



西北工业大学
NORTHWESTERN POLYTECHNICAL UNIVERSITY

C程序设计 Programming in C



1011014

主讲：姜学锋，计算机学院

编程处理各种格式的文件

- ◆ 1、从文件中读写二进制数据.....
- ◆ 2、成千上万的文件格式.....

10.3.5 读写数据块

- ▶ fread函数从文件中读入指定数目指定记录大小的数据块，其函数原型为：

```
size_t fread(void *buffer, size_t size, size_t count, FILE *stream);
```

10.3.5 读写数据块

- ▶ 参数stream是已打开的文件指针，该文件必须是以读或读写方式打开的。buffer是数组或能容纳成批数据的内存区起始地址（如动态分配得到的内存区），用于存储读取到的数据。参数size表示所读记录的大小，如读字节数据时，记录为字节，大小为1；又比如读结构体数据时，记录为结构体，大小为结构体长度。参数count表示欲读取多少个记录，如为单个记录，则count为1，如为一批记录（数组），则count为记录数或数组长度。因此，使用fread函数读到的总字节数等于。

10.3.5 读写数据块

► 例如:

```
int m , *p;
unsigned char array[100];
struct tagDATA a , B[10];
fread(&m,sizeof(int),1,fp); //读1个整型, 参数buffer是整型变量的地址
//读100个字节, 参数buffer是数组名 (地址)
fread(array,sizeof(unsigned char),100,fp);
//读结构体, 参数buffer是结构体对象的地址
fread(&a,sizeof(struct tagDATA),1,fp);
//读结构体数组, 参数buffer是数组名 (地址)
fread(B,sizeof(struct tagDATA),10,fp);
p=(int*)malloc(100*sizeof(int)); //动态分配能容纳100个整型的内存区
fread(p,sizeof(int),100,fp); //读100个整型, 参数buffer是指针值 (地址)
```

10.3.5 读写数据块

- ▶ `size_t`类型本质上是unsigned int型，函数返回实际读入的记录总数。由于读入错误或实际数据量的原因，返回值总是小于等于count。如果buffer参数实际内存长度不足以存储读入数据时，`fread`函数会导致崩溃性错误。

10.3.5 读写数据块

- ▶ fwrite函数将指定数目指定记录大小的数据块写入文件中，其函数原型为：

```
size_t fwrite(const void *buffer, size_t size, size_t count, FILE *stream);
```

- ▶ fwrite函数参数含义与fread相同，fwrite函数输出时不会进行任何编码转换、插入数据等操作，即直接将内存数据送到文件中。

10.3.5 读写数据块

- ▶ 使用fread和fwrite函数处理成批数据、数组、结构体等记录形式数据的二进制文件最方便，被广泛地应用于图形图像、音频视频、科学计算、格式文档文件应用中。
- ▶ 在处理成批数据二进制文件时，比fgetc和fputc函数有效率。但fread和fwrite函数不适合处理文本文件，因为它将空字符（\0）、换行符（\n）、回车符（\r）都一视同仁，读写数据不容易分辨出字符串信息。

10.3.5 读写数据块



【例10.6】

使用加大缓冲方式复制源文件为目的文件，支持命令行文件名输入。

10.3.5 读写数据块

例10.6

```
1 #include <stdio.h>
2 #include <string.h>
3 int main(int argc, char *argv[])
4 { //使用带参数的main函数版本获取命令行信息
5     char src[260],dest[260],buff[16384]; //读写缓冲达到16K
6     FILE *in,*out;
7     unsigned int rs;
8     if (argc<2) gets(src); //若无命令行参数输入源文件名
9     else strcpy(src,argv[1]); //否则第1个命令行参数为源文件名
10    if (argc<3) gets(dest); //若只有1个命令行参数输入目的文件名
11    else strcpy(dest,argv[2]); //否则第2个命令行参数为目的文件名
12    in=fopen(src,"rb"); //打开源文件读
13    if (in!=NULL) {
14        out=fopen(dest,"wb"); //创建目的文件写
15        while (!feof(in)) { //是否到源文件末尾
```

10.3.5 读写数据块

例10.6

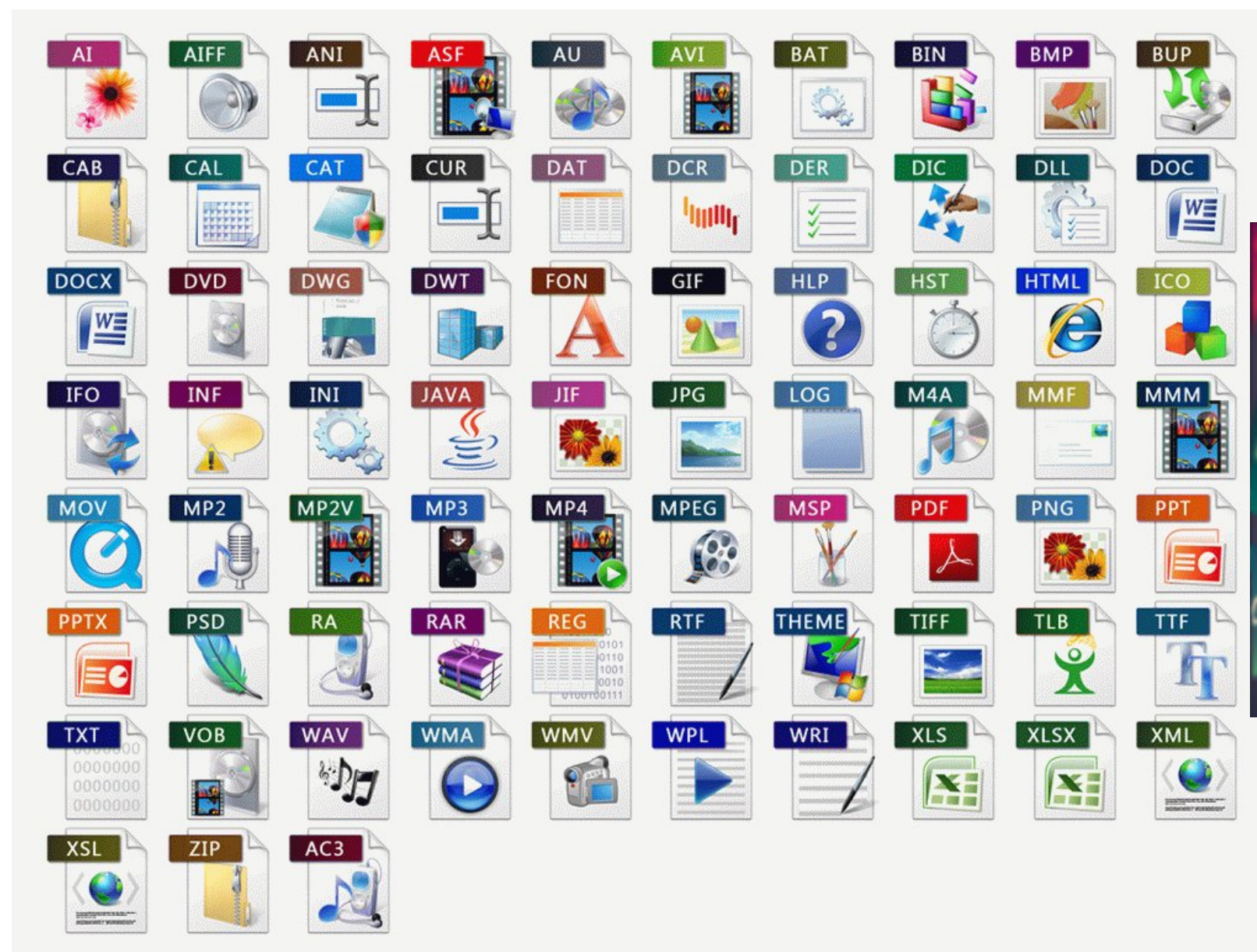
```
16      rs=fread(buff,sizeof(char),sizeof(buff),in);
17      fwrite(buff,sizeof(char),rs,out); //按实际读到字节数写入
18  }
19  fclose(out); //关闭目的文件
20  fclose(in); //关闭源文件
21  }
22  return 0;
23 }
```

10.5 编程处理各种格式的文件

▶ 1. 这世上，有成千上万种文件格式

- 压缩文件：
 - 7z,APK,ARC,ARJ,bzip2,cab,deb,DMG,GHO,gzip,jar,RAR,tar,Z,zip
- 图像设计文件：
 - 3dxml,ASM,DWF,DWG,GBR,AI,CDR,PSD,PDF,BMP,DIB,GIF,ICO,JPEG,MNG,PCX,EPS,SVG,3DS,C4D
- 文档文件：
 - DOC,DOCX,HTML,ABF,FNT,OTF,TTF,WOFF,XLS,XLSX,PPT,PPTX
-

10.5 编程处理各种格式的文件



10.5 编程处理各种格式的文件

▶ 2. 了解文件格式的网站

- 维基百科（支持中文）：
- https://en.wikipedia.org/wiki/List_of_file_formats
- [https://en.wikipedia.org/wiki/List_of_filename_extensions_\(alphabetical\)](https://en.wikipedia.org/wiki/List_of_filename_extensions_(alphabetical))
- Extension.info: <http://www.extension.info/>
- File extension database: <http://extensionfile.net/>
- File-Extensions.org: <http://www.file-extensions.org/>
- File.org: <http://file.org/>

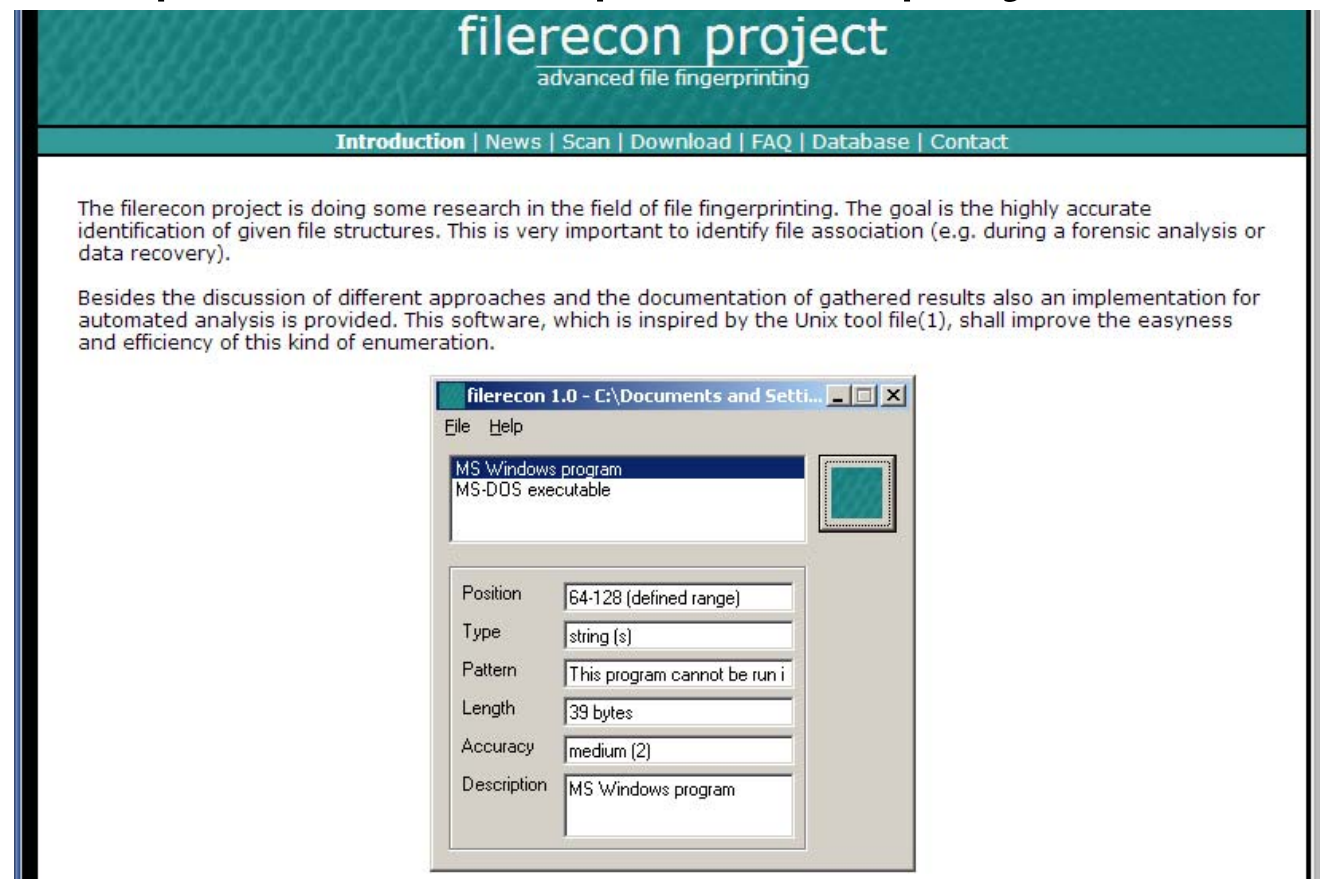
10.5 编程处理各种格式的文件

▶ 2. 了解文件格式的网站

- FileDesc（支持中文）：<http://www.filedesc.com>
- FileExtension.info：<http://fileextension.info/>
- Filefacts.net（支持中文）：<http://www.filefacts.net>
- FileInfo.com：<http://fileinfo.com/>
- Fmtz：<http://www.fmtz.com/>
- Uniblue File Extension：<http://www.file-extension.com/>
- Webopedia - Data Formats :
http://www.webopedia.com/Data/Data_Formats
- Wotsit.org：<http://devel.archefire.org/mirrors/www.wotsit.org/>

10.5 编程处理各种格式的文件

- ▶ 3. 这伙计，开源了一个检测文件格式的程序
 - <http://www.compute.ch/projekte/filerecon/>



10.5 编程处理各种格式的文件

- ▶ 4. 找到文件格式的说明，才能处理它
- ▶ ZIP文件
- MIDI文件
- RIFF文件

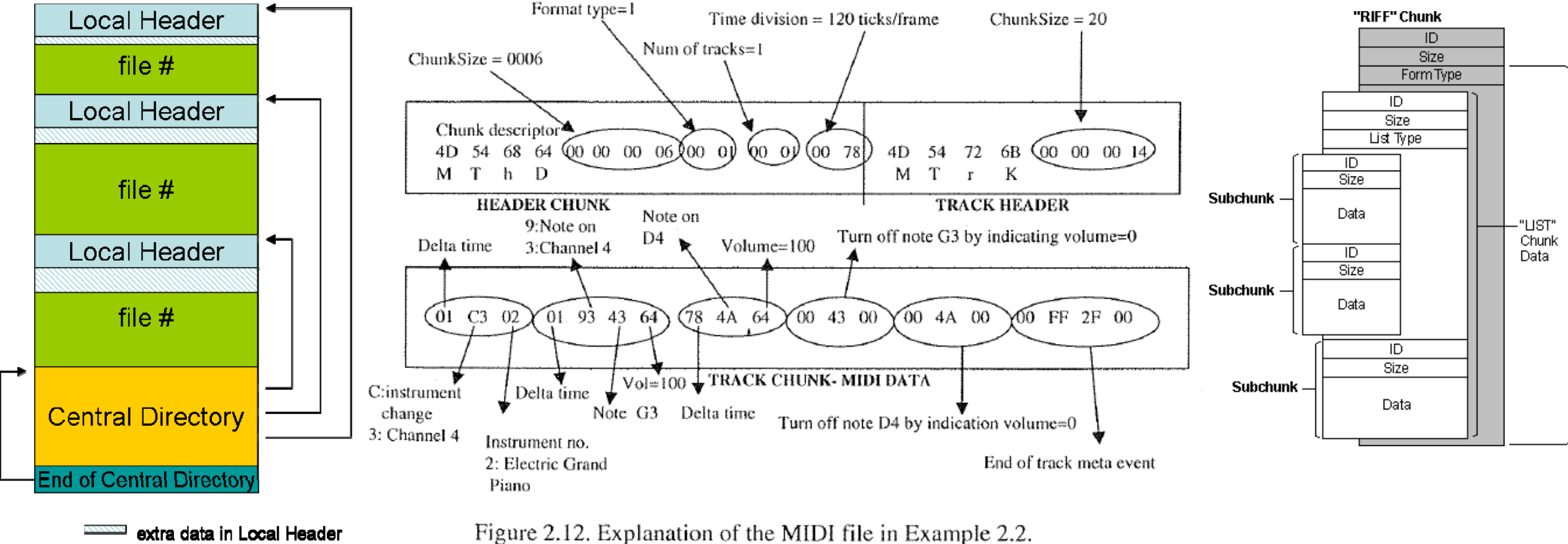


Figure 2.12. Explanation of the MIDI file in Example 2.2.

10.5 编程处理各种格式的文件



【例10.5】

显示24位位图文件信息。

10.5 编程处理各种格式的文件



例题分析

24位位图文件由文件头、位图信息头和图形数据三部分组成。文件头主要包含文件大小、文件类型、图像数据偏离文件头的长度等信息；位图信息头包含图像尺寸信息、图像像素字节数、是否压缩、图像所用颜色数等信息（<http://devel.archefire.org/mirrors/www.wotsit.org/>网站上有各种图形图像、音频视频、文档文件格式的说明）。根据位图文件格式定义文件头、位图信息头结构体类型，从位图文件读取结构体数据从而得到位图文件信息。

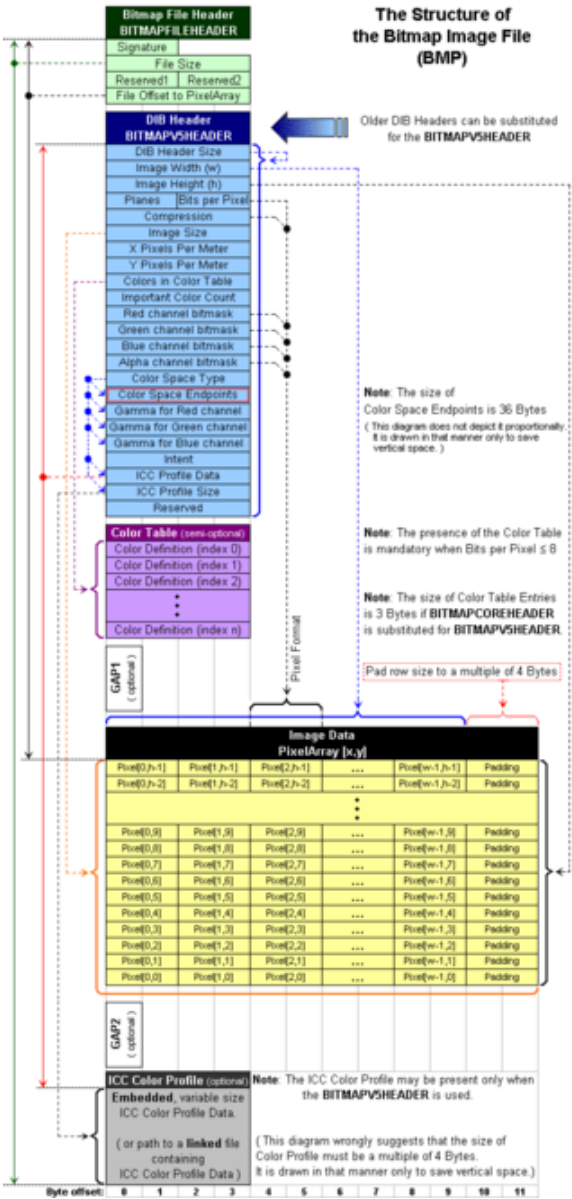
10.5 编程处理各种格式的文件

► 5. BMP位图文件格式

- https://en.wikipedia.org/wiki/BMP_file_format

Bitmap File Header BITMAPFILEHEADER	Bitmap File Header BITMAPFILEHEADER	Bitmap File Header BITMAPFILEHEADER	Bitmap File Header BITMAPFILEHEADER	Bitmap File Header BITMAPFILEHEADER
Signature	Signature	Signature	Signature	Signature
File Size	File Size	File Size	File Size	File Size
Reserved1	Reserved1	Reserved1	Reserved1	Reserved1
Reserved2	Reserved2	Reserved2	Reserved2	Reserved2
File Offset to PixelArray	File Offset to PixelArray	File Offset to PixelArray	File Offset to PixelArray	File Offset to PixelArray

DIB Header BITMAPV4HEADER	DIB Header (undocumented) BITMAPV3HEADER	DIB Header (undocumented) BITMAPV2HEADER	DIB Header BITMAPINFOHEADER	DIB Header BITMAPINFOHEADER
DIB Header Size	DIB Header Size	DIB Header Size	DIB Header Size	DIB Header Size
Image Width (w)	Image Width (w)	Image Width (w)	Image Width (w)	Image Width (w)
Image Height (h)	Image Height (h)	Image Height (h)	Image Height (h)	Image Height (h)
Planes	Planes	Planes	Planes	Planes
Bits per Pixel	Bits per Pixel	Bits per Pixel	Bits per Pixel	Bits per Pixel
Compression	Compression	Compression	Compression = BI_ALPHABITFIELDS	Compression = BI_BITFIELDS
Image Size	Image Size	Image Size	Image Size	Image Size
X Pixels Per Meter	X Pixels Per Meter	X Pixels Per Meter	X Pixels Per Meter	X Pixels Per Meter
Y Pixels Per Meter	Y Pixels Per Meter	Y Pixels Per Meter	Y Pixels Per Meter	Y Pixels Per Meter
Colors in Color Table	Colors in Color Table	Colors in Color Table	Colors in Color Table	Colors in Color Table
Important Color Count	Important Color Count	Important Color Count	Important Color Count	Important Color Count
Red channel bitmask	Red channel bitmask	Red channel bitmask	Red channel bitmask	Red channel bitmask
Green channel bitmask	Green channel bitmask	Green channel bitmask	Green channel bitmask	Green channel bitmask
Blue channel bitmask	Blue channel bitmask	Blue channel bitmask	Blue channel bitmask	Blue channel bitmask
Alpha channel bitmask	Alpha channel bitmask		Alpha channel bitmask	
Color Space Type				
Color Space Endpoints				
Gamma for Red channel				
Gamma for Green channel				
Gamma for Blue channel				



10.3.5 读写数据块

例10.5

```
1 #include <stdio.h>
2 #pragma pack(1) //结构体按1个字节对齐
3 typedef struct tagBITMAPFILEHEADER { //文件头
4     unsigned short bfType; //文件类型, 固定为BM
5     unsigned int    bfSize; //文件大小 (字节)
6     unsigned short bfReserved1; //保留
7     unsigned short bfReserved2; //保留
8     unsigned int    bfOffBits; //位图数据到文件头的偏移位置
9 } BITMAPFILEHEADER;
10 typedef struct tagBITMAPINFOHEADER { //位图信息头
11     unsigned int biSize; //位图信息头的长度
12     long biWidth; //位图的宽度, 以像素为单位
13     long biHeight; //位图的高度, 以像素为单位
14     unsigned short biPlanes; //位图的位面数
15     unsigned short biBitCount; //每个像素的位数
```

10.3.5 读写数据块

例10.5

```
16  unsigned int biCompression; //数据压缩类型
17  unsigned int biSizeImage; //位图数据大小（字节）
18  long biXPelsPerMeter; //水平分辨率（像素/米）
19  long biYPelsPerMeter; //垂直分辨率（像素/米）
20  unsigned int biClrUsed; //位图使用的颜色数
21  unsigned int biClrImportant; //调色板规范
22 } BITMAPINFOHEADER;
23 int main()
24 {
25     BITMAPFILEHEADER bmfh;
26     BITMAPINFOHEADER bmih;
27     FILE *bmp;
28     printf("%d,%d\n", sizeof(BITMAPFILEHEADER), sizeof(BITMAPINFOHEADER));
29     bmp=fopen("car.bmp", "rb");
```

10.3.5 读写数据块

例10.5

```
30  if (bmp!=NULL) {  
31      fread(&bmfh,sizeof(BITMAPFILEHEADER),1,bmp);  
32      fread(&bmih,sizeof(BITMAPINFOHEADER),1,bmp);  
33      printf("位图宽=%d,高=%d\n",bmih.biWidth,bmih.biHeight);  
34      fclose(bmp);  
35  }  
36  return 0;  
37 }
```

CP 程序设计