

· 多媒体技术及应用 ·

多声道音频编码AC-3算法原理

贺前华 韦 岗 帅 林

(华南理工大学电子与通信工程系 广州 510641)

摘要 数字音频压缩标准AC-3算法已在广泛领域得到应用,如DVD、HDTV等,是发展家庭影院的主要技术之一。简要介绍AC-3算法的特点、帧结构及其编、解码原理,并指出对其相关技术进行研究的必要性。

关键词 多声道音频编码 比特分配 时域混迭消除

Principle of AC-3 Algorithm for Multichannel Audio Coding

He Qianhua Wei Gang Shuai Ling

(South China University of Technology Guangzhou 510641)

【Abstract】 Digital audio compression standard AC-3 algorithm has been widely utilized in various fields, such as DVD and HDTV, and is regarded as one of the key techniques for home theatre. The characteristics, frame structure, and the principle of encoder/decoder of AC-3 are introduced in this paper, meantime, it points out that it is necessary to study techniques used in AC-3.

【Key words】 Multichannel audio coding; Bit allocation; Time domain alias cancellation

Dolby AC-3是第一个专门为编码多声道数字音频信号而设计的感测编码系统,它可以满足单声道到5.1声道数字音频的编码要求。AC-3的6个声道是完全独立的,6个声道全为数码传输,意味着从节目混编控制台到家庭播放系统不会丢失任何信息。低频效果声道(LFE)传输额外的低频信息,使爆破、撕裂等低频效果显著增强,其频响限于3Hz~120Hz,称之为"0.1"声道。与其它编码系统相比, Dolby AC-3的编码效率、恢复音质以及通用性均是史无前例的。与模拟Dolby环绕声(一个带宽受限的环绕声道, 100Hz~7000Hz)相比, Dolby AC-3有两个完全独立的环绕声道(左环绕、右环绕),每个环绕声道与三个前向声道(左、中、右声道)一样的全频带保真度,使声音定位更准确。因此AC-3可产生真正的立体环绕声效果,使听者如临真实境界一样。

AC-3算法采用时域混迭消除技术,并运用人耳掩蔽效应,从而可对PCM信源进行高效压缩,恢复质量与原音相差无几。例如,一个5.1声道节目源的PCM代码超过5Mbps ($6 \times 48\text{kHz} \times 18\text{bits} = 5.184\text{Mbps}$), AC-3算法可将其压缩为384kbps的比特流。由于AC-3算法的诸多特点,已在AV领域得到了广泛应用,详情请参考文献[1]。

1 AC-3的帧结构—比特流语法

AC-3码流由一系列同步帧构成,同步帧结构如图1。每个同步帧由同步信息开始,后接一个比特流信息块,6个编码的音频数据块,音频数据后可有一个辅助数据区,同步帧以循环码冗余效验字CRC结束, CRC字用于出错校验。

同步信息(SI) 同步信息包括: ① 同步字(0x0B77);

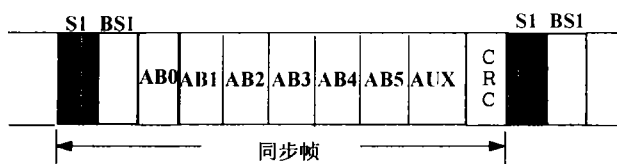


图1 AC-3同步帧结构

② 第一个循环码冗余效验字(应用于同步帧的前5/8部分); ③ 音频采样率码字; ④ 帧长码字。

比特流信息(BSI) 比特流信息由描述编码音频环境参数组成,共31个。包括比特流识别码、比特流模式、音频编码模式、中心混音电平、环绕混音电平、Dolby环绕模式、低频效果标志、压缩率存储标志、压缩率等信息。

音频数据块 每个音频数据块包含256抽样点(一个声道)的频率信息: 256频率系数,每个频率系数表示成指数和尾数两部分分别进行编码。具体表现为编码的指数,量化尾数,比特分配信息以及其它一些为达到更高压缩效率而采取的编码策略信息。第一个数据块须包括正确解码比特流所需的全部必要信息,而其它五个数据块可重用前面数据块的一些信息。

帧出错校验字 帧出错效验字包括1个CRC保留位和1个16位的循环冗余校验字,该校验字应用于整个同步

* 本研究得到东南大学移动通信开放实验室基金及广东省自然科学基金资助

贺前华 男, 33岁, 副教授, 主要从事语音、音频信息处理及自然语言理解等方面的研究

收稿日期: 1998-01-19

帧。当循环冗余校验字与同步字相同时，利用CRC保留位使CRC字与同步字不一样。

AC-3同步帧须满足以下约束条件，这些约束条件可减少AC-3解码器的输入存储缓冲区的大小。

- 1) 块0和块1的长度之和不大于帧长的5/8;
- 2) 块5的尾数与辅助数据长度之和不大于帧长的3/8;
- 3) 块0总包含有正确解码比特流所需的全部必要信息;
- 4) 当耦合标志从off变到on时，所有耦合信息都包含在此耦合标志为on的数据块中，而以前耦合标志为on的数据块中的耦合信息不再使用。

AC-3比特流语法对编码器没有提出具体要求，因此可根据具体要求设计复杂度不同的编码器，而解码器不需变动。

2 AC-3编码器

图2给出了AC-3编码的基本流程，下面介绍其中的几个主要模块。

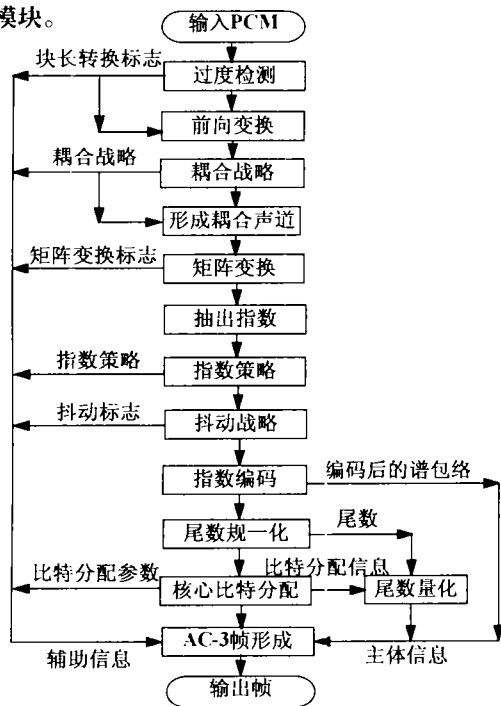


图2 AC-3编码的基本流程

2.1 过渡检测

过渡检测的目的是决定编码器何时使用长块(512样点)变换，何时使用短块(256样点)变换，使频域和时域分辨率的综合性能随信号的变化特征达到最佳。因为长块变换可使变化缓慢的时域信号得到最好的频率分辨率，而短块在时域信号变化快的情况下能获得更高的时间分辨率。过渡检测由高通滤波、块切分、峰值检测、阈值比较4步完成，检测结果用标记blksw表示。

2.2 时频变换

若标记blksw为0，进行一个512点TDAC变换得到256个频率系数，否则将512点的信号块分为两个256点的短块，分别进行TDAC变换，分别得到128个频率系数，将此两组系数一一交叠形成256个频率系数。以后的处理与一个单一的长块等同。设 $x[n]$ 表示加窗的输入时间序列，TDAC变换后的输出系数 $X_p[k]$ 由下式给出。

$$X_p[k] = -2/N \sum_{n=0}^{N/2} x(n) \cos(2\pi/4n(2n+1)(2k+1) + \pi/4(2k+1)(1+a))$$

$$0 \leq k < N/2 \quad (1)$$

若是长块变换， $N=512$ ， $a=0$ ，若是短块变换， $N=256$ ，对第一个短块 $a=-1$ ，对第二个短块 $a=1$ 。

2.3 声道耦合

由于人耳听觉系统主要根据临界带滤波后的信号谱包络，而不是信号本身定位信号的高频部分^[1]，因此可选择地对高频载体成分进行预先组合，从而删除了多余的高频位置信息，得到更高的编码增益。AC-3充分利用这一心理学现象，将高频子带信号分为谱包络和载波两部分，用较高的精度编码包络信息，如果需要，可选择地将不同声道的载波进行组合。由于定位信息保留在谱包络中，这一方法对恢复音质影响甚微。这一技术被称为声道耦合技术，具体是将参与耦合的声道的频率系数的平均值作为一个共同的耦合声道，而每个被耦合的声道保留一组独立的耦合系数，用来保存其原始声道的高频包络。

2.4 指数编码

时频变换的频率系数转换成浮点数形式：每个系数由指数和尾数组成。即使是对定点DSP芯片，这一表示方法能保证中间处理过程不损伤信号的动态范围。指数是频率系数二进制表示前导0的个数。其范围限定在0—24之间。对每一个声道，首先分析其指数序列随时频的变化情况确定所采用的编码策略，然后对指数序列进行差分编码(最大值为±2)，第一个指数总是4位表示。编码的指数代表编码的频谱包络，是计算比特分配的依据。

2.5 比特分配

AC-3的比特分配进程相对于人耳频率掩蔽效应和静听阈值分析音频信号的谱包络，以决定量化每个尾数的比特数。比特分配把所有声道作为一个整体考虑，根据信号特征，可将固定比特数在声道、频率、语音块之间进行最优分配。编码器反复调整粗SNR值(调整量为±6 dB)和细SNR值(调整量为±3/8 dB)，将允许的最大限度的比特数分配给所有的声道。如果当前分配的比特数超过允许分配的比特数，下一次计算时减小SNR的值，否则增加SNR的值，这种反复计算直到SNR值最大，且所分配的比特数不超过允许分配的比特数时才停止。其结果是获得的粗SNR值和细SNR值以及量化每个尾数的比特数。由于编码器将SNR值传送给解码器，解码器只需进行一次比特分配的计算流程。编解码器的比特分配计算均采用定点整数操作和查表操作，保证编解码器的比特分配结果完全一致。

掩蔽效应(masking effect)

当两个频率接近的不同强度的信号达到人耳时，弱信号(被掩蔽者)可能因为强信号(掩蔽者)的存在而不被人耳感觉到，这一心理学现象成为人耳的掩蔽效应。掩蔽阈值与掩蔽者的声压(SPL)、信号频率有关。相对而言，高频部分更容易被掩蔽。低于掩蔽阈值的信号对人耳而言是不可见的。因此利用掩蔽效应可对信道噪声进行屏蔽，信道噪声包括编码噪声、混迭失真或传输误差。

图3给出了掩蔽效应的定性示意图^[1]：掩蔽者的声压值与掩蔽阈值之差称为信号掩蔽比(SMR)，其最大值位

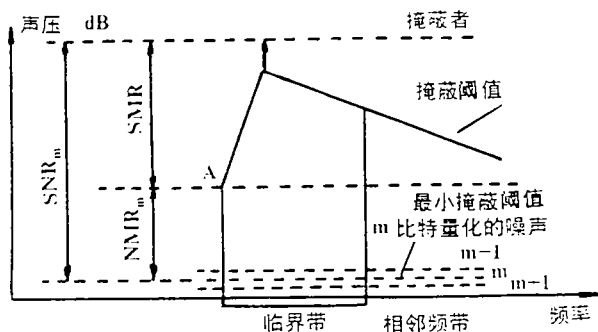


图3 掩蔽效应的定性示意图

于临界带最左端(A点)。在临界带内，只要信噪比(SNR)高于SMR，编码噪声不被察觉。若用SNR(m)代表m-bit量化后的信噪比，定义NMR(m)=SMR-SNR(m)(以db为单位)表示在给定的窄带中可觉察的失真。当量化比特数m减小时，NMR增大，但只要NMR小于零，失真便不会察觉到。

虽然声道间也存在掩蔽效应，但一个声道的信号屏蔽另一声道的噪声的能力很有限，并且与听者的位置有关，故一般只利用声道内的掩蔽效应。

2.6 尾数量化

所有归一化的尾数根据比特分配确定的量化精度进行量化，而不仅是保留尾数的n位最有意义比特作为其编码值。若量化比特数小于等于4，尾数采用对称均匀量化，以使量化误差最小。若量化比特数大于4，尾数采用非对称量化方式，采用一般的二进制补码形式表示。

3 AC-3 解码器

解码进程基本上是编码进程的逆过程，AC-3解码器的基本原理框图如图4。

帧同步及出错效验 解码器首先对AC-3比特流进行

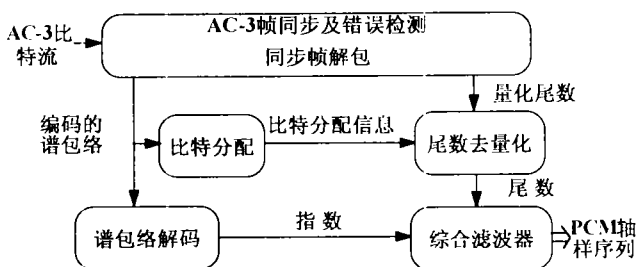


图4 AC-3解码器原理图

同步及CRC出错效验，若发现错误，解码器需采取一定的措施。另外解码器对帧内的每一个数据块进行比特流一致性检查，发现当前音频块有数据错误，解码器可使用上一数据块代替当前数据块，以消除数据传输错误。信号重建过程中的重迭相加特性确定这种出错处理方法构建的输出信号相当好，事实上一个好的音频数据块可重复使用数次。

帧解包 同步帧中的每一数据块的解包包括两个阶段，第一阶段将固定格式的数据解出，包括指数、耦合

系数，比特分配信息等。解码器利用这些数据恢复量化尾数的比特数。第二阶段利用比特分配计算出的指针解出变格式的量化尾数。频谱包络解码后产生指数序列。比特分配流程根据解出的指数序列及比特分配信息计算量化尾数的比特数，指数和尾数组合成转换系数并经逆TDAC变换形成时域中的PCM抽样序列。

实际的AC-3解码器比图4复杂得多，包括以下几种功能：

- 高频成分耦合的声道必须进行去耦合；
- 在两声道模式下，如果声道已经重组，必须进行反重组；
- 合成滤波器组的精度必须以编码器、分析滤波器组相同的方式进行动态调整。

· 标准的声音增减技术。该技术使任意声道数的AC-3码流可通过具有任意输出声道数的播出系统播放，且播出的音频质量是最佳的。

4 小结

数字音频压缩标准AC-3算法已在广泛领域得到应用，如VCR's、DVD、激光视盘、HDTV、多媒体等，是发展家庭影院的关键技术之一。Dolby AC-3的软硬件发展势头非常迅猛，如Zoran公司的ZR-38000芯片，其发展态势不但表明了杜比数码技术可满足人们对高质量音质的期望，更重要的是显示了人们对环绕立体声产品需求的庞大市场，这一需求不可避免地带来一场AV产业革命。当DVD大量上市及HDTV技术成熟之日，便是Dolby AC-3家庭影院普及之时。我国相关产业正处在起步阶段，因此对AC-3所采用的一些关键技术进行研究是非常必要的。

本文简要介绍了AC-3算法的特点、帧结构和编解码原理，供同行参考指正。另外我们在这方面已开展了一年多的工作，已用高级语言实现了AC-3编解码器，目前正将其转化为DSP语言，并打算用常规DSP设计实现其硬件系统。

参考文献

- 1 贺新华，韦岗. Dolby Digital (AC-3) 在AV领域中的应用. 电子科技导报，1998(4): 33-36
- 2 Yost A, Gourevitch G. Directional Hearing, Springer-Verlag, New York, 1987
- 3 Noll P. Wideband Speech and Audio Coding. IEEE Communications Magazine, 1993(11): 34-44
- 4 Digital Audio Compression Standard (AC-3). Doc. A/52, 1995-12-20

~~~~~

## Intel出台新的多媒体微处理器

Katmai是Intel公司继MMX技术之后多媒体技术的新发展，目的是加快高级软件应用的处理速度。Katmai中最重要的新指令是一种新的存储器streaming体系结构。据Intel称，如果没有这种新技术，当取数、执行和存储后，在从cache进行另一取数前要等待。有了Katmai中的新的存储器streaming体系结构后，基本上可以不用等待就可立即取数。性能的提高随应用而变化，但初步数据表明许多先进的应用提高10%左右。Katmai将出现在Intel公司1999年推出的多种微处理器中，其中第一种基于0.25 μm的工艺；运行速度为450—500MHz，预计1999年第1季度出货。1999年下半年将有两种采用Katmai指令的微处理器推出，采用0.18 μm的工艺，时钟频率大大高于500MHz，并具有综合的L2 cache存储器。（张）