重建 GPT 分区的研究与实现

陈培德,吴建平,钱文华,曹良坤,王林茂
(云南大学 信息学院 云南省高校数字媒体技术重点实验室,云南 昆明 650223)

摘 要: GPT 分区是目前硬盘普遍使用的一种分区形式,克服了 MBR 对分区管理不能超过 2 TB 的缺点。但是由于用户误操作、计算机病毒破坏、突然掉电等因素的影响,使得 GPT 分区被破坏的情况时有发生,从而导致存储在硬盘中的数据丢失。针对这一情况,以 Windows 7 为平台,WinHex 15.08 为分析工具,虚拟硬盘为实验对象,对虚拟硬盘 GPT 分区结构进行分析,通过 DBR 的特征值查找并获得 DBR 及 DBR 备份所在扇区,将 DBR 中存储的总扇区数转换为逻辑盘的总容量,以此为依据重建硬盘 GPT 分区,最后通过 DBR 备份所在扇区恢复 DBR。实验结果表明,当 GPT 分区被破坏后,只要获得逻辑盘总扇区数,便可成功恢复 GPT 分区,从而完整恢复各逻辑盘中的全部数据。

关键词: GPT 分区; MBR 分区; 数据恢复; FAT32 文件系统; NTFS 文件系统
中图分类号: TP311.12 **文献标识码:** A **文章编号:** 1673-629X(2019)02-0096-05
doi: 10.3969/j.issn.1673-629X.2019.02.020

Research and Implementation of Rebuilding GPT Partition

CHEN Pei-de, WU Jian-ping, QIAN Wen-hua, CAO Liang-kun, WANG Lin-mao
(Key Laboratory of Digital Media Technology of Universities and Colleges in Yunnan Province, School of Information Science and Engineering, Yunnan University, Kunming 650223, China)

Abstract: GPT Partition is a common partition form of hard disk at present, which overcomes the disadvantage that MBR cannot exceed 2TB in partition management. However, due to user's misoperation, computer virus damage, sudden power loss and other factors, the destruction of GPT partition occurs from time to time, resulting in the loss of data stored in the hard disk. In response to this situation, based on Windows 7 platform, WinHex 15.08 as the analytical tool and the virtual hard disk as experiment object, we analyze the GPT partition structure. The DBR and its backup in the sector are searched and acquired by the eigenvalues of DBR. The total number of sectors stored in DBR is transformed into the total capacity of the logical disk, based on which the hard disk GPT partition is rebuilt. Finally the DBR is restored by its backup in the sector. The experiment shows that when the GPT partition is damaged, as long as the total number of sectors of the logical disk is obtained, the GPT partition can be restored successfully, so as to completely restore all the data of each logical disk.

Key words: GPT partition; MBR partition; data recovery; FAT32 file system; NTFS file system

0 引 言

GPT 是 globally unique identifier partition table 的缩写,其含义是"全局唯一标识磁盘分区表"[1]。GPT 的出现是为了替代旧式的 MBR(master boot record),主要解决 MBR 分区表不支持容量大于 2.2 TB 的分区问题[2]。

目前,微软公司 Windows 8 使用了 GPT 磁盘分区格式,同时 Windows 8 不再支持 MBR。计算机如果使用 Windows 7 就必须采用 MBR 分区格式,这样不同分区表误操作、误转换的结果是硬盘中原有的磁盘分区表丢失,磁盘中的数据不能正常读取。在 Windows 7 和 Windows 8 用户数量庞大的今天,这种因 GPT 分区表问题导致硬盘中的数据无法读取和使用的问题较为突出。

当 GPT 存储磁盘出现误操作或操作系统本身故障导致的分区表损坏、数据不可见、不可读的数据等问题时,有可能是 GPT 分区表的逻辑出错,可以通过一定的技术手段,将保存在台式机硬盘、笔记本硬盘、服务器硬盘等设备上丢失的宝贵数据进行抢救和恢复。

当 GPT 分区被破坏后,恢复分区常用的方法是:

收稿日期: 2018-02-21 **修回日期:** 2018-06-26 **网络出版时间:** 2018-11-15
基金项目: 国家自然科学基金(61662087,61462093)
作者简介: 陈培德(1966-),男,工程师,研究方向为文件系统与数据恢复技术。
网络出版地址: http://kns.cnki.net/kcms/detail/61.1450.TP.20181115.1050.076.html

使用 DiskGenius 软件的“搜索已丢失分区(重建分区表)”功能来重新建立分区,但该方法只能重建 MBR 分区,不能重建 GPT 分区。

1　GPT 磁盘简介

从整体来看,GPT 磁盘主要由 6 大部分组成,即保护 MBR、GPT 头、GPT 分区表、GPT 分区区域(即文件系统所在区域)、GPT 分区表备份和 GPT 头备份[3]。大致结构如图 1 所示[4](注: 假设 GPT 磁盘的扇区号范围为 0~ n -1,其中 n 为 GPT 磁盘的总扇区数)。

保护 MBR	GPT 头	GPT 分区表	分区区域	GPT 分区表备份	GPT 头备份
0 号扇区	1 号扇区	2~33 号扇区	34~(n -35) 号扇区	(n -34)~(n -3) 号扇区	n -2 号扇区

图 1　GPT 磁盘的整体结构

(1) 保护 MBR。

保护 MBR 位于 GPT 磁盘的 0 号扇区,也是由主引导记录、磁盘签名、MBR 分区表和结束标志 4 个部分组成[5]。在 MBR 分区表中,分区标志为 0XEE,相对扇区为 1,总扇区数为 4 294 967 295,也就是分区总数的最大值,即该磁盘已经被 GPT 分区占用,不能再进行 MBR 分区[6]。

(2) GPT 头。

GPT 头位于 GPT 磁盘的 1 号扇区[6],该扇区是在将 MBR 磁盘转换成 GPT 磁盘后自动生成的,GPT 头定义了 GPT 分区各参数的基本信息[7]。

(3) GPT 分区表。

GPT 分区表位于 GPT 磁盘的 2~33 号扇区,共占用 32 个扇区,每个分区表占 128 字节,最多可以容纳 128 个分区表[7],由于第 1 个分区表为系统保留,所以用户在 GPT 磁盘上最多可以再建立 127 个分区,每个分区表管理一个分区。

(4) 分区区域。

GPT 分区区域是整个 GPT 磁盘中最大的区域,位于 GPT 磁盘的中间位置,GPT 分区区域的开始扇区和结束扇区由 GPT 头定义[7]。一般情况下,开始扇区为 34 号扇区,而结束扇区为 GPT 磁盘总扇区数减去 35。该区域由多个具体的分区组成,如: 微软保留分区、EFI 系统分区、LDM 元数据分区、LDM 数据分区、OEM 分区和主分区。各分区的开始扇区和结束扇区在各分区表中均有定义。

(5) 分区表备份。

一般情况下,分区表备份位于 GPT 磁盘的倒数 33 号扇区~倒数 2 号扇区,也是占用 32 个扇区,是 GPT 分区表位于 GPT 磁盘的 2~33 号扇区的备份。

(6) GPT 头备份。

GPT 头备份位于 GPT 磁盘的倒数 1 号扇区,该扇区也是在将 MBR 磁盘转换成 GPT 磁盘后自动生成的,GPT 头备份也是定义了 GPT 分区各参数的基本信息,但该扇区不是 GPT 头的简单备份,GPT 头备份对 GPT 分区各参数基本信息的定义与 GPT 头对 GPT 分区各参数基本信息的定义稍有不同。

2　实验环境及制作实验素材

2.1　实验环境

(1) 操作系统: Windows 7;

(2) 数据恢复软件及分析工具: WinHex 15.08。

2.2　制作实验素材

(1) 在 Windows 7 操作系统下,使用 Windows 7 的虚拟磁盘管理功能在 D 盘的根目录上建立一个名为 abcd.vhd 的文件,文件大小为 1 GB。

(2) 将 abcd.vhd 文件附加为虚拟磁盘 1,转换成 GPT 磁盘; 在磁盘 1 上依次建立 4 个分区,并对 4 个分区进行格式化操作,磁盘 1 中 4 个分区依次对应 4 个逻辑盘,如下所示:

- H 盘,文件系统: FAT32,容量: 200 MB;
- I 盘,文件系统: NTFS,容量: 160 MB;
- J 盘,文件系统: FAT32,容量: 300 MB;
- K 盘,文件系统: NTFS,容量: 330 MB。

(3) 分别在 H 盘、I 盘、J 盘和 K 盘中存储一些文件夹和文件; 磁盘 1 总体布局如图 2 所示。

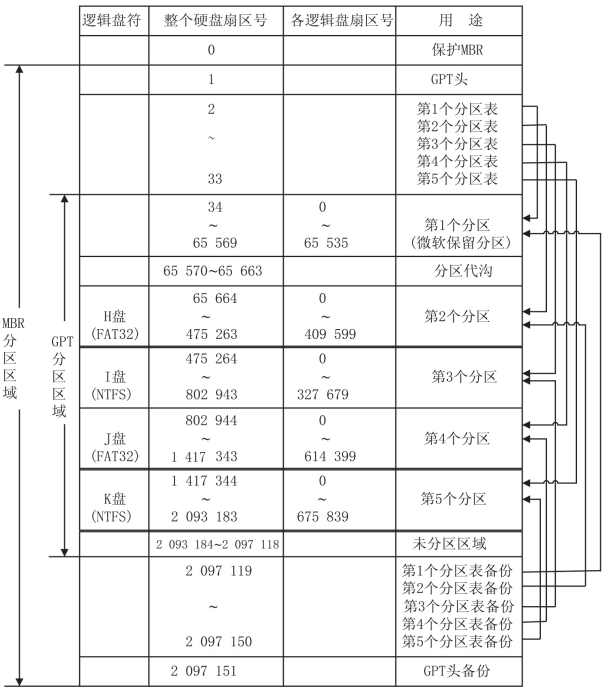

图 2　磁盘 1 的总体布局

(4) 使用 Windows 7 的磁盘管理功能将 H 盘、I 盘、J 盘和 K 盘的 GPT 分区删除。

(5) 将 GPT 磁盘转换为 MBR 磁盘; 磁盘 1 总体布局如图 3 所示。

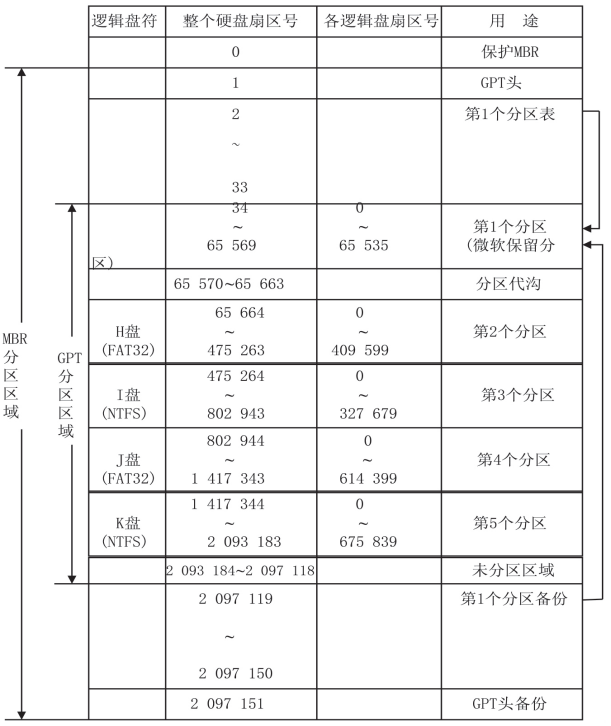

图3 GPT分区删除后的磁盘1总体布局

至此,实验素材已制作完成。

3 删除GPT分区前、后硬盘变化情况对比

删除GPT分区前,各逻辑盘在磁盘1中的分布情况如图2所示。即在磁盘1的2号扇区建立了4个分区(对应微软保留分区、H盘、I盘和J盘),在3号扇区建立了1个分区(对应盘符为K盘),各分区在硬盘中的位置如表1所示。

表1 磁盘1各逻辑盘在硬盘中的位置

分区	开始扇区号	结束扇区号	总扇区数	容量/MB	文件系统
微软保留分区	34	65 569	65 536	32	
H盘	65 664	475 263	409 600	200	FAT32
I盘	475 264	802 943	327 680	160	NTFS
J盘	802 944	1 417 343	614 400	300	FAT32
K盘	1 417 344	2 093 183	675 840	330	NTFS

从删除GPT分区前、后对比硬盘变化可知:

(1)删除GPT分区并转换为MBR磁盘后,0号扇区的保护MBR分区表已经被删除;

(2)GPT头仍然完好保存;

(3)第1个分区表(即对应微软保留分区)仍然保存;

(4)各逻辑盘中的数据仍然完好无损;

(5)第1个分区表备份(即对应微软保留分区)仍然保存;

(6)GPT头备份仍然完好保存。

由于各逻辑盘的GPT分区表和GPT分区表备份已经被删除,虽然各逻辑盘中的数据完好无损,但是通过磁盘管理附加磁盘1,在资源管理器中无法查看到各逻辑盘盘符,也就无法查看各逻辑盘中的文件和文件夹。

4 重建GPT分区的基本思路与方法

经过大量实验发现,在GPT磁盘中建立一个分区后,只要不将对应逻辑盘进行格式化操作,那么系统只将逻辑盘的开始扇区填充为"00",而逻辑盘的剩余扇区即完好无损地保存着。

根据这一特点,恢复GPT分区中各逻辑盘的基本思路与方法如下:

(1)通过开始扇区(即DBR所在扇区)的特征值,查找并记录下各逻辑盘的DBR和DBR备份所在扇区号[3];

(2)通过DBR中的BPB参数,获得各逻辑盘的总扇区数[3];

(3)通过各逻辑盘的总扇区数,计算各逻辑盘的总容量;

(4)通过各逻辑盘总容量,依次建立各逻辑盘(注:在建立各逻辑盘时,不要格式化各逻辑盘);

(5)最后通过各逻辑盘的DBR备份依次恢复各逻辑盘的DBR[3]。

通过磁盘管理附加磁盘1,在资源管理器中可以查看到各逻辑盘的盘符。

5 重建GPT分区的步骤

根据重建GPT分区的基本思路与方法,重建GPT分区的步骤如下:

1.获得4个逻辑盘的基本情况。

(1)在Windows 7操作系统下,启动WinHex;

(2)使用WinHex的文件功能,打开D盘根目录上的abcd.vhd文件,并映像为磁盘;

(3)通过FAT32_DBR的特征值,查找FAT32_DBR及其备份,分别在65 664、65 670、802 944以及802 950号扇区找到;

(4)经过确认,65 664号扇区为FAT32_DBR,而65 670号扇区为65 664号扇区的备份,即FAT32_DBR备份;802 944号扇区为FAT32_DBR,而802 950号扇区为802 944号扇区的备份,即FAT32_DBR备份;

(5)通过NTFS_DBR特征值,查找NTFS_DBR及其备份,分别在475 264、802 943、1 417 344、2 093 183号扇区找到;

(6)经过确认,475 264号扇区为NTFS_DBR,而802 943号扇区为475 264号扇区的备份,即NTFS_

DBR 备份; 802 943 号扇区为 NTFS_DBR,而 2 093 183 号扇区为 802 943 号扇区的备份,即 NTFS_DBR 备份;

(7) 从 65 664、475 264、802 944 和 802 943 号扇区(即各 DBR) 所获得的总扇区数分别为 327 680、614 399、675 840 和 409 599;

(8) 退出 WinHex。

综合步骤 4、步骤 6 和步骤 7,4 个逻辑盘的 DBR、DBR 备份、总扇区数、容量和文件系统如表 2 所示。

表 2　磁盘 1 各逻辑盘的基本情况

分区	DBR 所在扇区号	DBR 备份所在扇区号	总扇区数	容量/MB	文件系统
H 盘	65 664	65 670	409 600	200	FAT32
I 盘	475 264	802 943	327 680	160	NTFS
J 盘	802 944	802 950	614 400	300	FAT32
K 盘	1 417 344	2 093 183	675 840	330	NTFS

注: NTFS_DBR 中存储的总扇区数要比实际分区所占扇区数少 1 个扇区。

2.将 MBR 磁盘转换为 GPT 磁盘。

(1) 在 Windows 7 操作系统下,使用 Windows 7 的虚拟磁盘管理功能附加 D 盘根目录上的 abcd.vhd 文件为磁盘 1;

(2) 将光标移动到“磁盘 1 基本 1 023 MB 联机”处,右击,从弹出的快捷菜单中选择“转换成 GPT 磁盘(V)”。

至此,磁盘 1 由 MBR 磁盘转换为 GPT 磁盘。

3.重建 4 个逻辑盘 GPT 分区表。

(1) 将光标移动到“磁盘 1 的 992 MB 未分配”处,右击,从弹出的快捷菜单中选择“新建简单卷(I) …”;

(2) 出现“新建简单卷向导”第 1 个窗口,单击“下一步”按钮;

(3) 出现“新建简单卷向导”第 2 个窗口—指定卷大小,在“简单卷大小(MB) (S):”右侧的列表框中输入第 1 个逻辑盘的大小“200”,单击“下一步”按钮;

(4) 出现“新建简单卷向导”第 3 个窗口—分配驱动器号和路径,在“分配以下驱动器号(A):”右侧的列表框中选择“H”,单击“下一步”按钮;

(5) 出现“新建简单卷向导”第 4 个窗口—格式化分区,选择“不要格式化这个卷(D):”选项,单击“下一步”按钮;

(6) 出现“新建简单卷向导”第 5 个窗口,单击“完成”按钮;

(7) 重复步骤 1~6,共计 3 次; 在步骤 3 出现“新建简单卷向导”第 2 个窗口—指定卷大小,在“简单卷大小(MB) (S):”右侧的列表框中依次输入第 2 个、第 3 个和第 4 个逻辑盘的大小“160”、“300”和“330”; 在出现“新建简单卷向导”第 3 个窗口—分配驱动器号和路径,在“分配以下驱动器号(A):”右侧的列表框中依次选择“I”、“J”和“K”。

至此,磁盘 1 中的 4 个 GPT 分区已经建立。

将光标移动到“磁盘 1 基本 1023 MB 联机”处,右击,从弹出的快捷菜单中选择“分离 VHD”。

4.通过各自的 DBR 备份恢复 DBR。

由于 4 个逻辑盘的 DBR 均被破坏,所以 4 个逻辑盘的文件系统均为 RAW,需要通过各自的 DBR 备份来恢复,步骤如下:

(1) 启动 WinHex;

(2) 使用 WinHex 的文件功能,打开 D 盘根目录上的 abcd.vhd 文件,并映像为磁盘;

(3) 将 65 670 号扇区复制到 65 664 号扇区; 将 802 943 号扇区复制到 475 264 号扇区; 将 802 950 号扇区复制到 802 944 号扇区; 将 2 093 183 号扇区复制到 1 417 344 号扇区; 然后,存盘并退出 WinHex。

至此,各逻辑盘的 DBR 已经通过 DBR 备份恢复成功。磁盘 1 总体布局如图 2 所示。

通过磁盘管理附加磁盘 1,在资源管理器中可以查看各逻辑盘的盘符,并且可以看到各逻辑盘中的所有文件和文件夹。

6　通过重建 MBR 分区来恢复各逻辑盘中的数据

如果硬盘总容量小于 2.2 TB,且分区总数小于或等于 4 个时,可以使用在硬盘 0 号扇区建立对应 MBR 分区表的形式来恢复各逻辑盘中的数据。

由于磁盘 1 正好满足这一条件,所以可以通过在硬盘 0 号扇区建立 4 个 MBR 分区表[10]。方法如下:

1.从表 2 中 4 个逻辑盘的 DBR 所在扇区号和总扇区数可以计算出磁盘 1 中 H 盘、I 盘、J 盘和 I 盘在硬盘 0 号扇区的分区表,如表 3 所示。

表 3　分区表

序号	盘符	对应分区表
1	H 盘	00 0101 00 0C FE FF FF 80 00 01 00 00 40 06 00
2	I 盘	00 0101 00 07 FE FF FF 80 40 07 00 00 00 05 00
3	J 盘	00 0101 00 0C FE FF FF 80 40 0C 00 00 60 09 00
4	K 盘	00 0101 00 07 FE FF FF 80 A0 15 00 00 50 0A 00

对 4 个分区表说明如下:

(1) 由于 H 盘、I 盘、J 盘和 K 盘均不引导系统,各分区表中第 1 个字节的值为“00”[11];

(2) 目前硬盘的存取方式均为 LBA,分区表中第

2~4字节未定义，可以填充任意值，这里填充“01 01 00”[12]；

(3) 由于H盘和J盘的文件系统为FAT32，所以，分区标志为“0C”[13]；而I盘和K盘的文件系统为NTFS，所以，分区标志为“07”[14]；

(4) 分区表中第6~8字节未定义，可以填充任意值，这里填充“FE FF FF”；

(5) 分区表中第9~12字节为相对扇区，即各逻辑盘DBR所在扇区号；

(6) 分区表中第13~16字节为总扇区数，即各逻辑盘所占扇区数。

2.将这4个分区表填入到硬盘0号扇区偏移0X01BE~0X01FD处，然后存盘并退出WinHex。磁盘1总体布局如图4所示。

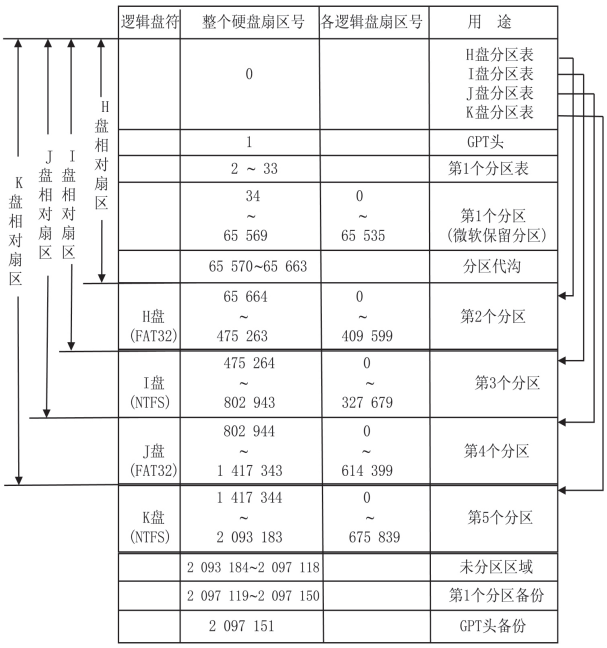

图4 恢复4个MBR分区后的磁盘1总体布局

通过磁盘管理附加磁盘1，在资源管理器中可以查看各逻辑盘的盘符，并且可以看到各逻辑盘中的所有文件和文件夹。

7 结束语

综上所述，当GPT磁盘中的分区被删除，并将GPT磁盘转换为MBR磁盘后，只要查找并获得各逻辑盘的DBR，通过DBR中的总扇区数，计算出各逻辑盘的总容量；依次重建各逻辑盘，在建立各逻辑盘时，只要不格式化各逻辑盘，最后通过各逻辑盘的DBR备份恢复各自的DBR，便可以恢复GPT磁盘中各逻辑盘的全部数据；如果GPT磁盘总容量小于2.2 TB，且分区总数小于4个时，也可以在硬盘0号扇区通过重建MBR的形式来恢复各逻辑盘中的全部数。

参考文献：

[1] 刘 伟.数据恢复技术深度揭秘[M].北京：电子工业出版社，2010：179.

[2] 杨 倩.数据恢复备份实训教程[M].北京：电子工业出版社，2016：92.

[3] 马 林.数据重现—文件系统原理精解与数据恢复最佳实践[M].北京：清华大学出版社，2009：84.

[4] 汪中夏，张京生，刘 伟.RAID数据恢复技术揭秘[M].北京：清华大学出版社，2010：150.

[5] 陈培德，吴建平，王丽清.重建NTFS的DBR及分区表的研究与实现[J].实验科学与技术，2016，14(6)：56-59.

[6] RUSSINOVICH M E，SOLOMON D A，LONESCU A.Windows internals[M].Beijing：Post & Telecom Press，2009：938.

[7] CARRIER B.File system forensic analysis[M].[s.l.]：Addison Wesley Professional，2005.

[8] 陈培德，吴建平，王丽清.NTFS文件系统实例详解[M].北京：国防工业出版社，2015：10.

[9] SOLOMAN D A.Inside Windows NT[M].2nd ed.Washington，U.S.A：Microsoft Corporation，1998：330.

[10] 戴士剑，涂彦晖.数据恢复技术[M].北京：电子工业出版社，2005：226.

[11] IVENS K，GARDINILE R K.Windows 200：the complete reference[M].Beijing：China Machine Press，2000：532.

[12] 陈培德，吴建平，王丽清.Ghost后数据恢复的研究与实现[J].计算机技术与发展，2017，27(1)：112-116.

[13] 陈培德，吴建平，王丽清.重建分区表与FAT32_DBR研究与实现[J].计算机技术与发展，2016，26(10)：188-191.

[14] 刘乃琦，郭建东，张 可.系统与数据恢复技术[M].北京：电子科技大学出版社，2008：46.