

文章编号 :1001-2486(2001)02-0094-05

MPEG 音频层 III 解码算法的仿真及定点化*

陈洪光 林嘉宇 唐朝京

(国防科技大学电子科学与工程学院,湖南长沙 410073)

摘要 :应用 MPEG 音频层 III 压缩算法的 MP3 音乐格式,以其较高的压缩比及近乎 CD 的音质得到广泛流行。论文在完成 MPEG 音频层 III 解码算法 C++ 语言仿真的基础上,对算法进行了定点化,以实现基于定点 DSP 的音频层 III 实时解码。

关键词 :MPEG audio Layer III ;解码 ;定点化 ;MP3

中图分类号 :TN911.21 **文献标识码** :A

Simulating and Fixed Point Programming of MPEG Audio Layer III Decoding Algorithm

CHEN Hong-guang, LIN Jia-yu, TANG Chao-jing

(College of Electronic Science and Engineering, National Univ. of Defense Technology, Changsha 410073, China)

Abstract :MP3 music format which uses MPEG audio Layer III compressing algorithm is popular for its high compression rate and CD-like quality. The fixed point programming of the MPEG audio Layer III decoding algorithm is implemented after being simulated with C++ language, which can be used to realize the real time decoding of MPEG audio Layer III based on digital signal processor.

Key words :MPEG audio Layer III ;decoding ;fixed point programming ;MP3

MPEG(Motion Picture Expert Group)音频压缩标准 ISO/IEC 11172-3 总共规定了三个层次的子带音频编码算法 :Layer I、Layer II 和 Layer III。它们各具不同的性能和编码复杂度,对应于不同的应用要求。MPEG 音频层 III 就是其中的 Layer III,它的算法最为复杂而压缩性能比也最好(压缩比高达 12 倍),目标编码速率为 64 kbit/s,特别适用于音频存储和 ISDN 网络传输。我们在 PC 机上对 Layer III 解码算法做了软件仿真和定点化,并对仿真以及定点化的结果进行了分析测试。

1 音频 Layer III 解码算法的基本原理及其仿真实现

MPEG 音频压缩算法运用了谱分析技术和频率相关比特分配技术,使量化噪声与听觉特性相匹配,是一种声音听觉编码。MPEG 音频 Layer III 的码流数据以帧为单位封装,一帧包含 1152 个音频样值(48k 采样时,相当于 24 ms)帧格式如图 1 所示。

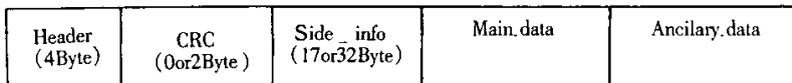


图 1 MPEG 音频 Layer III 帧结构

Fig.1 Frame rtructure of MPEG audio layer III

帧头(Header)包含同步码和状态信息。错误检查(CRC)部分包含错误检测信息,这部分依据帧头信息中保护位(Protection_bit)的状态可选。边信息(Side_info)中包含解码所需的辅助信息,其长度依声道数不定,单声道为 17 字节,双声道为 32 字节。主数据(Main_data)是真正的子带样值编码数据,它的长度可变。辅助数据(Ancillary_data)由用户自己定义。

MPEG 音频 Layer III 解码器结构框图如图 2 所示,解码以帧为单位进行,其过程简述如下:首先对输

* 收稿日期 2000-09-03

作者简介 陈洪光(1976-)男,硕士生。

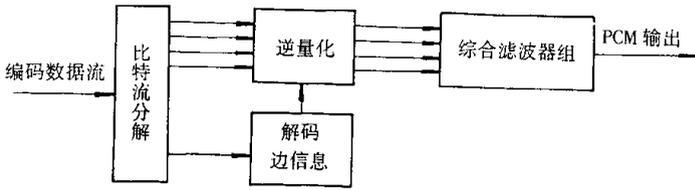


图 2 MPEG 音频 Layer III 解码结构框图

Fig.2 Block Structure of MPEG Audio Layer III Decoder

入解码器的音频数据按照帧格式进行比特流分解,恢复出各种信息,并对量化样值进行霍夫曼解码。然后根据码流分解得到的边信息、总增益、窗口类型等信息,对哈夫曼解码后的频谱量化值作逆量化处理。处理后的数据经综合滤波器组恢复出 PCM 音频信号。

在以上的解码过程中,比特流分解以及解码边信息模块是根据音频 Layer III 的帧格式进行的,它们在仿真实现中大多是文件操作以及查表操作,没有涉及复杂运算,比较容易实现。解码的运算量主要集中在后面的哈夫曼解码、逆量化频谱、IMDCT 变换以及合成子带滤波器。哈夫曼解码是一个查表过程,ISO/IEC 11172-3 标准提供了音频编解码所需的哈夫曼码表。详细码表请参见^{[1][2]}。逆量化模块的运算稍为复杂一些,有非整数指数的幂运算。在 C 语言的浮点算法实现中可直接作幂运算,而定点算法中要将其转换成查表运算。逆量化频谱过程可由下面的公式完整描述:

$$xr_i = \text{sign}(is_i) * |is_i|^{4/3} * 2^{k_1} * 2^{k_2} \quad (1)$$

公式中 is_i 、 xr_i 分别是逆量化频谱的输入输出, $\text{sign}()$ 为求符号运算,系数 k_1, k_2 与所处理的块长有关。

最后的综合滤波器组模块包含 IMDCT 变换和矩阵变换。IMDCT 变换是规则的标准运算,计算公式如下:

$$X_i = \sum_{k=0}^{n/2-1} X_k \cos\left(\frac{\pi}{2n}\left(2i+1+\frac{n}{2}\right)(2k+1)\right) \quad i = 0, 1, \dots, n-1 \quad (2)$$

IMDCT 变换可采用快速算法提高运算效率。矩阵变换也是一些常规的乘累加运算,DSP 指令中有专门的乘累加器可以完成。

一帧包含 $36 * 32 = 1152$ 个子带样值,经解码后即可得到 1152 个音频 PCM 样点。将解码程序输出的 PCM 样点数据封装成 WAV 格式文件,即可通过声卡播放。我们对解码仿真结果进行了主观测试,八位听音人一致认为,仿真程序解码后的音质与 Winamp 等流行 MP3 软播放器感觉不出差异。

2 Layer III 解码算法的定点化

我们对音频 Layer III 的解码运算量进行了分析估计,得出双声道的解码约需 15 MIPS 运算量。ADSP2181 的运算速度为 33 MIPS > 15 MIPS,故采用一片 ADSP2181 即可完成 Layer III 的实时解码。ADSP2181 是 AD 公司的 16 位定点 DSP 芯片,其处理速度快且成本低。为能在定点 ADSP2181 上完成 MPEG 音频 Layer III 的实时解码,需先将 MPEG 音频 Layer III 的解码算法做定点化处理。

定点解码仿真可直接着手进行,也可在完成了浮点解码算法的基础上对解码程序进行改造。我们采用后者,首先用

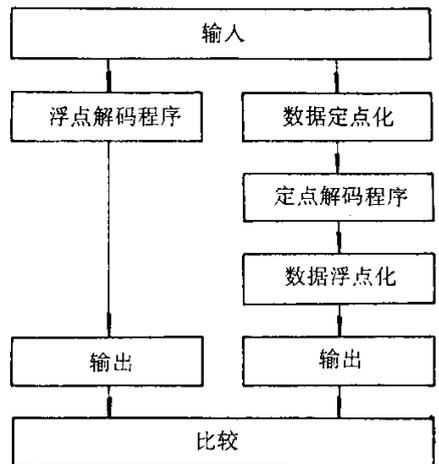


图 3 定点化工作流程

Fig.3 Fixed point programming flow

C语言实现 Layer III 的浮点运算解码,并分析浮点算法中各参数的动态范围以确定定点的运算精度,然后对浮点算法进行定点化的改造,最后对定点算法进行性能分析,将其与浮点算法作比较。若人耳分辨不出两者解码音质的差别,并且信噪比较低,则可认为定点化比较成功。整个工作过程见图 3。

定点算法处理信号的动态范围小,且乘法运算会引入舍入或截尾误差,因此定点化改造会损伤信号的精度。在定点处理字长固定的情况下,信号处理的精度与动态范围成反比。在实际处理时,精度往往要受信号动态范围的控制。因而在不太影响运算速度的条件下,如何尽量不损伤精度又兼顾动态范围,成为定点化的关键。

表 1 一些关键变量的动态范围和量化精度

Tab.1 Dynamic ranges and quantizing precisions of some key variables

变量名	动态范围	处理精度	变量意义
is	[- 511, 511]	Q_0	哈夫曼解码的输出值,是整数
$ is_i^{4/3} $	{0} [1, 4085.3]	Q_3	
gain	(0, 1)	Q_{15}	全局增益
scfcr	[2^{-18} , 1]	Q_{15}	缩放因子
xr	[0, 151.81]	Q_7	逆量化后的频谱值
cs ca	[0, 8575, 1] [0.00369, 0.5145]	Q_{15}	混迭减除处理的系数
xar	[0, 229.92]		混迭减除处理后的输出
x	[0, 1]	Q_{15}	IMDCT 的输出
N	[0, 1]	Q_{15}	合成滤波器的系数
S	[0, 32767]	Q_0	合成滤波器输出

在定点化之前,我们根据原理和统计两方面估计了浮点程序中所有全局参量的动态范围,从而确定了这些参数的处理精度。表 1 列出一些关键变量的动态范围和量化精度。

整个定点化过程中逆量化频谱模块的定点化较为复杂些,其中包含有非整指数的幂运算,这些运算要转换成查表运算,并且这当中有些变量值的动态范围超过了 16 位定点所能表达范围,如 $|is_i^{4/3}|$ 的动态范围为 [1, 165112.49] {0},有必要对其动态范围进行压缩,以保证处理精度。

逆量化频谱公式(1)中 $|is_i^{4/3}|$ 的计算可转换成查表运算。哈夫曼解码输出的频谱量化值 is_i 是整数,无需作精度处理,定为 Q_0 。理论上, is_i 的动态范围是 [- 8191, 8191],经大量测试,我们发现 99.9999875% 以上的 is_i 绝对值小于 511,因此,在实际中我们取其动态范围为 [- 511, 511],这样可提高 $|is_i^{4/3}|$ 的处理精度, $|is_i^{4/3}|$ 的动态范围是 {0} [1, 4085.3],精度取为 Q_3 。另外,式(1)中的增益参数 $gain = 2^{k/4}$ 和缩放因子 $scfcr = 2^{k/2}$ 也要转换成查表运算。 $gain$ 的理论动态范围是 (0, 2896.3)。大量的实际测量表明,99.999999% 以上的值在 (0, 1) 内,因此,在实际定点化改造中, $gain$ 的动态范围取为 (0, 1),这样就可大大提高该变量的处理精度: Q_{15} 。 $scfcr$ 的动态范围是 [2^{-18} , 1],量化精度定为 Q_{15} 。逆量化后的频谱值 xr 的动态范围为 [0, 151.81],精度定为 Q_7 。

主要变量的精度确定之后,就开始对运算过程进行分析,判断那些地方的会出现 16 位以上变量的运算。对这些运算有两种处理方法:可以将运算进行高低字节分开处理,以转换成 16 位的运算,最终的结果要视其动态范围重新定点化;也可以根据这些变量的动态范围,将其定点化成 16 位以内的精度之后,再作运算。

对 DSP 指令集中存在的运算可直接进行定点改造,如加、减、乘、除等运算,而对 DSP 指令集中不存在的运算(如开方、对数等)要转换成查表处理,或者将这些运算进行泰勒级数展开处理,以便作定点运算。

作者在进行定点化改造时,是分模块单独进行的,完成一个模块的定点改造后,将其嵌入原浮点程序中,以替代相应的浮点模块。这样运行程序,通过与浮点程序解码结果对比,就可以检查该模块改造的效果。若效果不如浮点解码程序或有明显的差别,则可确定该模块的定点化存在问题,而不必涉及其它

模块,这样做一定要注意已改造的定点模块与未改造的浮点模块之间接口参数的一致性,为此,要把定点模块的输出和输入参数浮点化。

3 定点化性能分析

定点化完成后,我们从主客观两方面对定点程序的解码性能进行了分析测试。主观听音测试中,听音人感觉不出定点程序与浮点程序解码结果的差异。我们做了一个有趣的试验,首先让听音人分别听一遍浮点解码音质和定点解码音质,并预先告知他们所听的是浮点还是定点的解码音质;然后,按随机次序再播放一遍两种程序的解码音质,这次不告诉听音人所播放的是哪一种解码音质,之后让他们分辨前后播放的分别是何种解码音质。测试结果表明,几乎没有人能分辨出两种音质。

客观分析中,由于 layer III 基本上属于波形编码,我们可以计算定点程序解出的音频信号相对浮点程序的信噪比,计算式如下:

$$SNR = 10 \log \frac{\sum S_{float}^2}{\sum (S_{fix} - S_{float})^2} \quad (3)$$

式中, S_{float} 、 S_{fix} 分别是浮点程序和定点程序解码恢复的 PCM 样值。我们有代表性地选取了八种音乐材料,它们分别是男女声摇滚乐、男女声流行乐、小提琴协奏曲(弦乐)、小号独奏曲(管乐)、鼓音较强的进行曲(打击乐)、交响曲。计算结果列于表 2 中,它们信噪比的平均为 32.2 dB。

表 2 八种音乐材料的信噪比

Tab.2 SNRs of 8 Types of Music

男声 摇滚乐	女声 摇滚乐	男声 流行乐	女声 流行乐	小提琴 协奏曲	小号 独奏曲	进行曲	交响曲	平均值
32.4	32.8	33.3	32.1	33.6	31.9	30.4	31.3	32.2

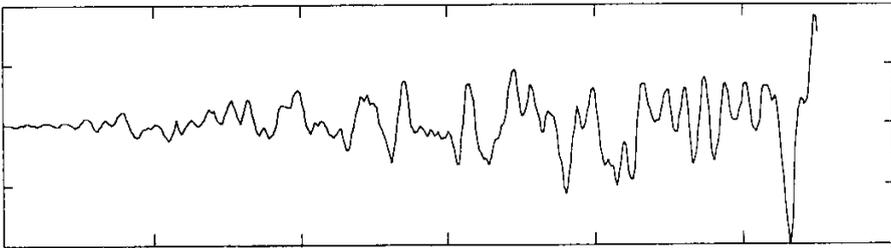


图 4(a) 定点程序解码 0.5 帧数据得到的 PCM 信号波形

Fig.4(a) PCM Signal waveform derived from fixed point programme

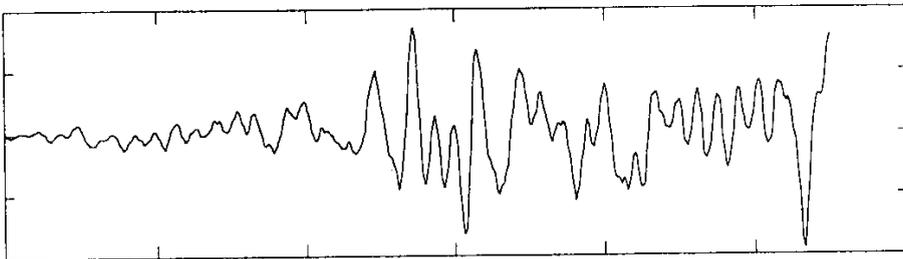


图 4(b) 浮点程序解码 0.5 帧数据得到的 PCM 信号波形

Fig.4(b) PCM Signal waveform derive from floating point programme

上面的波形图描绘了浮点程序与定点程序解码所得到的 PCM 样值以及两者比较的误差。其中图 4

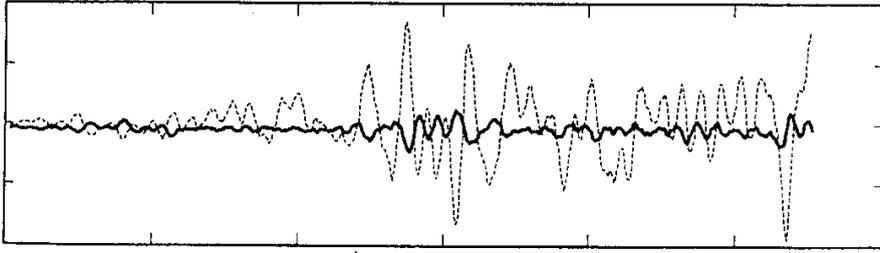


图 4(c) 定点解码与浮点解码数据的差值

Fig.4(c) Decoding difference of fixed point and floating point programme

(a)(b)分别是定点和浮点程序解码 0.5 帧(12 MS)MP3 音乐的结果,图 4(c)中的实线为误差曲线,虚线是浮点程序解码结果。从图中的误差曲线可以看出,定点程序解码的 PCM 信号波形与浮点程序解码波形相比失真较小,基本一致。仔细比较,可以发现定点程序解码波形的抖动不如浮点程序明显,即高频分量不如浮点程序解码波形丰富。这说明定点化处理相当于一个低通滤波器,抑制了高频分量。横向比较看,幅度较大的 PCM 信号的失真要比幅度较小 PCM 信号的失真大些。这是由于量化精度受动态范围的控制,动态范围大则精度要低些,反之亦然。在定点化处理中,我们为保证信号的整体处理精度(照顾小信号),信号的动态范围被人为地压缩了(如表 1 中的变量 is 和 $gain$),也就是说定点处理的动态范围比实际的动态范围(即浮点动态范围)要小,致使一些超过定点处理动态范围的较大信号被强制限幅,定点处理相当于限幅器。

参考文献:

- [1] ISO/IEC 11172-3 Coding of Moving Picture and Associated Audio for Digital Storage Media at up to About 1.5Mbit/s Part 3 :Audio[S],1993.
- [2] 杨品等译. MPEG 运动图象压缩编码标准[M]. 北京:机械工业出版社,1995.
- [3] Steinmetz R, Nahrstedt K. Multimedia :Computing ,Communications and Applications[M]. 北京:清华大学出版社,1997.
- [4] 苏涛等. 高性能数字信号处理器与高速实时信号处理[M]. 西安:西安电子科技大学出版社,1999.
- [5] A. V. 奥本海姆. 数字信号处理[M]. 北京:科学出版社,1981.

