Identification of Different Patterns of MP3 and Duration Calculation

Yaqi Li[1]

Communication University of China
Beijing, China
E-mail: lyqegbupt@126.com

Yizhen Cao[2]

Communication University of China
Beijing, China
E-mail: caoyizhen@cuc.edu.cn

Abstract- With the development of technology and internet, mp3 has become a recognized music data format due to the advantages of its convenience. At present, the developers use some methods to analyze header and calculate the duration of mp3 with different header format. However some of the calculated durations are not accurate which will cause a series of serious software error. This paper compares the difference of CBR and VBR, ABR MP3 files and proposes a method to calculate the accurate duration of MP3 audio file.

Keywords- mp3; duration; VBR ; CBR ; ABR

I. INTRODUCTION

With the rapid development of Internet, the features of high density and needing smaller transmitting band width make MP3 occupy the header position of music area.MPEG standard provide that each version of MPEG, such as MPEG-1 or MPEG-2,must have three different Layers, which are named LayerI,LayerII and LayerIII. The LayerIII of MPEG-1 and MPEG-2 is called MP3.The one which is most common is MPEG-1 LayerIII. So nowadays, MP3 which is known by us generally refers to MPEG-1 LayerIII [1].

Bitrate is one of the most important factors of MP3 audio files. It is the number of bits that are processed per unit of time (bit per second). The higher this value, the quality of the music is better [2]. As the algorithm is different, there are three different kinds of MP3 compression, the CBR, VBR and ABR.

Firstly, VBR might cause timing difficulties for some decoders, the MP3 player might display incorrect timing information or none at all. Secondly, CBR is often required for broadcasting, which initially was an important purpose of the MP3 format [3].

II. COMPARISON OF CBR VBR AND ABR AUDIO FILE

CBR means constant bitrate; the number of bits of each second in mp3 audio file is the same. Though Music files have both complex and simple parts, the encoder use the same bitrate to encode the file .The result is that quality of complex parts is worse than the simple parts. The advantage of CBR file is that the size of files can be estimated. On the other hand, compared to VBR and ABR, size of CBR is bigger but the quality is not significantly improved.

VBR means variable bitrate; the number of bits of each second in mp3 audio always changes. According to different data, the encoder compress audio file with different bitrate on real time. This algorithm is provided by a company named XING [4]. They use high bitrate to compress the complex parts and low bitrate to compress the simple parts. This idea is intelligent, but the algorithm of VBR used by encoder of XING is unsatisfactory. People can't endure its low efficiency. Fortunately, a company named Lame perfectly optimizes the VBR algorithm and makes it the best pattern of encoding for MP3. The VBR, which gives attention to both quality and size of file, is a recommended encoding pattern. The advantage of this pattern is when the VBR file is taking a little space. It also can guarantee the quality of the audio. The disadvantage is the size of file after compressed can't be estimated before encoding. Complexity of encoding increase and it need more time to encode or decode the VBR file.

ABR means average bitrate, which is a kind of interpolation of VBR parameters. Aiming at big size of CBR file and uncertain size of VBR file, Lame creates this new pattern. ABR is also called "Safe VBR". Within the specified average Bitrate, it uses relatively low bitrate in parts of low frequency or frequency which is not sensible and relatively high bitrate in parts of high frequency or dynamic performance per 50 frames. For example, when encoding a WAVE file with specified 192kbps, 85 percent of the file will be encoded with 192kbps, the other part will be encoded in variable bitrate. Complex parts will be encoded with higher than 192kbps and the simple parts will be encoded with lower than 192kbps. Compared with 192kbps CBR file, their size are almost the same and difference of size between them is usually from 2 percent to 3 percent, but the quality of ABR audio file is much more better. Compared with VBR, although the speed of encoding is 2 times of VBR, the quality of audio is worse.

III. CALCULAITION DURATION OF MP3

A. MPEG Frame Header

MPEG audio file itself doesn't have file header. It consists of many data blocks, which is called data frame. Frame is the basic unit in audio technology. Each frame contains a header at its beginning followed by the audio data. This audio data always contains a fixed number of samples. Frame header is some data of a certain length to

describe the parameters of the audio data. The Audio decoders analyze the frame header for decoding the audio file. Frame = Frame Header + Audio Data, The header which is at the beginning of each frame is 32 bits long and has the following format.

TABLE1   MPEG AUDIO FRAME HEADER[5]

Position	Length	Meaning	Example
0	11	Frame sync to find the header (all bits are always set)	1111 1111 111
11	2	Audio version ID 00 MPEG Version 2.5 01 reserved 10 Version MPEG 2 11 Version MPEG 1	00 MPEG Version 2.5 01 reserved 10 Version MPEG 2 11 Version MPEG 1
13	2	Layer index 00 reserved 01 Layer III 10 Layer II 11 Layer I	01
15	1	Protection bit 0 - protected by 16 bit CRC following header 1 - no CRC	0
16	4	Bitrate index (please search the table2)	1001
20	2	Sampling rate index (please search the table3)	11
22	1	Padding bit	0
23	1	Private bit (only informative)	1
24	2	Channel mode 00 Stereo 01 Joint Stereo 10 Dual channel (Two mono channels) 11 Single Channel (Mono)	00
26	2	Mode extension (Only used in Joint Stereo)	00
28	1	Copyright bit (only informative)	1
29	1	Original bit (only informative)	1
30	2	Emphasis 00 none 01 50/15 ms 10 reserved 11 CCIT J.17	00

TABLE2   BITRATES

Bitrate Index	MPEG 1			MPEG 2, 2.5 (LSF)	
	Layer I	Layer II	Layer III	Layer I	Layer II & III
0000	free				
0001	32	32	32	32	8
0010	64	48	40	48	16
0011	96	56	48	56	24
0100	128	64	56	64	32
0101	160	80	64	80	40
0110	192	96	80	96	48
0111	224	112	96	112	56
1000	256	128	112	128	64
1001	288	160	128	144	80
1010	320	192	160	160	96
1011	352	224	192	176	112
1100	384	256	224	192	128
1101	416	320	256	224	144
1110	448	384	320	256	160
1111	reserved				

TABLE3   MPEG VERSIONS AND SAMPLING RATES

Sampling Rate Index	MPEG 1	MPEG 2 (LSF)	MPEG 2.5 (LSF)
00	44100 Hz	22050 Hz	11025 Hz
01	48000 Hz	24000 Hz	12000 Hz
10	32000 Hz	16000 Hz	8000 Hz
11	reserved		

We can conclude that two Bytes at the beginning of the frame of MP3 have to be "FFFA"or"FFFB".Before we calculate the duration, we need to make sure whether the file is CBR or VBR.ABR is one pattern of VBR. VBR file have 2 kinds of appended Header: XING and VBRI. Most VBR files add one of these headers. This header stands after the first MPEG audio header at a specific position (0x24). The whole first frame which contains the XING header is a valid but empty audio frame, so even decoders which don't consider this header can decode the file.

TABLE4 XING HEADER

Position	Length	Meaning	Example
0	4	VBR header ID in 4 ASCII chars, either 'Xing' or 'Info', not NULL-terminated	"Xing"
4	4	Flags which indicate what fields are present; flags are combined with a logical OR. Field is mandatory. 0x00000001 - Frames field is present 0x00000002 - Bytes field is present 0x00000004 - TOC field is present 0x00000008 - Quality indicator field is present	0x0007 (means Frames, Bytes & TOC valid)
8	4	Number of Frames as Big-Endian DWORD (optional)	7344
8 or 12	4	Number of Bytes in file as Big-Endian DWORD (optional)	45000
8, 12 or 16	100	100 TOC entries for seeking as integral BYTE (optional)	
8, 12, 16, 108, 112 or 116	4	Quality indicator as Big-Endian DWORD from 0 - best quality to 100 - worst quality (optional)	0

XING header position = MPEG first Frame header position + 4 (Bytes) + padding

MPEG first Frame header position: if there is no ID3 information, this value is 0. On the contrary, we need use frame sync to find the header.

Padding: the padding size of MP3 which we always use is 32 Bytes.

ID3 is a metadata container most often used in conjunction with the MP3 audio file format. It allows information such as the title, artist, album, track number, and other information about the file to be stored in the file itself. But what need to pay attention to is that these information maybe include the sync '11111111111', which may affect us on locating the first frame header.SO we need to rule out the interference of this part of the information.ID3 is divided into two versions, ID3v1 and ID3v2. ID3V1 is 128 bytes at end of mp3 file with 'Tag' in the first three characters, followed by music information. So it doesn't affect us on location the first frame header.ID3v2 generally exists at the beginning of the file, which can store large capacity of information such as album, lyrics, ID3v2 composed of header and some frames. The header structure is shown as below [6].

Char Header [3] /*must be 'ID3'*/

Char Ver /*version*/

Char Revision/*vice version*/

Char Flag /*flag of what be stored*/

Char Size [4] /*size of tag. （do not include the 10 bytes of header）*/

So the size of whole tag can be achieved by reading the header of tag.

We can use the size as the offset to search the sync for finding first data frame header.

VBR also has another header named VBRI, which is only used by Fraunhofer Encoder. VBRI Header is located exactly 32 bytes after the end of the first MPEG audio header in the file.

TABLE 5 VBRI HEADERS

Position	Length	Meaning	Example
0	4	VBR header ID in 4 ASCII chars, always 'VBRI', not NULL-terminated	'VBRI'
4	2	Version ID as Big-Endian WORD	1
6	2	Delay as Big-Endian float	7344
8	2	Quality indicator	75
10	4	Number of Bytes as Big-Endian DWORD	45000
14	4	Number of Frames as Big-Endian DWORD	7344
18	2	Number of entries within TOC table as Big-Endian WORD	100
20	2	Scale factor of TOC table entries as Big-Endian DWORD	1
22	2	Size per table entry in bytes (max 4) as Big-Endian WORD	2
24	2	Frames per table entry as Big-Endian WORD	845
26		TOC entries for seeking as Big-Endian integral. From size per table entry and number of entries, you can calculate the length of this field.	

So, we can check file whether CBR or VBR by the steps.
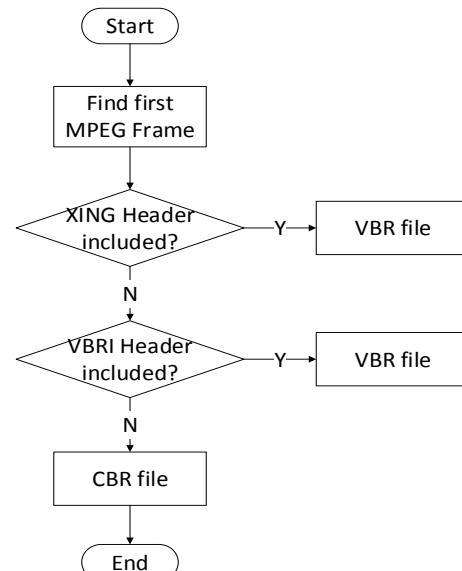
Pseudo-code is as follows:
While(offset <= file size)
{

```
    Char sync = read 1 byte from offset;
    If(sync&0xFE == 0xFE)
    {
        First_Frame_header_position = offset;
        Break;
    }
    Else
    {
        ++offset;
    }
}
If(offset == file size)
    Output("this file maybe broken");
Else
{
    int  Xing_VBRI_header_position =
First_Frame_header_position+36;
    Char info[4]= read 4 bytes from
Xing_VBRI_header_position;

    If(info equals "Xing")
        This is a VBR MP3 with Xing header;
    Else if (info equals "VBRI")
        This is a VBR MP3 with VBRI header;
    Else
        This is a CBR MP3;
}
```

B. Calculate the durition of mp3

No matter VBR audio file or CBR audio file, the sample frequency is same in each frame. We know that sample number of MPEG1-LayerIII is 1152. With constant sample frequency (ex.44100hz), duration of each frame is constant. Time_Per_Frame = Sample_Number/Sample_Frequency.

1) Duration of CBR file

Strictly speaking, size of audio file equals to file whole size minus Tag size. But compared with whole file, size of Tag is so small that it can be generally ignored. So we can use Whole size of file to calculate directly. AS CBR file use the constant Bitrate, so we can analyze first frame and achieve the bitrate index and sample frequency index. Then we can get the value of them through checking the table.

Frame_Size = Frame_Header_Size + Data_Size
            =4(Bytes) + Data_Size
            =4*8+ Time_Per_Frame*Bitrate*1000(bit);
Total_Frame_Number = File_Size/Frame_Size;
File_Duration = Time_Per_Frame * Total_Frame_Number;

We can also use another method to calculate the duration directly. As we know the size and bitrate of the CBR file, we can easily calculate the duration by the formula.

CBR Duration = File_Size*8/(Bitrate*1000);(unit of Bitrate is Kbps)

But Bitrate is only relative to audio data, it doesn't conclude Frame header (4 Bytes).Through the above formula, we can't get exact duration of MP3.The error Range is about 1%.

2) Duration of VBR file

We can adopt method named 'average bitrate' to calculate duration of VBR file. Due to the bitrate of each frame is not a fixed value, so we need to scan the whole file for getting the summation of bitrate from every frame and the total number of frames. Then we use this number divided by summation of bitrate that we can get the average bitrate. Now we can use the method to calculate the duration of VBR as follow.

File_Duration = File_Size/Average_Bitrate

But this method is inefficient. We have to scan the whole file and analyze header of each frame. It is too complex. So the following method is much better.

As we have said before, we know that most of VBR files conclude XING Header or VBRI Header. We can get the number of total frames, Bitrate index and sample frequency index through locating Xing Header or VBRI Header and analyze them. After checking relative table, we can directly get the bitrate and sample frequency. So here is the formula which can help us calculate the duration of VBR MP3 efficiently.

File_Duration = Time_Per_Frame * Total_Frame_Number

IV.   CONCLUSION

No matter CBR or VBR file, this sampling rate is fixed so the duration of each frame is also fixed. We can locate the first frame of data by the flag "11111111111". Analyzing the first frame and checking file whether CBR or VBR is the second step. The value of sampling rate, bitrate and the total number of frames were achieved at last so that we can calculate the exact duration of Mp3 audio file.

REFERENCES

[1] Cervera,Teresa,The effect of MPEG audio compression on multidimensional set of voice paramet ers, Logopedics Phoniatrics Vocology, 2001. 8, Vol. 26,

[2] Alberto Corrales-Garc á,Jos é Luis Mart ñez,Gerardo Fern ández-Escribano,Francisco Jos é Quiles. Variable and constant bitrate in a DVC to H.264/AVC transcoder[J]. Signal Processing: Image Communication,2011,266:.

[3] Raissi R. The theory behind MP3[J]. MP3'Tech, 2002.

[4] Gonzalez J,Cervera T. The effect of MPEG audio compression on multidimensional set of voice parameters.[J]. Logopedics Phoniatrics Vocology,2002,263:.

[5] ISO/IEC 11172-3:1993

[6] ID3v2 Chapter Frame Addendum. Available at: http://id3.org/id3v2-chapters-1.0