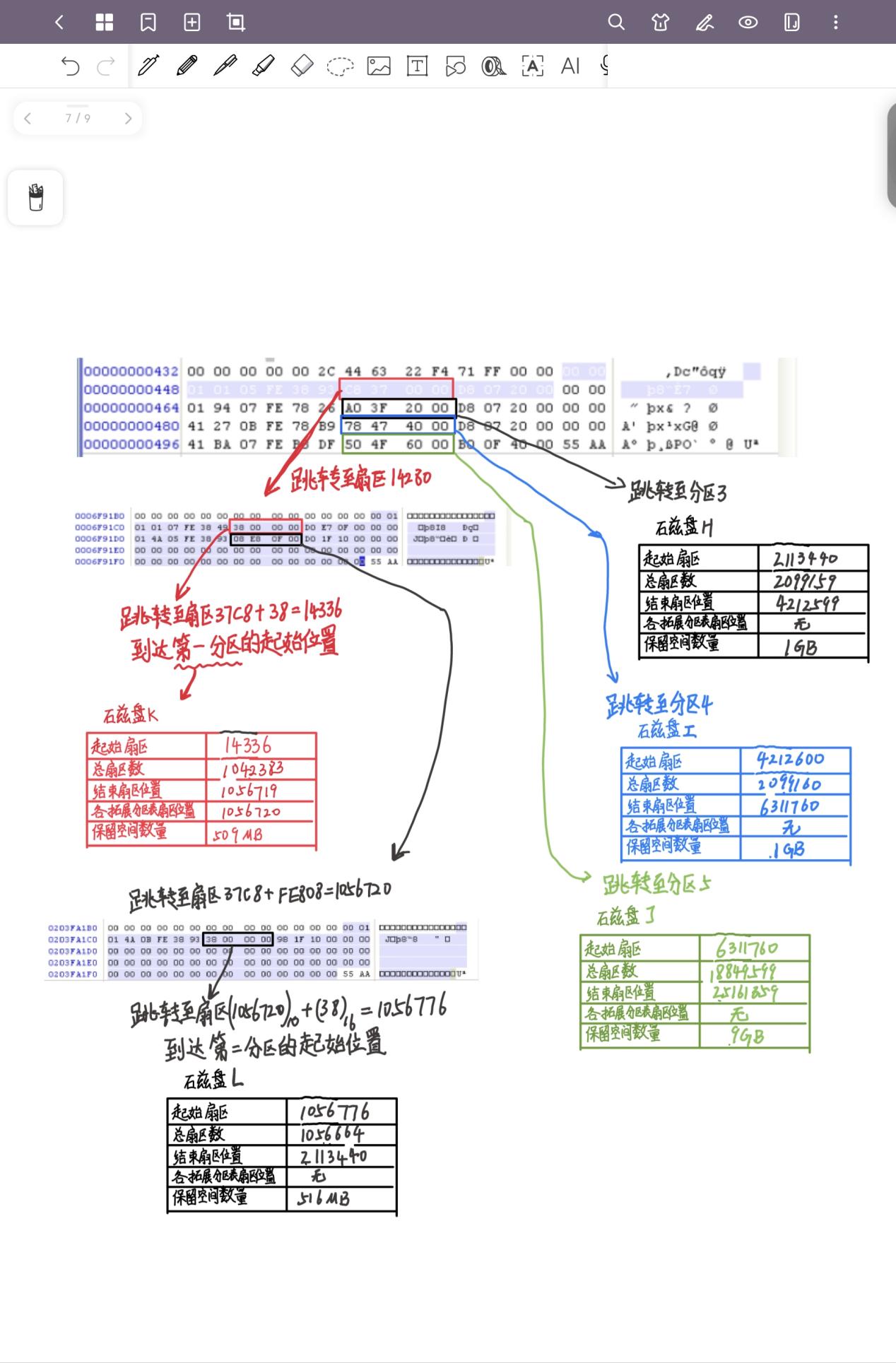
在本次实验中分析磁盘3每个逻辑盘的位置大小和逻辑结构，首先在磁盘管理器中查看逻辑盘结构如下图所示。



打开其主分区表如下所示

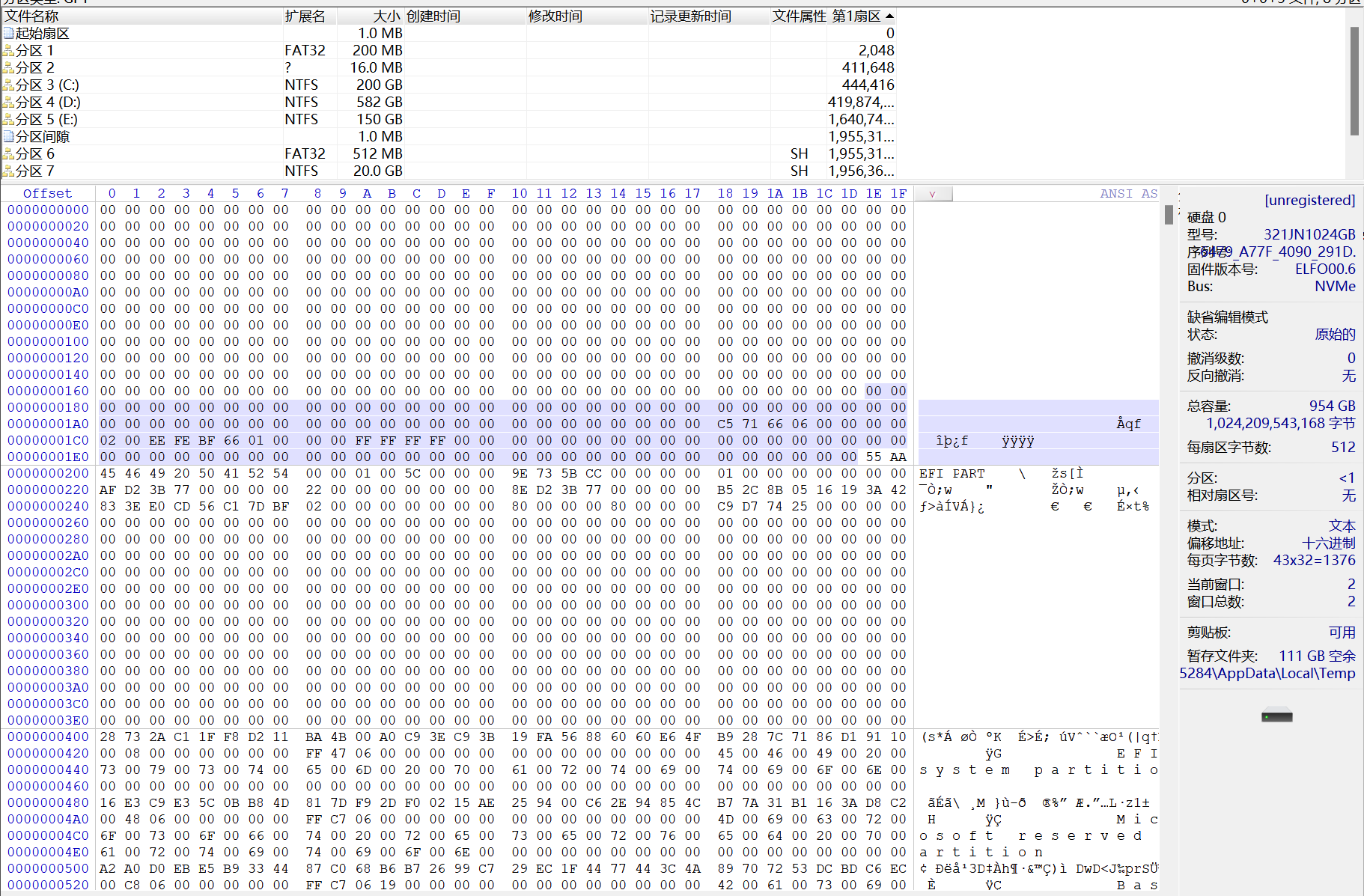


则每一个逻辑盘的信息如下图所示。

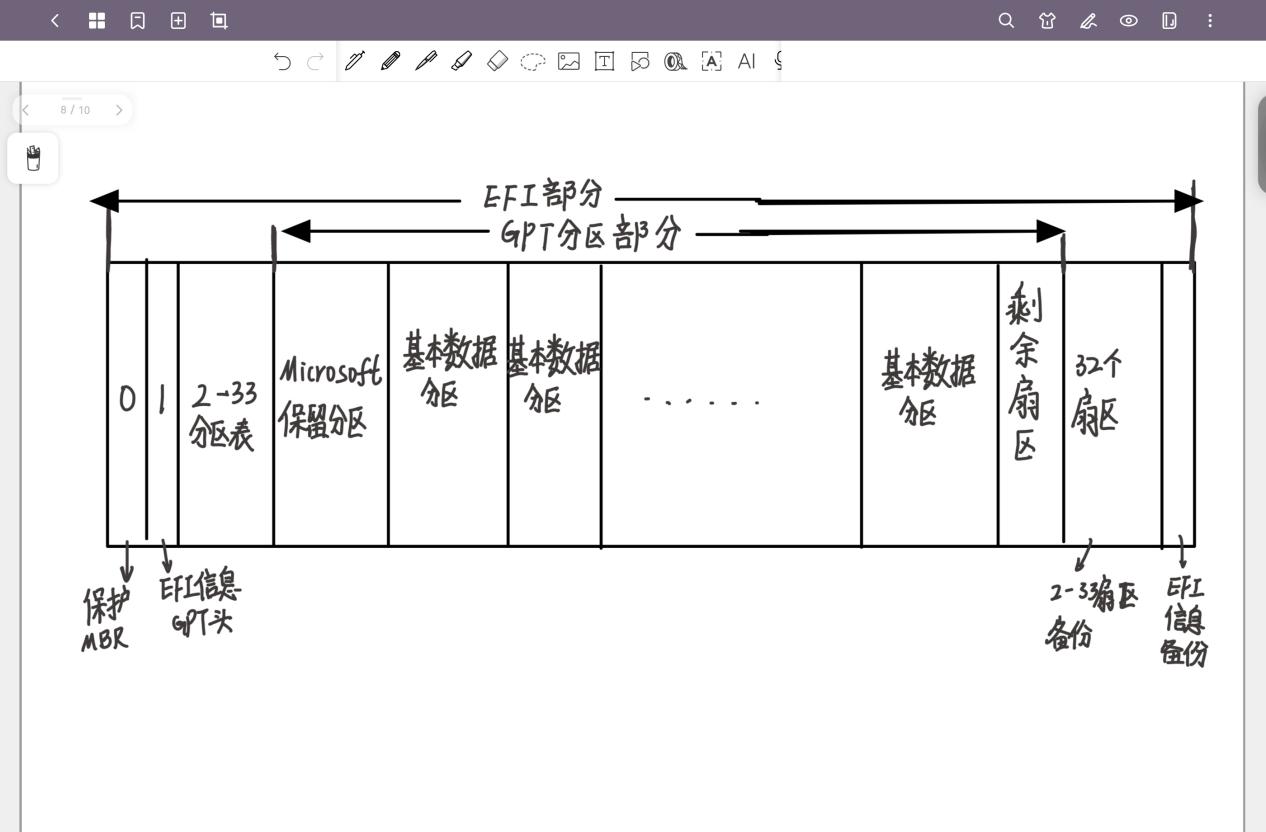


#### 1.4.6分析GPT分区的格式。

使用WinHex在物理机上打开一个硬盘后硬盘内容如下图所示



GPT分区的整体结构如下图所示。



#### 1.4.7GPT分区结构与MBR的具体差异有哪些？

1.MBR中分区表的组织将磁盘的可寻址存储空间限制为2TB（232×512字节）。

2.MBR只支持4个主分区，或者3个主分区和1个扩展分区的组合。如果要创建更多分区，则需要将其中一个主分区设为“扩展分区”，然后在扩展分区内创建更多逻辑分区。否则，磁盘将转换为动态磁盘。

3.在MBR中，分区表的大小是固定的；在GPT分区表头中可自定义分区数量的最大值，也就是说GPT分区表的大小不是固定的。

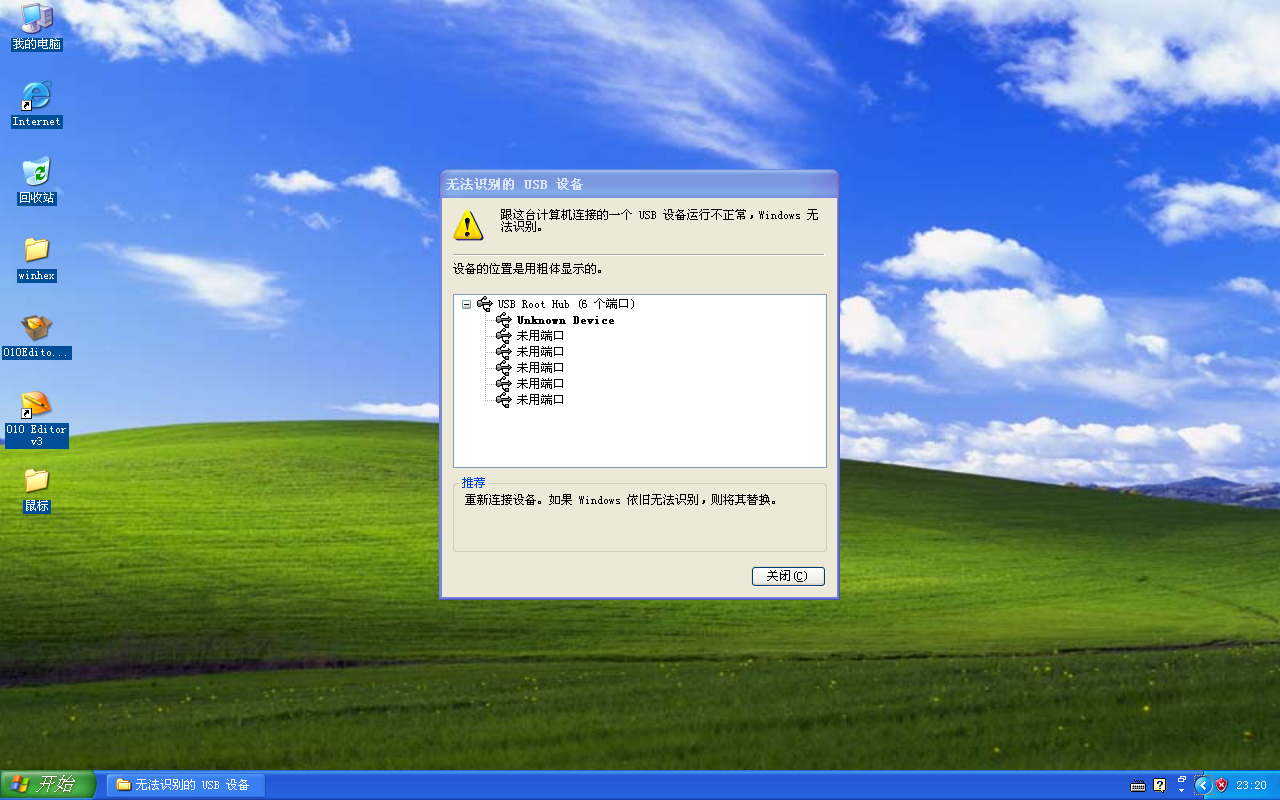
4. GPT磁盘具有冗余的主分区表和备份分区表，可以优化分区数据结构的完整性。

通常来说，MBR和BIOS（MBR+BIOS）、GPT和UEFI（GPT+UEFI）是相辅相成的。这对于某些操作系统（例如 Windows）是强制性的，但是对于其他操作系统（例如 Linux）来说是可以选择的。

#### 1.4.8主分区头所在扇区包括哪些重要内容，验证这些重要内容的有效性。

主分区头所在的扇区有一个标识为0xEE的分区，以此来表示这块硬盘使用GPT分区表。不能识别GPT硬盘的操作系统通常会识别出一个未知类型的分区，并且拒绝对硬盘进行操作，除非用户特别要求删除这个分区。这就避免了意外删除分区的危险。另外，能够识别GPT分区表的操作系统会检查保护MBR中的分区表，如果分区类型不是0xEE或者MBR分区表中有多个项，也会拒绝对硬盘进行操作。

找到一个GPT分区类型的U盘之后将其连接到WinXP操作系统中，WinXP能够识别MBR类型的磁盘，连接好后会显示如下错误。由此验证GPT分区头所在的扇区能够拒绝不能识别GPT分区的操作系统对其进行读取操作。



#### 1.4.9通过分区节点分析自己硬盘的各分区信息。

打开GPT分区的磁盘的分区表后如下图所示。

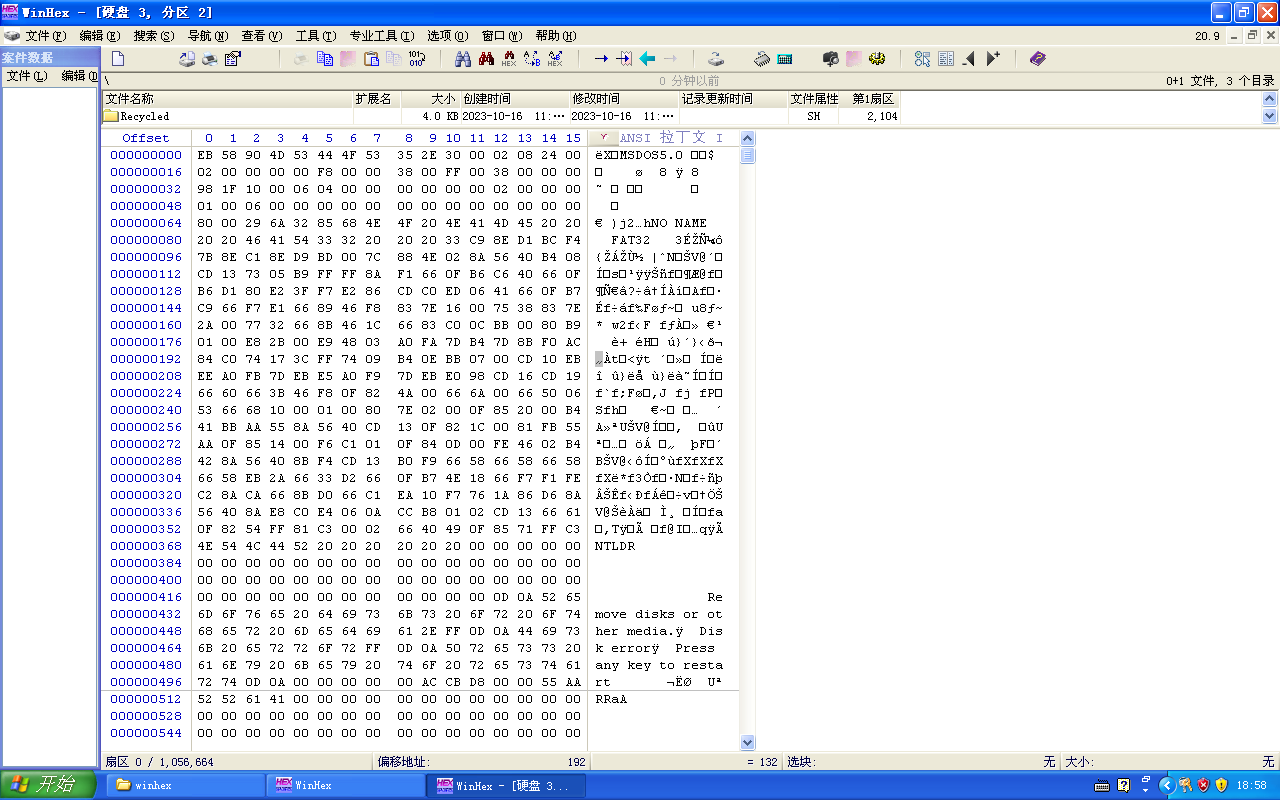


以下分析第一分区和第二分区的信息如下表所示

|  |  |  |
| --- | --- | --- |
| 内容 | 分区1 | 分区2 |
| 分区类型GUID | 28 73 2A C1 1F F8 D2 11 BA 4B 00 A0 C9 3E C9 3B | 16 E3 C9 E3 5C 0B B8 4D 81 7D F9 2D F0 02 15 AE |
| 分区GUID | 19 FA 56 88 60 60 E6 4F B9 28 7C 71 86 D1 91 10 | 25 94 00 C6 2E 94 85 4C B7 7A 31 B1 16 3A D8 C2 |
| 分区起始地址 | 2,048 | 411,648 |
| 分区结束地址 | 411,647 | 444,415 |
| 分区属性 | 00 00 00 00 00 00 00 00 | 00 00 00 00 00 00 00 00 |
| 分区名 | EFI system partition | Microsoft reserved partition |

#### 1.4.10 分析FAT32分区格式的逻辑盘的起始扇区。

用WinHex在虚拟机中打开FAT32分区格式的逻辑盘I的起始扇区如下图所示。



打开其BIOS Paraneter Block如下图所示，



分析其BIOS Paraneter Block的组成和在该逻辑盘中的值如下表所示。

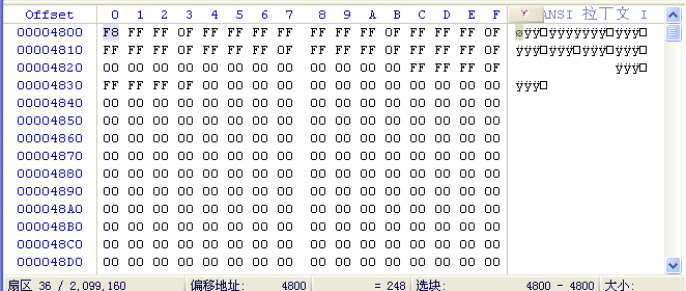
|  |  |  |  |
| --- | --- | --- | --- |
| 偏移量 | 字节数 | 含义 | 值 |
| 0x00B | 2 | 每扇区字数 | 0x0200 |
| 0x00D | 1 | 每簇扇区数 | 0x08 |
| 0xOOE | 2 | 保留扇区数 | 0x0024 |
| 0x010 | 1 | FAT个数 | 0x02 |
| 0x011 | 2 | 根目录项数，FAT32以突破该限制，无效 | 0x0000 |
| 0x013 | 2 | 扇区总数，小于32M使用 | 0x0000 |
| 0x015 | 1 | 存储介质描述 | 0x0F8 |
| 0x016 | 2 | 每FAT表占用扇区数 ，小于32M使用 | 0x0000 |
| 0x018 | 2 | 逻辑每磁道扇区数 | 0x0038 |
| 0x01A | 2 | 逻辑磁头数 | 0x00FF |
| 0x01C | 4 | 系统隐含扇区数 | 0x00000038 |
| 0x020 | 4 | 扇区总数，大于32M使用 | 0x00101F98 |
| 0x024 | 4 | 每FAT表扇区数，大于32M使用 | 0x00000406 |
| 0x028 | 2 | 标记 | 0x0000 |
| 0x02A | 2 | 版本 (通常为零) | 0x0000 |
| 0x02C | 4 | 根目录起始簇 | 0x00000002 |
| 0x030 | 2 | Boot占用扇区数 | 0x0001 |
| 0x032 | 2 | 备份引导扇区位置 | 0x0006 |
| 0x034 | 14 | 保留 | 0x00800000000000  00000000000000 |
| 0x042 | 1 | 扩展引导标记 | 0x29 |
| 0x043 | 4 | 序列号 | 0x6885326A |
| 0x047 | 10 | 卷标 | 转成字符即“NO NAME” |
| 0x052 | 8 | 文件系统 | 转成字符即“FAT32” |

#### 1.4.11查看该逻辑盘的FAT1和FAT2的内容和大小。

打开FAT1和FAT2如下图所示。该逻辑盘中一个扇区占512字节，一簇8个扇区

于是一簇8\*512=4kb

F盘1GB，计算1\*1024\*1024/4得簇数共有262144个，一簇在FAT表中占4字节，故计算得FAT表大小应为4\*262144B=1MB。





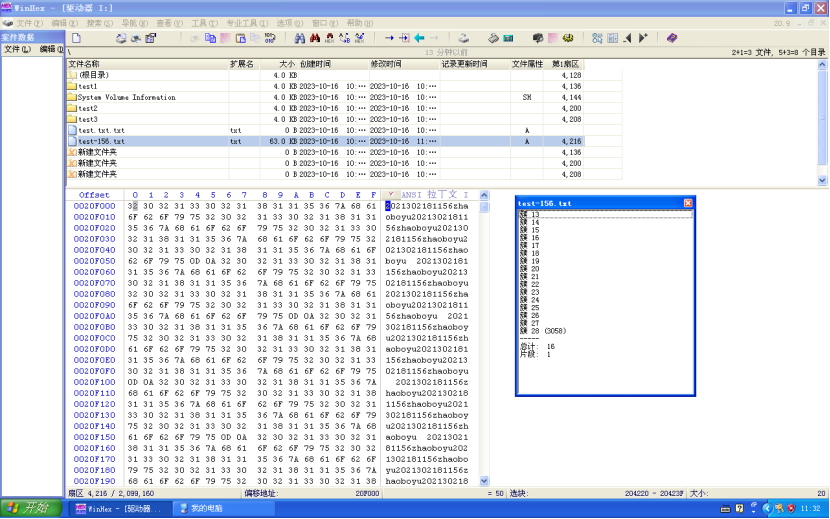
#### 1.4.12查看该逻辑盘的根目录区。

该逻辑盘的根目录去如下图所示。



#### 1.4.14新建文件并分析其目录项

在test-156.txt文件中存有63KB的数据后观察其簇的结构，如下图左下角显示的是该文件的具体存储位置，右下角显示出的是该文件的簇的格式。



得到的该文件的目录项如下所示。00-07代表文件的拓展名，08-0A代表文件的拓展名，14H-15H为起始簇号高16位（0000）。1AH-1BH为起始簇号低十六位（0D00）。1CH-1FH为文件字节长度（0000FBF2）,为64498字节。起始簇号0000000DH，为13号。

IMG_256

#### 1.4.15分析新建文件的簇相关信息

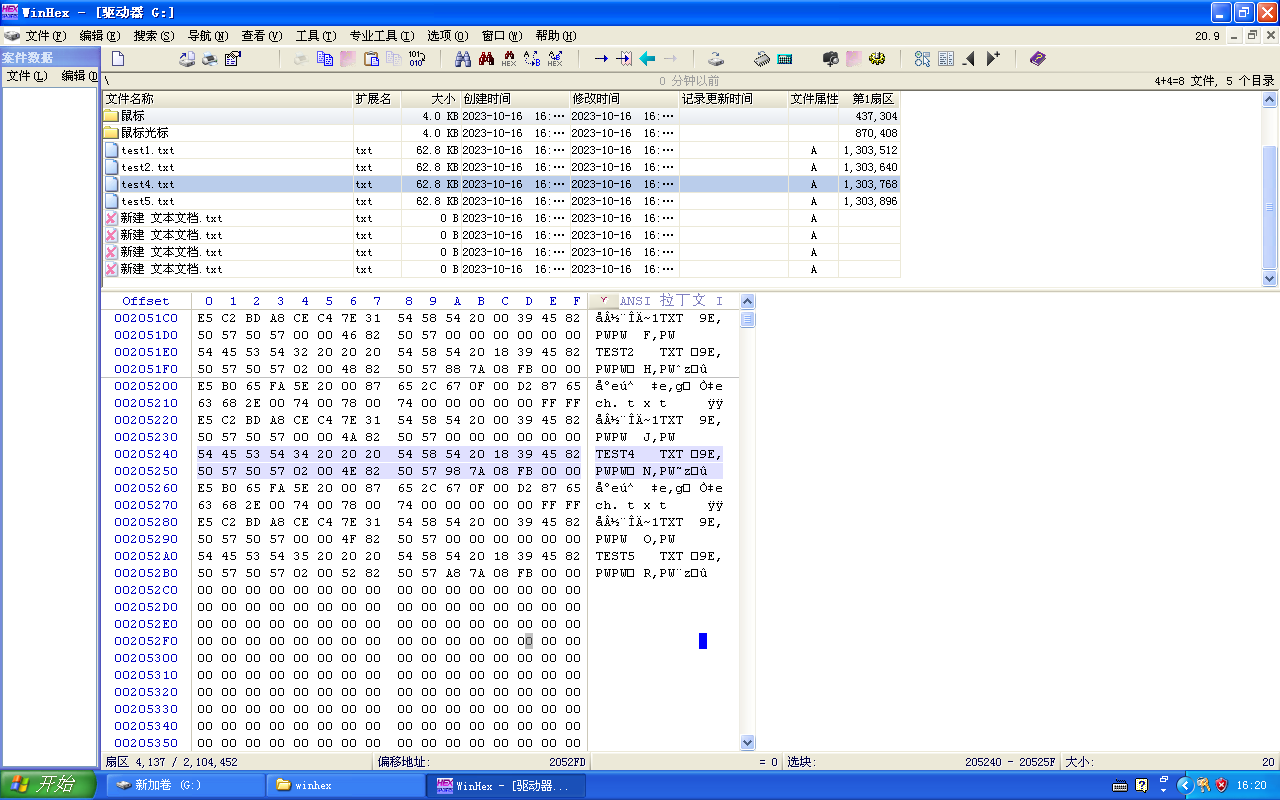
分析得簇列表如下所示，该文件共占有16个簇，因此其簇列表有着16\*4=64B的内容，其首簇数据为0000000E=14指向下一个簇号。最后一个簇的内容为FFFFFF0F

IMG_256

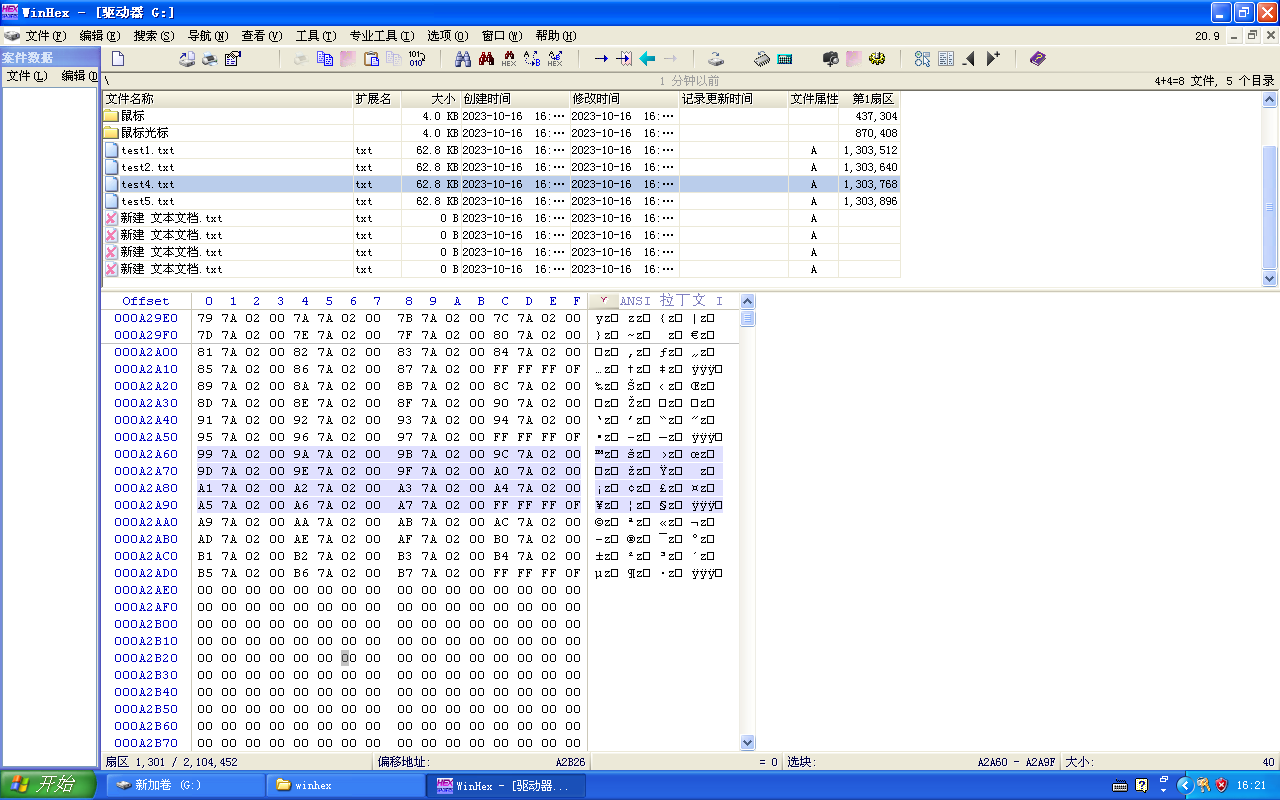
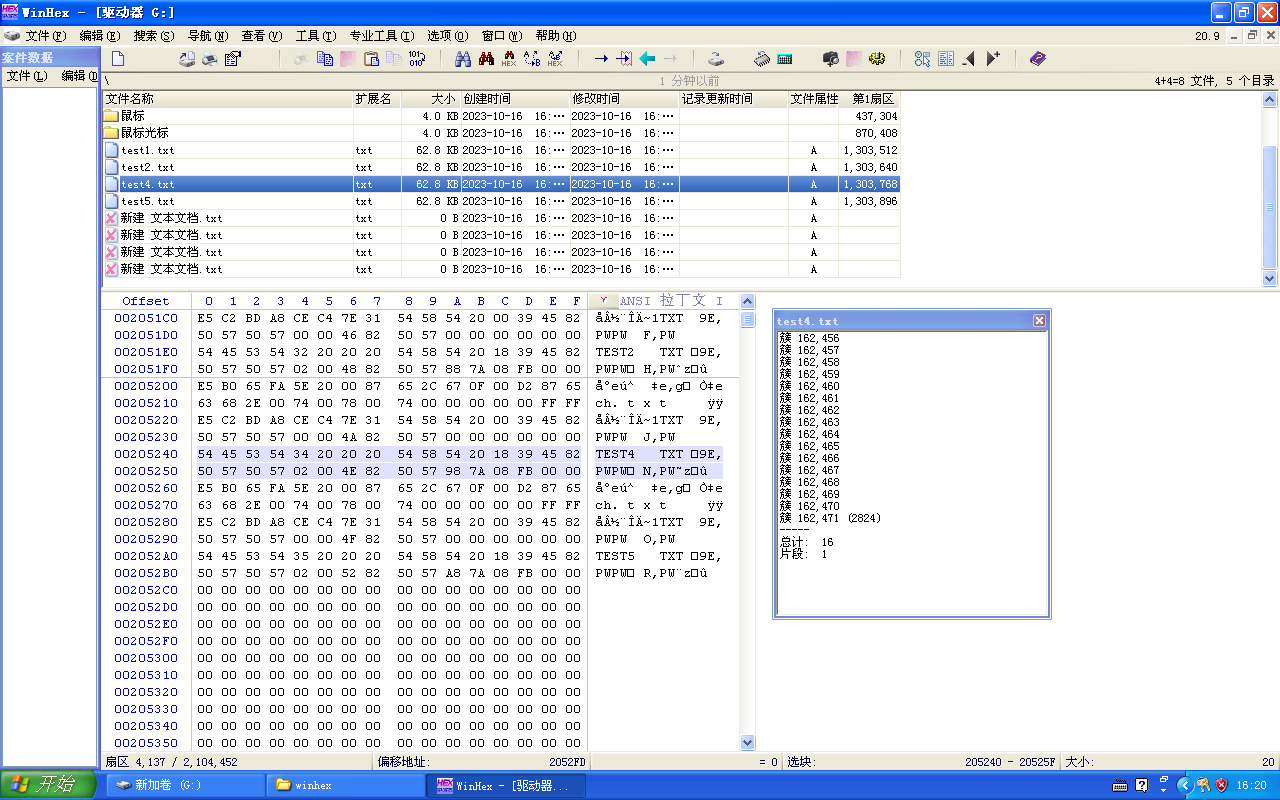
#### 1.4.16删除文件的恢复

（1）恢复前文件信息。

在删除文件之前首先得到该文件的目录表项的内容如下图所示，观察该文件的信息我们可以看到，该文件的簇高位为0002。文件名为test4。

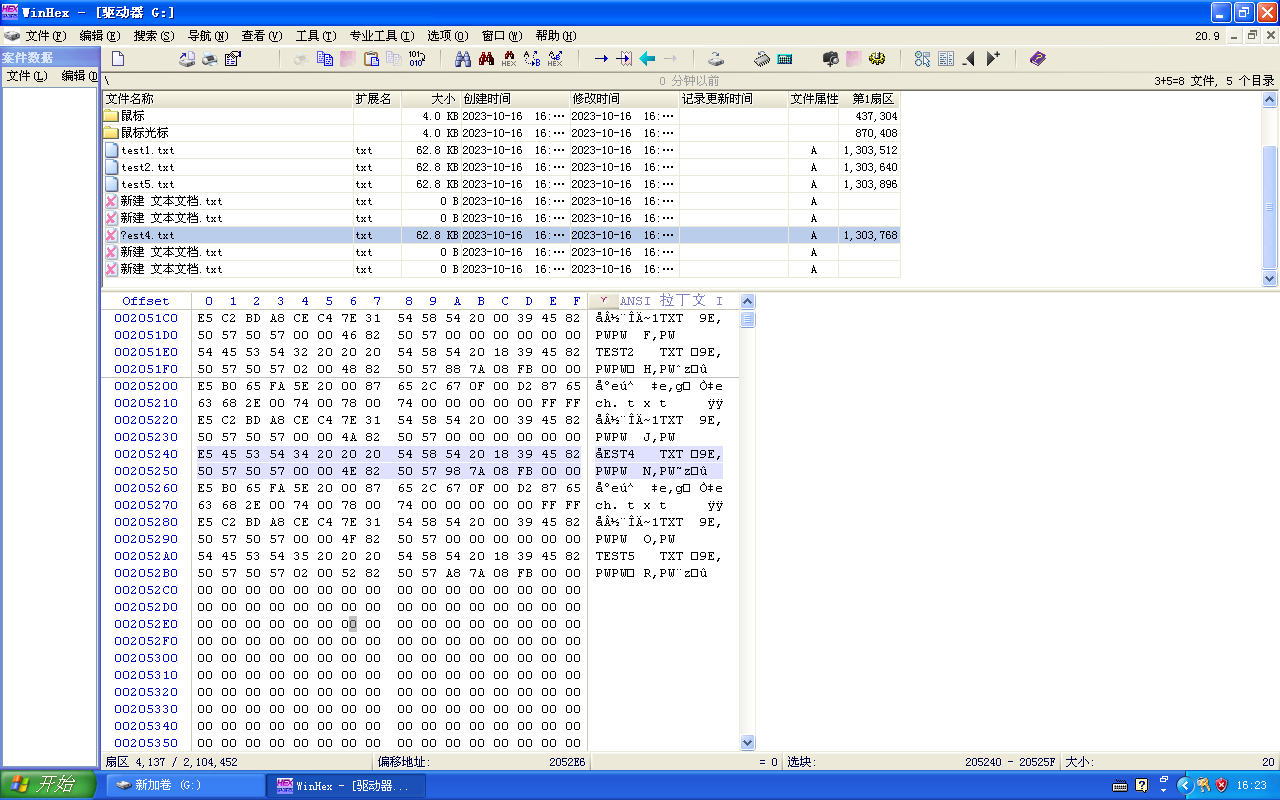


查看该文件的簇列表如下图所示，该文件一共存储在16个簇中，粗链表在内存中保存在A2A60-A2A90区域内。



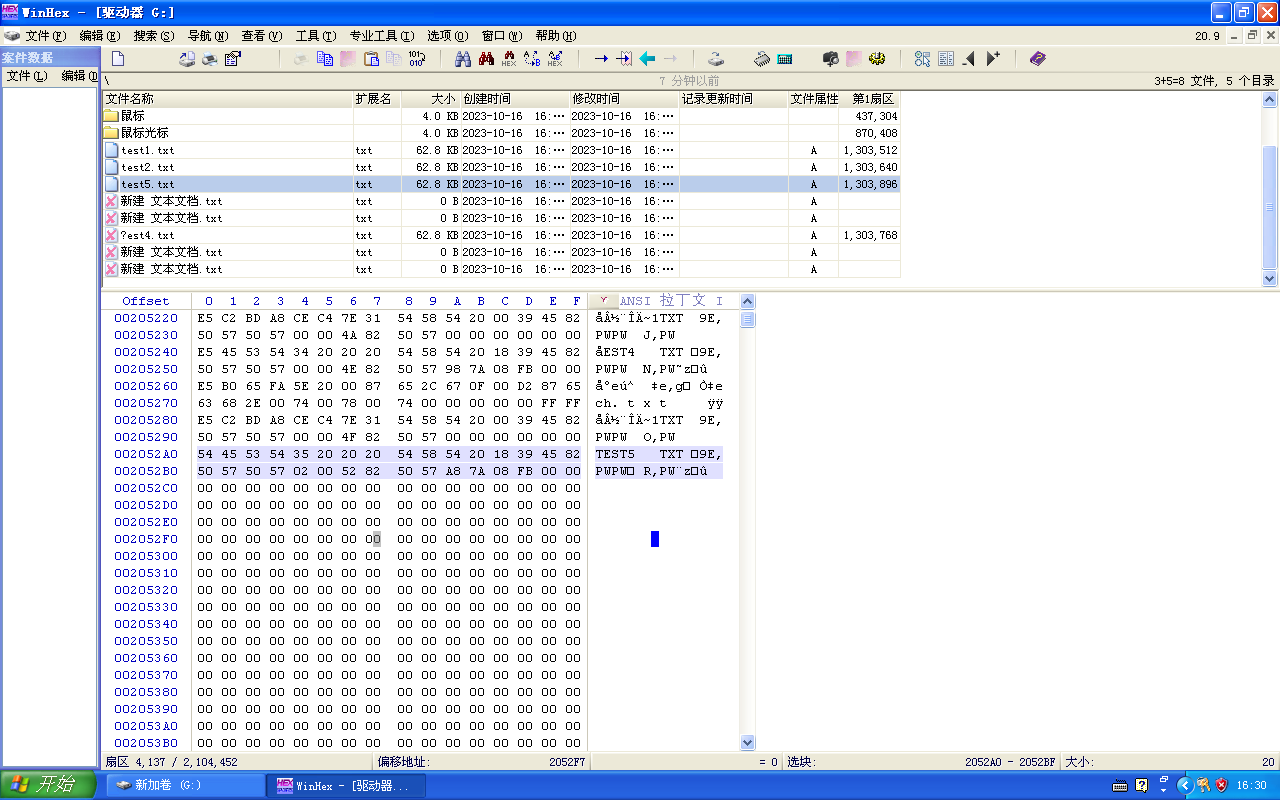
（2）恢复后文件信息。

将文件使用shift+del删除之后查看文件的目录项发现簇高位由原来的0002被清零，簇列表被全部清零。但在文件目录项中仍然能够查看到文件其他位置的目录项。



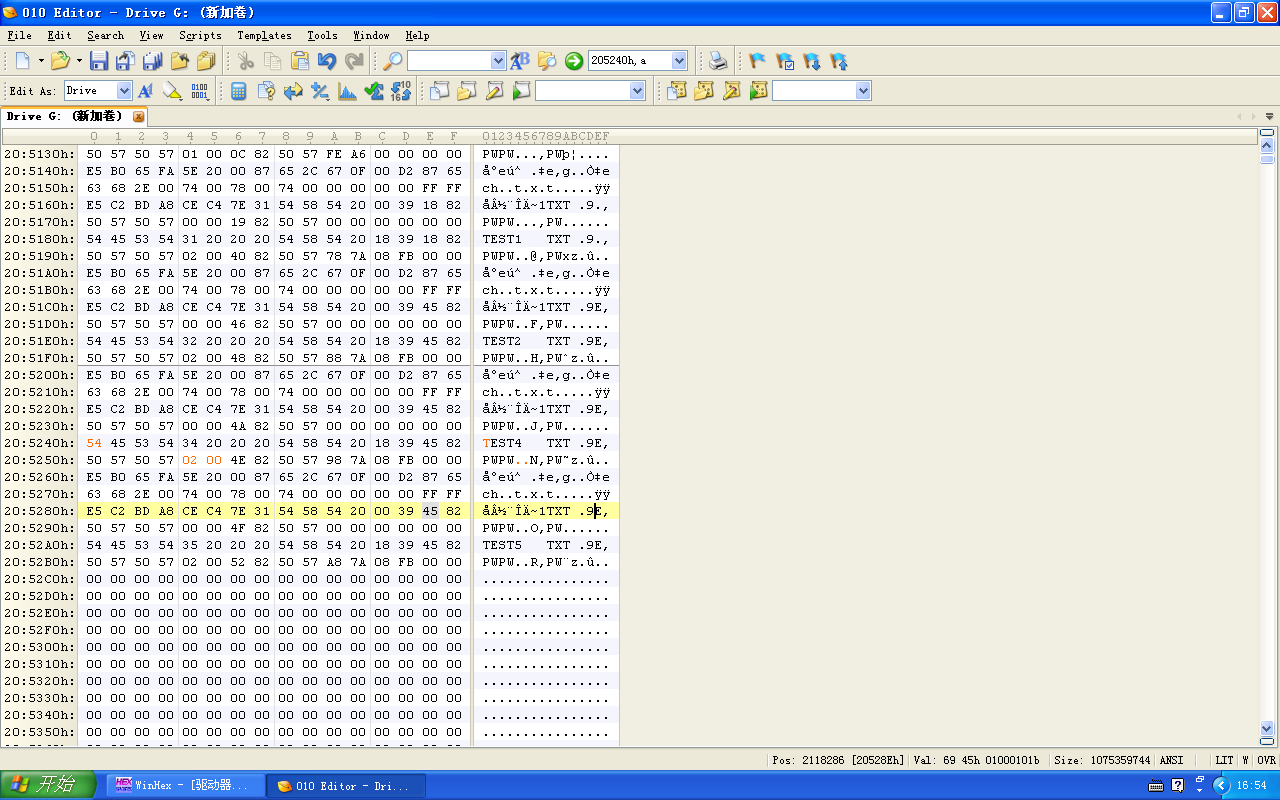
（3）对test4进行数据恢复操作

首先恢复文件目录项，打开test4之前和之后的文件如下图所示，观察其簇高位的数值。



由此推测出该文件的簇高位也应该为02，在010Editor中修改簇高位的值，将其该为我们的推测值02。

然后观察文件目录项可以发现文件的头由原来的54被改为了E5，在此我们将其改回，修改完成的文件目录项如下图所示。

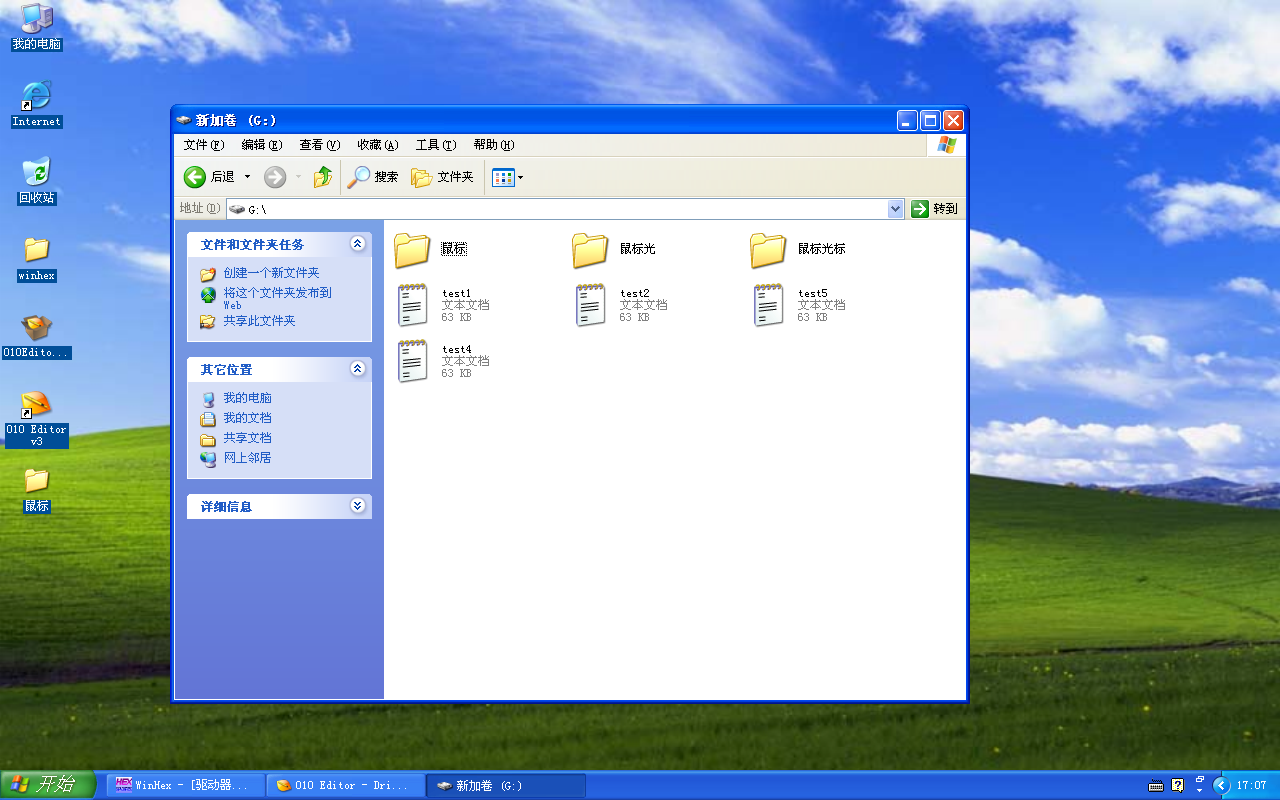


接下来修改簇列表，已知簇高位为02，簇低位为7A 98，由此，我们将簇低位加1后写入第一个簇的前两个字节中去，然后写入簇高位的一个字节，在最后一个字节补上00，由此修改完第一个簇，将后面的簇的簇地位依次加1后填入簇链表中。将簇列表修改完成后的结果如下图所示，



（4）验证结果

重新打开删除文件test4所对应的盘区发现被删除的文件已经被恢复。



#### 1.4.17 课后习题思考

#### （1）在磁盘分区过程中，用户提供了哪些信息？分析分区工具的工作原理。

用户需要提供的信息仅有新分区的文件习统和新分区的空间大小。

工作原理就是首先分区工具读取硬盘上的分区表，获取该磁盘中已知的分区结构和信息，然后根据用户输入的参数计算出新分区的起始位置和大小，并在分区表中创建相应的分区项，同时在磁盘上为新分区分配空间。

然后根据用户提供的文件系统对新分区进行格式化，在新分区的位置创建新的文件系统结构，存储元数据，并准备分区以存储文件和目录。

最后分区工具将对分区表进行更新，将新分区的信息写入分区表中。另外如果是在主拓展分区中还需要创建对应的EBR表项。

#### （2）高级格式化与低级格式化的具体原理和区别是什么？

原理：

低级格式化是指对物理磁盘介质进行处理，重新建立磁道和扇区的布局。在低级格式化过程中，磁盘驱动器会对磁盘表面进行物理处理，包括将磁道划分为扇区、标定磁道和扇区的位置以及校准磁头等操作。低级格式化通常由磁盘驱动器的制造商在生产过程中完成，用户很少需要手动进行低级格式化。

高级格式化是指在低级格式化之后对磁盘进行逻辑上的处理，包括创建文件系统、分配文件系统结构和元数据等。高级格式化通常由操作系统或格式化工具提供，用户可以在使用新磁盘或重新格式化磁盘时执行高级格式化。

区别：

在原理层面，低级格式化处理物理磁盘介质，重新建立磁道和扇区的布局；而高级格式化处理逻辑上的文件系统结构和元数据。

在作用上，低级格式化主要用于磁盘驱动器的生产阶段，由制造商完成，目的是初始化磁盘；高级格式化用于创建文件系统、分区和准备磁盘以存储数据。

在可控性上，低级格式化通常由磁盘驱动器制造商进行，用户很少需要手动进行；而高级格式化由用户或操作系统控制，并提供了更多的选项和灵活性，例如选择文件系统类型、分区大小等。

#### （3）查找资料，对NTFS分区的总体结构进行分析，尝试对NTFS下删除的文件进行手工恢复。

NTFS分区的总体结构包括以下几个部分。

主引导记录。NTFS分区通常位于硬盘的主引导记录中。MBR包含引导代码和分区表，用于引导操作系统和标识磁盘上的分区。

NTFS引导扇区。NTFS分区的第一个扇区是NTFS引导扇区（NTFS Boot Sector），也称为扇区0。它包含了引导代码和关键的文件系统参数，如OEM标识、扇区大小、分区偏移等。

文件分配表。NTFS使用了一种称为文件分配表（File Allocation Table）或MFT（Master File Table）的结构来管理文件和目录的分配和存储。MFT是NTFS中的核心组成部分，记录了文件和目录的属性、大小、权限等信息。

文件记录。NTFS中的每个文件和目录都在MFT中有一个相应的文件记录。文件记录包含了文件的属性、文件名、数据的物理位置等信息。对于较小的文件，文件记录直接存储在MFT中；对于较大的文件，文件记录中存储了指向数据存储位置的指针。

数据存储区。NTFS使用数据存储区来存储文件和目录的实际数据。对于较小的文件，数据存储区可以直接包含文件数据；对于较大的文件，数据存储区中存储了数据的扇区位置。

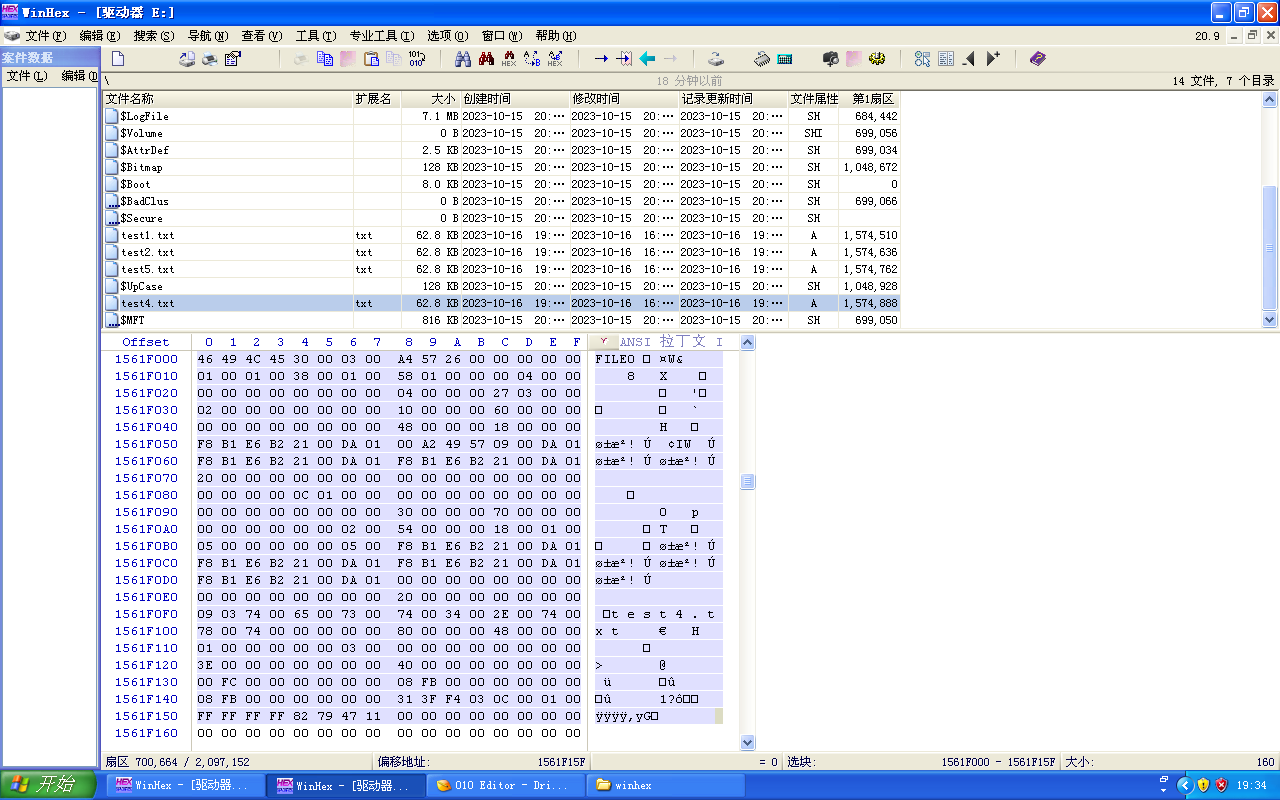
文件属性。NTFS支持各种文件属性，包括标准属性（如文件大小、创建时间、修改时间等）和扩展属性（如压缩属性、加密属性等）。这些属性存储在文件记录中，并提供了对文件的详细描述和管理。

安全描述符。NTFS使用安全描述符来管理文件和目录的访问权限。安全描述符包含了访问控制列表（Access Control List, ACL），用于定义允许或拒绝访问文件的用户和组。

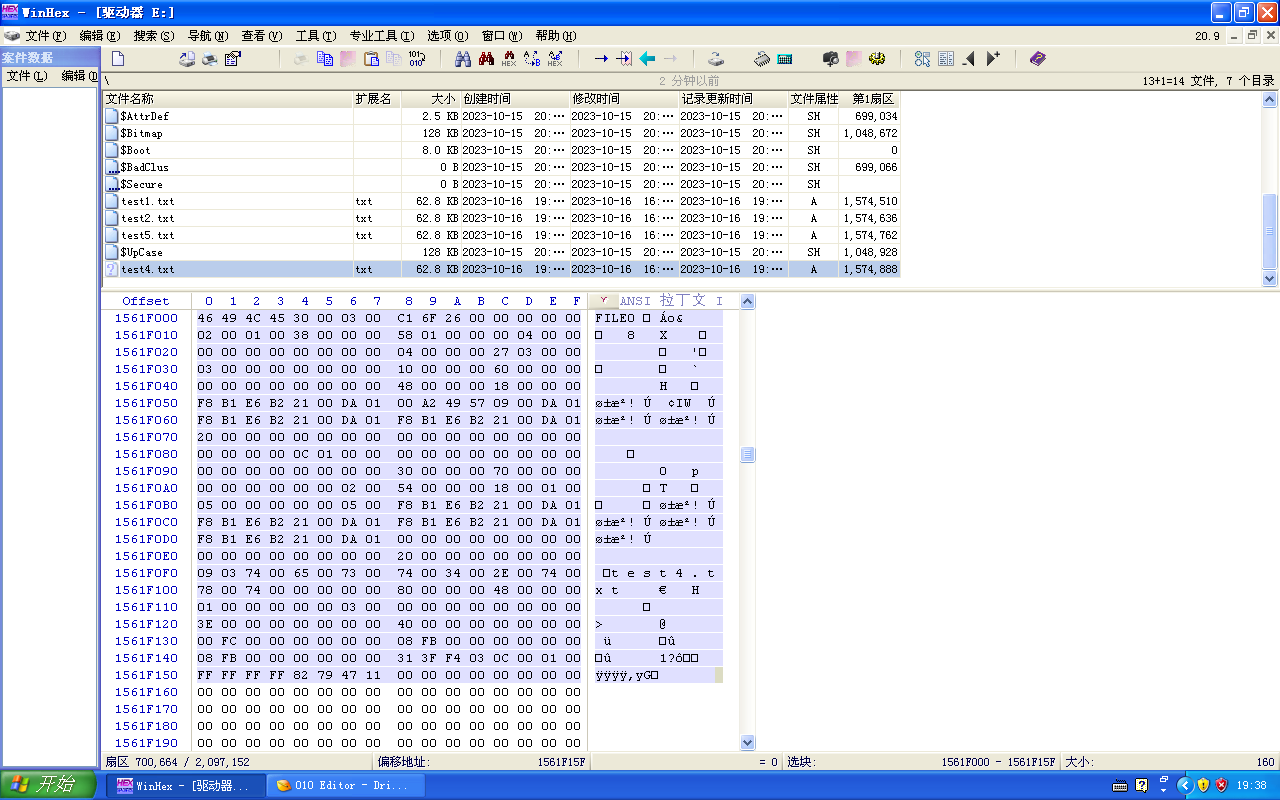
日志文件。NTFS使用日志文件（Log File）来记录对文件系统的更改操作。日志文件可以用于恢复文件系统的一致性，并提供了对系统崩溃或意外断电的保护。

以下对NTFS文件系统中文件进行数据恢复。

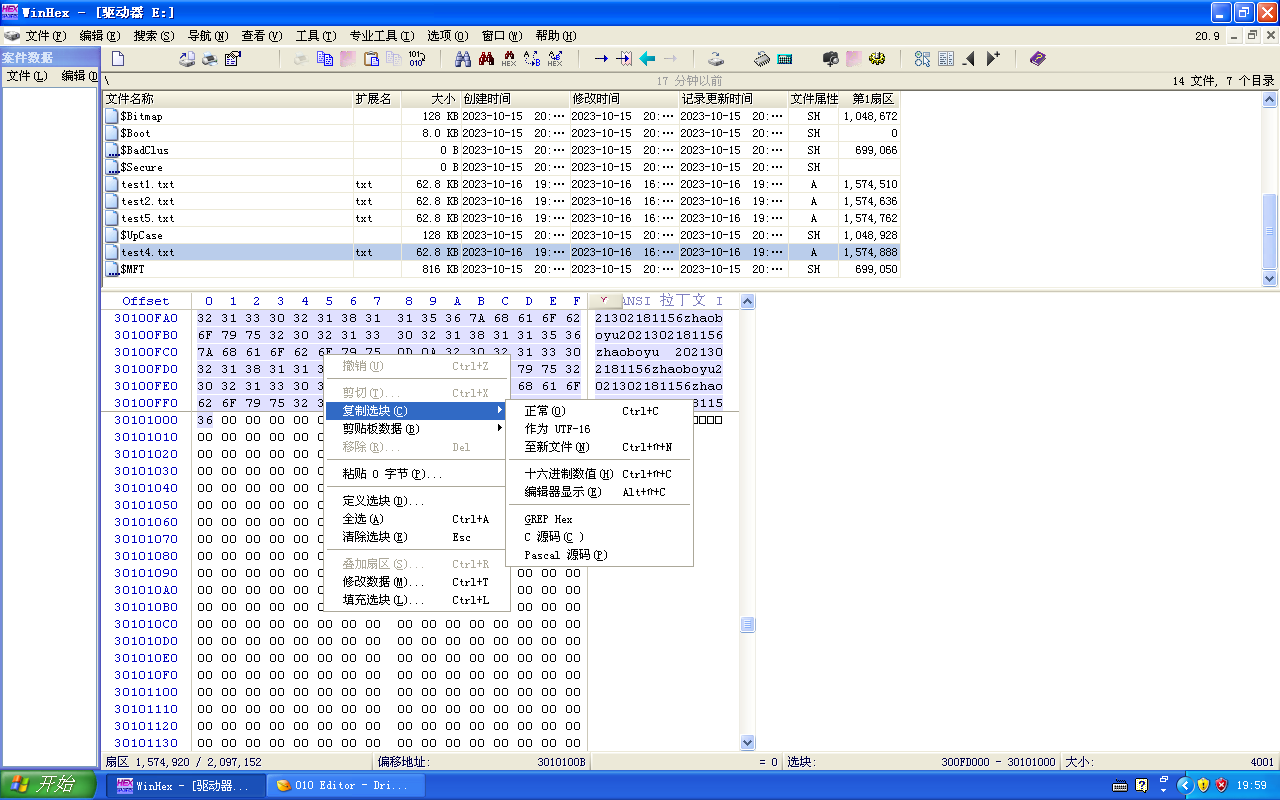
在删除前test4的文件目录如下图所示。



将test4彻底删除后其目录表如下图所示。

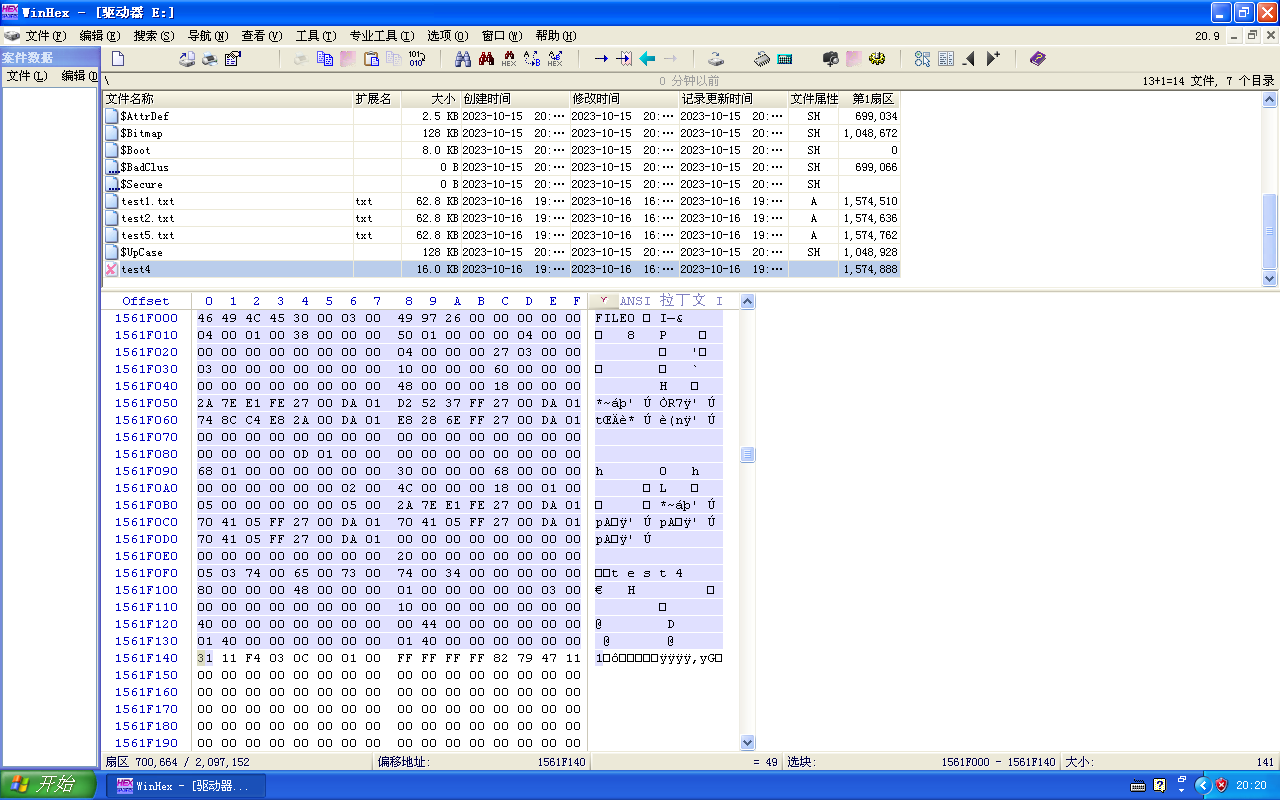


对目录表进行分析，在80H中保存有文件的内容，其中01代表这是非常驻属性，找到文件的簇流信息为31 3F F4 03 0C 可得起始簇位置为0C03F4H=787444,簇的大小为10H=16，使用winhex跳转到起始簇787444的位置，并将其选择为块起始位置，尾部位置为787444+16=787460，将选块复制进新文件即可完成数据恢复。



#### （4）用数据粉碎工具粉碎指定文件，分析其数据粉碎原理。

使用数据粉碎工具对test4进行删除后的结果如下图所示



在test4的原文件目录下的所有数据都已经被重写为其他数据，由此推断数据粉碎工具使用一种或多种覆盖算法来重复写入特定的数据模式到要删除的文件或磁盘空间上。这些算法通常使用随机数据、零值或者特定模式的数据进行覆盖。

#### （5）通过分区表看到的分区字节数为何与资源管理器中看到的分区字节数有差异？

在分区表中，分区字节数通常是通过读取分区表中的元数据来计算得出的。分区表中的元数据包含了有关分区的信息，如起始位置、结束位置等。通过计算这些信息，可以得到分区的字节数。

而资源管理器中看到的分区字节数则是通过读取文件系统的信息来计算得出的。文件系统包含了文件和目录的组织结构以及其在磁盘上的分布情况。资源管理器通过读取文件系统的信息，包括文件分配表和目录结构等，计算得出每个分区的字节数。

#### （6）如果删除的文件是长文件名，如何恢复所有文件名。

如果删除的是长文件名则可以通过查阅文件系统日志来进行数据恢复，某些文件系统会记录文件系统操作的日志，包括文件的创建、修改和删除等操作。通过分析文件系统日志，可以查找已删除文件的相关信息，包括文件名和位置。

或者可以借助文件标识符，在某些文件系统中，文件名和文件数据是分开存储的。文件名通常存储在目录中的文件标识符或类似的数据结构中，而文件数据则存储在不同的位置。通过查找文件标识符，可以找到已删除文件的元数据，包括文件名和其他属性。

#### 1.5实验体会和拓展思考

1.在我们平时使用电子设备时的删除功能可能并不是我们理解的删除，更像是将文件的标签进行了删除，而文件的本体还保存在硬盘中。

2.市面上所存在的众多文件恢复和文件彻底粉碎产品也是借用了这一原理，文件恢复产品通过恢复文件的标签或者说头目录表对文件本体进行恢复，而文件粉碎产品也是通过读取文件本体数据的各个位置将其覆盖为新的随机数据对文件进行改写。