

地震仪与地质雷达的数据存储格式及其读取

杨天春 王齐仁 刘新华

(湖南科技大学土木工程学院, 湖南湘潭 411201)

【摘要】 近年来,我国从国外引进了不少浅层地震仪和地质雷达设备,但厂家一般都不提供其数据文件存储格式。根据 SEG 公布的 SEG-2 数据存储格式,笔者对它进行了认真的探讨,并在 Matlab 语言环境下编制了相应的数据读取软件。实际应用结果表明,该程序的读取方法是正确的;且对于以 SEG-2 方式存储的任何类型的仪器设备数据,只需对读取字符串部分的语句稍加修改,就可满足其数据读取的需要。

【关键词】 SEG-2 格式;地震仪;地质雷达;Matlab

【中图分类号】 P 631.443

File Storage Format of Seismograph and GPR and Its Reading Method

Yang Tianchun Wang Qiren Liu Xinhua

(College of Civil Engineering, Hunan University of Science and Technology, Xiangtan Hunan 411201 China)

【Abstract】 Recently, many instruments of seismograph and GPR are imported. Manufacturers unusually provide the file format. The authors study the file format of SEG-2 data, and write a corresponding routine for reading the data in Matlab programming language. A practical application is used for the reading routine. The results show that the route is correct, and this route is easy to maintain and require a minimum of revision for a kind of instruments' SEG-2 files.

【Key Words】 SEG-2 format; seismograph; ground penetrating radar(GPR); Matlab

0 引言

随着工程物探的迅速发展,浅层地震、地质雷达方法被广泛应用于工程与环境问题的勘察中。这也使得浅层物探方法及其仪器的研制进入了一个崭新的时代。由于各仪器生产厂家对数据的记录都有自己特定的存储格式和其相应的软件包,为此,1987年10月,SEG(The Society of Exploration Geophysicists)工程地球物理会议规定了一种统一的数据文件存储格式,即 SEG-2 格式^[1]。在此之前,SEG 曾推荐过 SEG-Y、SEG-D 数据存储格式^[2-4]。但这两种格式存在一些缺陷,因此需要一种新的、更实用的存储标准。SEG 建议自 1987 年 10 月之后生产的浅层地震仪、地质雷达的数据存储格式均采用 SEG-2 标准,若厂家不采用 SEG-2 存储格式,就必须提供能将其数据转换成 SEG-2 格式的相应软件。

我国近年来购进了许多国外的数字地震仪(如 TSP203)和地质雷达系统,在应用国外仪器设备的基础上,我国各科研生产单位也逐步在研制国产的仪器设备,为此,我们必须弄清 SEG-2 的数据存储

方式,便于国产仪器的研制开发工作与国际标准接轨,也有利于国产仪器的改进。

1 SEG-2 存储格式

地震与雷达的原始数据与处理结果是分离的,但目前 SEG-2 格式可将原始数据与处理后的结果存放在一起,如瑞士安伯格测量技术有限公司研制的 TSP(Tunnel Seismic Prediction)²⁰³ 采集系统就是如此。同时,SEG-2 格式不仅适用于 DOS 支持的 IBM 微机及兼容机,它还适用于 OS/2 操作系统、Macintosh 系列微机和 UNIX 工作站。

SEG-2 采用了自由格式存储数据,它不依赖于任何编程语言,可为所有处理地震(或雷达)数据的计算机和 workstation 提供共享数据,并且该数据结构标准也便于维护和今后的修订。

SEG-2 数据文件由一个文件描述块(File Descriptor Block)、一个或多个道描述块(Trace Descriptor Block)及相应的数据块(Data Block)组成(见图 1)。文件的开头部分是文件描述块,接着为第 1 道描述块(Trace Descriptor Block 1),紧接着的是相

基金项目: 湖南科技大学博士基金资助项目(E50607);西部交通建设科技项目(200331879802)

作者简介: 杨天春,1968 年生,男,汉族,湖南津市人,副教授,主要从事地球物理学与岩土工程方面的教学和科研工作。E-mail: ytc6803@yahoo.com.cn

应的第1道的数据块(Data Block 1),再后面的即为其余各道的描述块及其相应数据块。

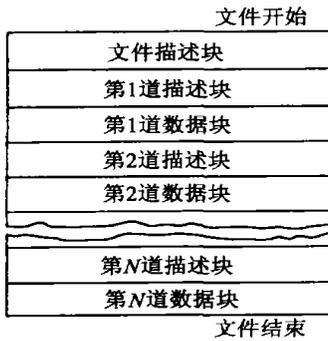


图1 SEG-2文件总结构

1.1 文件描述块的结构

文件描述块的内容说明了数据文件本身的结构,它主要由3部分组成(见图2):①前32字节存放了块标识符、版本号、道指针子块大小、采样道数、字符串结束符和行结束符等;②道指针子块(Trace Pointer Subblock)存储了每道描述块在文件中的起始位置;③一系列与本数据文件相关的操作字符串,如采集日期、采集时间、仪器号、操作员等。

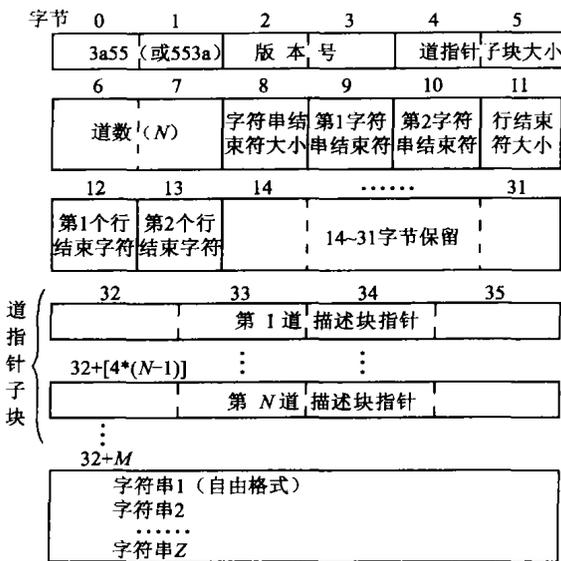


图2 文件描述块结构

文件描述块的第0、1字节存有十六进制数3a55h(或553ah),这是一个文件描述块标识符(File Descriptor Block ID),该整数表明了该地震(或雷达)文件的存储方式。如果文件的第1个字节(即0字节)是55h,则表明文件的存储是低字节在前、高字节在后,这就如同IBM系列微机所采用的存储方式;反之,若第1个字节为3ah,则文件的存储方式是低字节在后、高字节在前,如UNIX操作系统下的68000系列机的数据文件就采用该存储方式。

第2、3字节含有一个整数(integer),它指明了SEG标准的版本号(Revision Number)。第4、5字节是一个无符号整数(unsigned integer)M,它指明了道指针子块的大小(单位为字节),M必须能被4整除,因为每道的指针是一个无符号长整数(unsigned long integer),占4个字节,M的取值范围为[4,65532]。第6、7字节是一个无符号整数N,它指明了数据的采样道数,N必须小于或等于M/4,其取值范围为[1,16383]。

第8字节存放的为01h或02h,这主要取决于第9、10字节所存放的字符串结束符(String Terminator)个数,如果字符串结束符只用一个字符,则该字符存放于第9字节中,第10字节为空;否则,结束符被存放于9、10字节中。字符串结束符是1个或两个不可打印的ASCII字符,其十进制代码所取的范围为0~31,该结束符起着将文件描述块和道描述块中的多个字符串分开的作用。

行结束符(Line Terminator)是一个或两个不能打印的ASCII字符,一般采用回车(CR)或同时采用回车、换行(LF),该字符主要是将NOTE字符串中的多个文本行相互分离开来。与字符串结束符的存储方式相似,第11存放的是行结束符的大小(01h或02h),第12、13字节存放着行结束符,如果行结束符只用一个字符,则只存放于第12字节。

第14~31字节是SEG-2保留部分。第32~31+M字节为道指针子块,每个指针数为占4个字节的无符号长整数,其含义前面已述。

道指针子块后面存放着一系列自由格式的字符串,其存放方式后面再详述。

1.2 道描述块的结构

道描述块中记录了相应道数据块的采样点数、数据存储格式等诸多信息(见图3)。

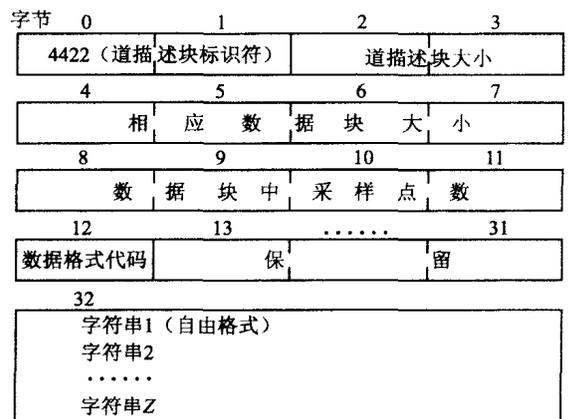


图3 道描述块结构

SEG-2 规定数据文件中至少含有一个道描述块和相应的数据块。

道描述块中第 0、1 字节存储一个无符号整数 4422h(十六进制),这是道描述块开始的标识符(Trace Descriptor Block ID)。第 2、3 字节含有一个无符号整数,它指明了该道描述块的大小(单位为字节),其值的取值范围为[32, 65532]。

第 4~7 字节含有一个无符号长整数,它指明了紧随其后的数据块大小(单位为字节),该数必须能被 4 整除。第 8~11 字节也是一个无符号长整数,其值大小指明了该道数据块中的采样点数。

第 12 字节定义了数据块中数据的存储格式代码(Data Format Code):①01h,数据格式为 16 位整数(16-bit fixed point);②02h,数据格式为 32 位整数(32-bit fixed point);③03h,20 位浮点数(20-bit floating point)(SEG-D 标准);④04h,32 位浮点数(32-bit floating point)(IEEE 标准);⑤05h,64 位浮点数(64-bit floating point)(IEEE 标准)。

第 13~31 字节是保留的,第 32 字节至道描述块结束为一系列的与数据采集和处理相关的信息,

它们以字符串的形式存储,其存储方式与文件描述块内字符串存储方式相同。

1.3 字符串的存储格式

文件描述块和道描述块中的字符串包含有许多必要的操作信息,每个字符串前 2 个字节为一个整数,其大小就是该字符串所占的字节数(前面整数所占的 2 个字节被包括在内),字符串中包含有关键词和与之相关的数据,它们之间用空格(spaces)或制表符(tabs)分开。关键词中所有的希腊字母全为大写,关键词(keywords)中的字(word)与字之间是用下划线分开;与关键词相关的变量(arguments)内不存在空格,并用大写字母表示。除 NOTE 字符串之外,所有字符串的存储是按字母顺序进行排列,NOTE 字符串被存放在描述块的最后。在所有字符串的最后是一个值为 0 的整数,占 2 个字节,这是字符串结束的标志。

2 Matlab 下 SEG-2 数据文件的读取

根据 SEG-2 数据文件的存储格式,笔者在 Matlab 语言环境下编制了读取地震原始数据的程序,数据来源于某防爆地震仪的野外采集结果。由于篇幅所限,下面仅给出某些关键的语句:

```

fid01=fopen(file1,'rb')           %打开 SEG-2 数据文件
.....
M=fread(fid01,1,'uint16');        %从文件描述块中读取道指针子块大小
N=fread(fid01,1,'uint16');        %从文件描述块中读取采样道数(N)
.....
temp=fread(fid01,1,'int16');       %读取道描述块的 0,1 字节
char10=dec2hex(temp);
B_size=fread(fid01,1,'uint16');    %读取道描述块的字节大小
D_size=fread(fid01,1,'ulong');     %读取数据块的字节大小
S_number=fread(fid01,1,'ulong');   %读取数据块中的采样点数(S_number)
D_format=fread(fid01,1,'int8');     %读取数据存储格式代码(D_format)
chx=dec2hex(D_format);
.....
temp=fread(fid01,1,'int16');
sample=fread(fid01,temp-2,'char'); %在道描述块中读取采样间隔字符串(SAMPLE_INTERVAL)
sample=char(sample');
[chx,chxx]=strtok(sample,'0');     %从采样间隔字符串中分离出采样间隔
interval=str2num(chxx);            %将采样间隔转换为浮点数据(interval)
.....
%下面是对数据的读取
for i=1:N
    frewind(fid01);                %将读数指针拨至文件开始位置

```

```

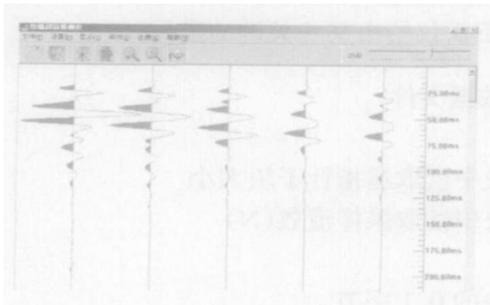
fseek(fid01,Pointers(i),'cof');    %将读数指针拨至第 i 道的道描述块起始位置
temp=fread(fid01,1,'int16');      %读取道描述块的 0,1 字节
char10=dec2hex(temp)
B_size=fread(fid01,1,'uint16');  %读取道描述块的大小
fseek(fid01,B_size-4,'cof');      %将读数指针拨至数据块的起始位置
for j=1:S_number
Data(i,j)=fread(fid01,1,format);
end
end
fclose(fid01);
.....                               %以文本文件格式存储数据结果(略)

```

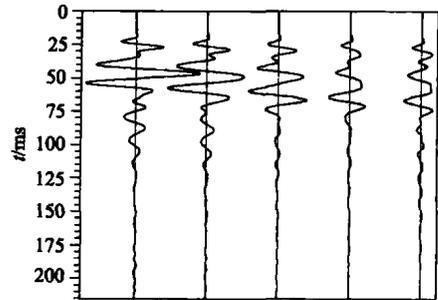
从以上的程序来看,在读取信息和数据时,除了要将读数指针拨至文件的正确位置外,还比较重要的一点是需要从文件描述块和道描述块中读出采样道数(N)、采样点数(S_number)、数据的存储格式、采样间隔(interval)等几个重要的参数,因为它们是正确的读取(或存储)数据的几个关键变量。

为说明程序的正确性,我们将程序应用于实践,

并将读出的某个数据文件在 Golden Software(GRAPHER)软件中绘制曲线图,再将结果与仪器处理软件的读取结果进行对比(见图4),其中图4(a)为仪器软件读出的前5道数据曲线,图4(b)为笔者根据读出的数据用 GRAPHER 绘制的前5道曲线,这两幅图都只显示前200ms左右的结果。由图4的曲线对比来看,编制的 Matlab 读数软件程序是正确的。



(a)原始数据曲线



(b)读出的数据结果曲线

图4 原始数据曲线与读出的数据结果曲线

3 结论

本文详细介绍了 SEG-2 数据文件的存储格式,并根据某地震仪的采集结果,在 Matlab 语言环境下编制了其数据读取程序。由于篇幅所限,笔者仅给出了关键语句部分,读者在此基础上很容易扩充完善。

SEG-2 数据文件的格式是目前 SEG 公布的最新的地震、雷达数据文件存储方式,为与国际接轨,国内仪器生产厂家也开始逐渐采用此存储格式。弄清了国外仪器的数据存储方式之后,便可更好地利用好国外的仪器设备及其软件系统,从中学习其长处,为国产仪器的研制和软件开发具有很好的促进作用。

参 考 文 献

- [1] Pullan S E. Recommended standard for seismic (/radar) data files in the personal computer environment [J]. Geophysics, 1990, 55(9): 1260-1271.
- [2] Northwood E J, Weisinger R C, Bradley J J. Recommended standards for digital tape formats [J]. Geophysics, 1967, 32(6): 1073-1084.
- [3] Meiners E P, Lenz L L, Dalby A E, et al. Recommended standards for digital tape formats [J]. Geophysics, 1972, 37(1): 36-44.
- [4] Barry K M, Cavers D A, Kneale C W. Recommended standards for digital tape formats [J]. Geophysics, 1975, 40(2): 344-352.