

电子防震系统音频编解码器及编解码方法研究

王景军

(杭州士兰微电子股份有限公司, 浙江 杭州 310012)

摘要: 提出一种高效的编解码实现方案。在时频分析模块中, 利用 MDCT 的对称性和奇偶性, 快速得到 MDCT 变换输出值, 节省运算时间和计算资源提高编码效率。在量化/熵编码模块中, 采用非线性量化和 Huffman 编码来压缩数据, 使量化的循环迭代次数减少, 在保证量化误差最小化的前提下, 提高编码速度。

关键词: 电子防震; 时频分析; 量化编码; MDCT

中图分类号: TP391

文献标识码: A

Electronic shock protection system audio encode/decode technology

WANG Jing Jun

(Hangzhou Silan Microelectronics CO., LTD., Hangzhou 210012, China)

Abstract: This paper proposes an encode/decode algorithm which works effectively. In time frequency analysis module, the N-point MDCT transform is employed to improve the computation rate and the data throughput. Finally, to meet the bit rate/distortion critical, a novel technique through pre-processing and adjustment of scale factor in quantization block is used.

Key words: ESP; time frequency analysis; quantization; MDCT

电子防震 ESP(Electronic Shock Protection) 系统在便携式或者车载伺服播放器等设备中应用非常广泛, 防止这些设备由于遭受机械震动的影响而导致音频数据不能正常连续播放, 电子防震时间是其中最重要的指标。电子防震系统的防震时间由 DRAM 容量和音频编码算法的压缩率决定, 电子防震系统对音频编解码算法有 3 个方面的要求: (1) 音质基本无损失; (2) 压缩比越高, 相同容量的 DRAM 支持的防震时间越长, 或者相同防震时间下所需的 DRAM 更小, 成本更低, 一般要求达到 4:1 以上的压缩率; (3) 由于在峰值处理阶段, ESP 系统需要并行完成 2 次编码运算和 1 次解码运算, 如果采用软件实现, 要求其编解码算法的运算复杂度较低。

传统的电子防震系统一般采用自适应差分脉冲编码调制(ADPCM)进行音频编解码。但是 ADPCM 编码器有 2 个显著的缺点: (1) ADPCM 编解码算法采用逐点计算的结构, 运算量大, 只能使用专用集成电路(ASIC)实现, 不适合软件实时实现, 采用这种方案就需要在系统中添加专用 ESP 芯片, 导致成本增加; (2) ADPCM 编码器的压缩效率较低, 高于 4:1 的压缩率会导致音质明显恶化。

目前还有一种技术提出了采用 MPEG1 音频编解码算法作为 ESP 的编解码算法, 能够在音质基本没有损失的前提下提高压缩比。但由于该方法采用心理声学模型并且使用混合型的滤波器组进行时频分析和逆时频分析, 运算复杂度高, 硬件成本开销大。

为了解决上述问题, 本文提出了一种新的高效编解码方法, 它具有计算复杂度低、编码速度快和适合软件实现的优点, 能够有效地延长电子防震时间。在时频分析模块中, 利用改进型离散余弦变换 MDCT(Modified Discrete Cosine Transform)的对称性和奇偶性, 快速得到 MDCT 变换输出值, 节省运算时间和计算资源, 提高编码效率。在量化/熵编码模块中, 采用非线性量化和 Huffman 编码来压缩数据, 使量化的循环迭代次数减少, 在保证量化误差最小化的前提下, 提高编码速度。

1 系统及编解码器架构

常用的 ESP 系统架构如图 1 所示, 包含了 ESP 编码器, DRAM 缓冲器和 ESP 解码器。其工作机制是, 当开启 ESP 系统时, 音频样点依次通过 ESP 编码器进行编码, 编码后的码流存储在 DRAM 缓冲器中, DRAM 缓冲区中

技术与方法 Technique and Method

的码流经 ESP 解码器解码输出,从而实现一定时间的电子防震效果。



图1 ESP系统

现行的 ESP 编码器一般都采用心理声学模型,并且使用混合型的滤波器组进行时频分析和逆时频分析,导致 ESP 编码器运算效率降低。本文提出的 ESP 编码器舍弃了心理声学模型,直接采用时频分析模块和量化/熵编码模块(或增加封装模块)架构,其结构框图如图2所示。音频样点输入时频分析模块后转换为频率系数,频率系数经过量化/熵编码模块压缩成比特流,比特流输出到 DRAM 缓冲器。

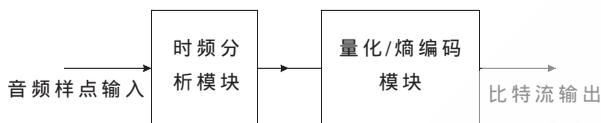


图2 ESP编码器

ESP 编码器架构还可以进一步扩展,增加封装模块,其架构如图3所示:



图3 扩展的ESP编码器

与 ESP 编码器结构类似,ESP 解码器的结构框图如图4所示,从 DRAM 缓冲器中读取1个比特块长度的比特流,然后将比特块长度的比特流送入熵解码/反量化模块的输入缓冲区,进行熵解码和反量化运算后得到重建的频率系数送入逆时频分析模块,进行逆改进型离散余弦变换 IMDCT(Inverse Modified Discrete Cosine Transform)变换,重建时域信号并输出播放。

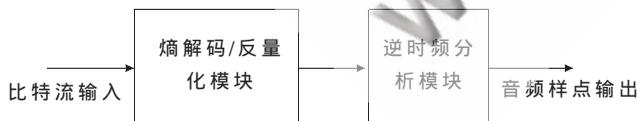


图4 ESP解码器

2 编解码算法

2.1 编码算法原理

ESP 系统中编码步骤包括:

- (1) 音频样点通过时频分析转换为频率系数;
- (2) 频率系数经过量化/熵编码压缩成比特流。

2.1.1 时频分析

时频分析模块采用 MDCT 变换,该变换是基于时域

混迭消除 TDAC(Time-domain Aliasing Cancellation) 技术的改进型 DCT,变换以帧为单位,相邻帧间有一半的样点重叠,能够很好地消除分帧引起的时间域混叠,降低边界效应,被广泛应用于数字音频编码系统的分析滤波器组中。因此,快速的 MDCT 实现算法是提高实时编码的关键技术之一。

本文提出的 MDCT 算法定义如下^[1]:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cos\left(\frac{\pi(2n+1+N/2)(2k+1)}{2N}\right) \quad (1)$$

$$k=0, 1, \dots, \frac{N}{2}-1$$

其中, $x(n)$ 表示音频样点, n 为音频样点的标号,取值范围为: $n=0, 1, \dots, N-1$ 。

$X(k)$ 表示频率系数, k 为频率系数的标号,取值范围为: $k=0, 1, \dots, \frac{N}{2}-1$ 。

N 表示每个帧包含音频样点的数量,选取需考虑如下因素:

- (1) N 选取为 2 的整数次幂,有利于快速算法的实现;
- (2) N 越大频谱分辨率越高,有利于提高压缩率;
- (3) N 越大,运算复杂度越高,平均每个样点的运算复杂度 C 与 N 值的关系可以表示为 $C \propto \log_2 N$;
- (4) 音频信号的特性是短时平稳的, N 值选取过大没有意义。

基于以上考虑, N 值可以在 128, 256, 512, 1 024, 2 048 中选取。

2.1.2 量化/熵编码

量化编码模块采用循环结构,根据压缩率的要求,采用非线性量化和 Huffman 编码来压缩数据,在保证音质的条件下,至少能够达到 5:1 的压缩率。具体实现步骤如下:

- (1) 时频分析模块输出的频率系数乘以定标因子 (Scaling Factor);
- (2) 对定标后的频率系数进行非线性量化;
- (3) 将量化后的频率系数两两组合成频率系数对(每帧共 $N/4$ 个频率系数对);
- (4) 选取 Huffman 码表,用该码表对频率系数对进行 Huffman 编码,生成比特流;
- (5) 判断 Huffman 编码产生的比特数是否超出指定压缩率的要求,如果超出就缩小定标因子,跳转到步骤(1),直至满足压缩率的要求退出迭代。

步骤(2)所述量化公式可以为: $y = \text{int}[x^{3/4}]$, 式中的 $\text{int}[\cdot]$ 函数表示取整数部分。步骤(4)所述码表可以选取为 1 个固定的码表。

2.1.3 封装

本文提出的 ESP 编码器还可以包括封装模块。一方面由于音频信号本质上是非平稳的,在一段时间内

技术与方法 Technique and Method

(10 ms 量级)可以看作是短时平稳信号,不同帧经量化编码输出的比特数有很大变化,通过将 M 帧音频信号量化编码后的比特流封装在 1 个比特块中,可以降低峰值比特流的影响;另一方面由于比特块的长度固定,ESP 解码器很容易确定每个比特块的起始位置,无需费时地进行比特块同步搜索,减小了系统的运算开销。

该模块将量化/熵编码模块输出的 M 帧比特流依次封装到 1 个固定长度的比特块中,然后再输出到 DRAM 缓冲器。若 M 帧比特流的长度之和小于比特块的固定长度,比特块的剩余部分可以用零填充。考虑到存储空间、编解码延迟、比特率平滑要求, M 可在 2~16 间选取。

2.2 解码器算法原理

ESP 系统中解码步骤包括:

(1)将 1 个比特块长度的比特流进行熵解码/反量化,得到重建的频率系数。

(2)对频率系数进行逆时频 IMDCT 分析,得到重建后的时域信号并输出播放。

2.2.1 熵解码/反量化

同量化/熵编码过程相反,熵解码/反量化模块的具体实现步骤如下:

(1)使用与 ESP 编码器相同的码表对输入缓冲区的 1 个比特块的比特流进行 Huffman 解码,解码得到 $M \times N/4$ 个频率系数对,其中每个频率系数对包含 2 个频率系数,共 $M \times N/2$ 个频率系数,以及 M 个反定标因子 $ISF[m]$ ($m=1, \dots, M$),其中每帧 1 个反定标因子。

(2)对解码得到的 $M \times N/2$ 个频率系数进行反量化处理,得到线性的频率系数。

(3)使用第(1)步解码得到的反定标因子 $ISF[m]$ ($m=1, \dots, M$),对第(2)步得到的线性频率系数进行反定标(二者相乘),得到反定标后的频率系数。

(4)将 $M \times N/2$ 个反定标后的频率系数,依次分为 M 帧,每帧 $N/2$ 个频率系数,输出至逆时频分析模块。

其中所述反量化的公式为 $y = \text{int}[x^{3/4}]$,其中 $\text{int}[\cdot]$ 函数表示取整数部分。

2.2.2 逆时频分析

逆时频分析模块将熵解码/反量化模块输出的频率系数转换为 PCM(Pulse Coding Modulation)信号。该模块采用 IMDCT 变换,其平均每个样点的运算复杂度 C 与 N 值的关系为 $C \propto \log_2 N$,本文提出的 IMDCT 算法定义如下^[2]:

$$\hat{x}(n) = \sum_{k=0}^{N/2-1} X(k) \cos\left(\frac{\pi(2n+1+N/2)(2k+1)}{2N}\right) \quad (2)$$

$$n=0, 1, \dots, N-1$$

其中, $X(k)$ 表示频率系数, k 为频率系数的标号,取值范围为: $k=0, 1, \dots, \frac{N}{2}-1$; $\hat{x}(n)$ 表示 IMDCT 变换后的

音频样点, n 为其标号,取值范围为: $n=0, 1, \dots, N-1$; N 表示每个帧包含音频样点的数量。

本文所提出的算法能够在使 CD 播放器音质基本无损失的情况下,舍弃了心理学模型,用较小的运算量,实现较高的压缩率从而达到较长的防震时间,提高动态随机存储器的使用效率,延长防震时间。同时算法采用基于块的运算结构,适合软件实现,运算复杂度低,有利于降低系统主频,从而降低系统功耗并提高系统的稳定性。该技术适合软件实现,能有效延长防震时间。

参考文献

- [1] CHENG Mu Huo. Fast IMDCT and MDCT algorithms a matrix approach[J]. IEEE Transactions on Signal Processing, 2003,51(1).
- [2] BYCOND G L. A new algorithm to compute the discrete cosine transform [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing.1984,32(6):1243-1245.

(收稿日期:2009-04-28)