

# 高频重建技术 SBR 的研究与实现

党 辰, 戴 葵, 王苏峰, 刘 芸, 王志英

(国防科技大学计算机学院计算机系, 湖南长沙 410073)

**摘 要:** AAC 是 MPEG-4 标准的音频编码规范, 为了提高音频数据的压缩率, 往往忽略了音频中所包含的高频成分, 只对低频部分进行编码, 因此 AAC 解码器设计中高频重建技术对提高音频质量起着非常重要的作用. 本文论述了高频重建技术的原理, 并采用 SBR 技术, 实现了 AAC 音频编码的高频重建, 同时性能评测结果表明所实现的 AAC+SBR 系统满足实际应用的要求.

**关键词:** 高频重建; AAC; SBR; aacPlus

**中图分类号:** TP302 **文献标识码:** A **文章编号:** 0372-2112 (2004) 12A-189-03

## The Research and Implementation of SBR High Frequency Reconstruction Technology

DANG Chen, DAI Kui, WANG Shu-feng, LIU Yun, WANG Zhi-ying

(School of Computer, National University of Defense Technology, Changsha, Hunan 410073, China)

**Abstract:** Advanced Audio Coding (AAC) is the standard audio coding specification of MPEG-4. As description of AAC audio coding, the high frequency part of audio is always omitted, and only the low frequency part of audio will be coded to enhance the audio compression efficiency. The high frequency reconstruction technology plays an important role in the AAC decoder design to improve the output audio quality. This paper discusses the basic principles of high frequency reconstruction based on spectral bandwidth replication (SBR) technology. The implementation of an AAC decoder with SBR is presented in this paper. Some evaluation results of the implementation show its efficiency of audio decoding.

**Key words:** high frequency reconstruction; advanced audio coding (AAC); spectral bandwidth replication (SBR); aacPlus

### 1 引言

高频信号的损失是目前数字音频编码存储技术不可避免的缺点, 而失去高频成分会使声音听起来沉闷、不明亮. 而高频重建技术的目的就是数字音频解码时能够恢复原始音频信号在进行数字编码的过程中所失去的高频成分, 从而提高声音播放的品质.

MPEG-4 (Moving Picture Experts Group) AAC (Advanced Audio Coding)<sup>[1]</sup>是目前最有效的感知音频编解码方案. 在 MPEG-4 AAC 数字音频编码方案中, 为了提高音频数据的压缩效率, 它往往抛弃音频数据中的高频信息, 从而使得原始音频信息中的高频成分受到损失, 造成音质的下降.

Coding Technologies 公司所提出的 SBR (Spectral Bandwidth Replication) 频率带宽复制技术是一种非常有效的数字音频高频重建技术, 已被 MPEG-4 的数字音频实现参考模型<sup>[2]</sup>采用, MPEG-4 AAC 数字音频编解码方案和 SBR 技术相结合, 从而构成了当前一种既具有很好的压缩效果又具有很好输出音质的数字音频编解码方案: aacPlus.

本文在论述高频重建技术原理的基础上, 给出了一种实用的 AAC+SBR 编解码器体系结构, 并在此基础上, 设计实

现了一种 AAC+SBR 的数字音频解码器. 解码器的解码结果显示解码后的音频质量与源音频质量相比非常接近, 性能评测结果表明该解码器能够满足实际系统的应用要求.

### 2 高频重建技术

音频中的高频成分代表了声音的细腻度, 是人们追求的高品质音频的必不可少的重要成分. 高频的损失有两方面的原因. 一方面由于人耳的听觉频率范围在 20Hz~20KHz, 对 1~3KHz 频率的音频信号较敏感, 而对高频信号相对来说并不十分敏感. 因此, 许多音频数字编码技术中在进行信号压缩处理时为了达到较好的压缩效果, 往往抛弃原始音频信号中的高频成分. 另一方面, 由于抽样定理要求输入信号的最高频率要小于采样率的二分之一, 因此在许多音频数字编码技术中采用低通滤波器来限制输入信号的频率, 从而造成音频数据中高频成分的丢失. 为了重建音频信号中的高频成分, 提高数据音频的品质, 当前各种数字音频解码器都是根据声源的物理特性所决定的声音信号的谐波 (harmonic) 特性, 使得低频信号与高频信号有着高度的相关性, 然后根据这种相关性来重建高频成分.

高频重建技术 (如图 1 所示) 根据音频数字编码端是否传

送相应的高频重建参数信号而分为“盲式”和“非盲式”两类。“非盲式”重建法必须根据编码端传送的信息来辅助高频信号的重建,这些辅助信息会占用一定的存储空间,但该方法能够较精确地实现高频重建。“盲式”<sup>[3]</sup>重建法不需要额外的辅助信息,而直接进行高频重建,其重建的语音信号中的高频成分的品质显然要低于“非盲式”高频重建方法。

SBR 是“非盲式”(Non blind)重建方法的一种,它必须根据编码端传送的少量参数信息来重建语音信号中的高频成分。由于高频重建有辅助信息支持,因此该方法具有较佳的重建品质。

### 3 AAC+ SBR 编解码系统

本文 AAC+ SBR 解码器的实现是基于 AAC 低复杂度(LC)实现框架,它是一种高品质音频框架。语音编解码系统的原理图如图 2 所示。

在编码端,在对音频信号进行 AAC 编码的同时,也对其进行 SBR 编码。经过 SBR 编码可以获得解码器为进行高频部分重建所需要的某些附加信息。而输入给 AAC 音频编码器的原始语音信号,通过滤波器组将时域信号变为频域信号。同时利用听觉心理模型计算各子带的当前掩蔽阈值。同时还输出时域噪声整形(TNS)、中间边立体声(M/S)和强度立体声所需的控制信息,TNS 用来控制量化噪声的瞬时形状,解决掩蔽阈值和量化噪声的错误匹配,强度立体声模块利用人耳对高频信号不敏感的特点,只传送高频信号的包络,大大降低了编码所需的比特数;M/S 模块利用左右声道的相关性,以左右声道的和信号 M 与差信号 S 代替左右声道,因为 S 基本上为零,从而可以

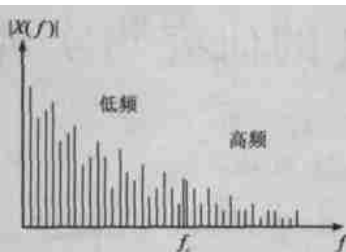


图 1 高频重建示意图

其中:  $f_c$  为低频的截止频率,接收端只收到低频部分,然后重建高频部分,并根据所传的参数对高频部分进行调整,从而恢复出完整带宽的频带。

有效减少编码后的音频流的数据量; AAC 将整个频谱划分为若干个比例因子带,每个比例因子带拥有一个作用于所有谱线的比例因子,从而改变谱线的振幅大小;量化模块进行比特分配,量化后的信号与比例因子进行哈夫曼编码,所编码的 AAC 码流最后与 SBR 编码的信号复合成 AAC+ SBR 码流。

解码是编码的逆过程, AAC 音频解码器的低频输出馈送给 SBR 解码单元,借以再生高频,并产生一个全带宽的音频信号。以上核心 AAC 编解码器的原理及其现在许多文章<sup>[4,5]</sup>中都有论述,技术已经很成熟,就不再详述。

### 4 SBR 解码和低频重建

AAC+ SBR 解码过程如图 3 所示。其中分析滤波器组采

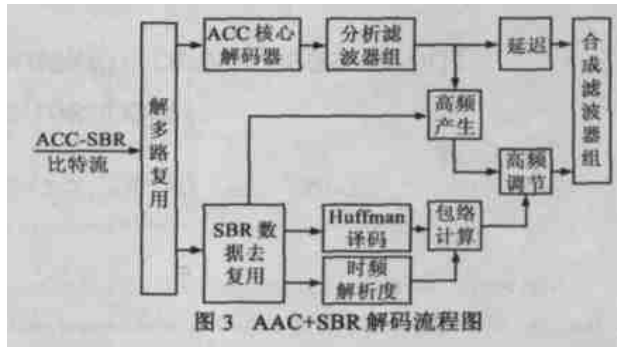


图 3 AAC+SBR 解码流程图

用了 32 个子带正交镜像滤波器组(QMF Quadrature Mirror Filter),对核心 AAC 解码得到的时域信号进行子频带分析。采用 QMF 的优点是可以抵消由于子带分割所带来的混叠效应。当声音信号经过滤波器组后会被转成 32 个等频宽的子频带信号。每输入 32 个时域信号就做一次子频带分析,从而可以得到 32 个输出(每一个子频带对应一个输出结果)。

对于不同长度的 AAC 帧,每个子带输出的样点值是不同的,设为 numSamples。输出保存在矩阵  $X_{low}(k, l + t_{HFGen})$ ,  $0 \leq k < 32, 0 \leq l < \text{numSamples}$ , 其中  $k$  表示第  $k$  个子频带,  $l$  表示第  $l$  个样点,  $t_{HFGen}$  是为了同步信号而加入的延迟。

与 QMF 输出的子带相关的是频率波段表,它描述了在频域上 QMF 子带的分组情况,对每一个包络都有两个表,高频解析表  $f_{TabHigh}$  和低频解析表  $f_{TabLow}$ ,对于噪声也有相应的频率波段表  $f_{TabNoise}$ ,以上各表通过主频带表  $f_{Master}$  导出。

Huffman 解码功能部件的任务就是把编码字转化成已量化的包络和噪声采样点。时频解析度功能部件则产生包络和噪声数量以及每一包络和噪声相关的开始和结束时间边界。此处,它还将给出每一个包络所采用的频带表。

高频产生功能部件首先将低频带的音频信号完全复制到高频带,然后根据编码器所传送的 SBR 重建参数对高频带进行“逆向滤波”(inverse filtering),从而适应各帧的信号频谱特性。最后所得的补偿高频信号计算如下:

$$X_{High}(k, l + t_{HfAd}) = X_{Low}(p, l + t_{HfAd}) + bwFac \times \alpha_0(p) \times X_{Low}(p - 1, l + t_{HfAd}) + [bwFac]^2 \times \alpha_1(p) \times X_{Low}(p - 2 + t_{HfAd}) \quad (1)$$

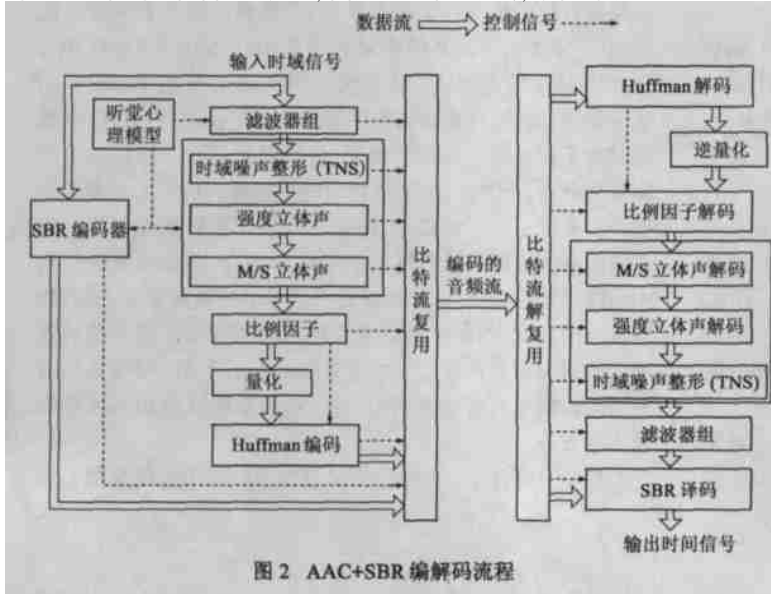


图 2 AAC+SBR 编解码流程

其中  $buFac$  为共振因子 (chirp factors), 由噪声波段表  $f_{TableNoise}$  定义,  $\alpha_0, \alpha_1$  为逆向滤波器系数,  $0 \leq l < numSamples$ ,  $k$  为要补偿的高频带,  $p$  为与高频相关的低频带,  $t_{HfAdj}$  是为同步信号而加入的延迟。

包络调节功能部件最后根据包络参数, 对高频带进行调整, 使语音输出信号的高频频谱与源语音信号相同, 从而完成整个 SBR 高频重建工作。当前的 SBR 的频谱包络在时间间隔上以时频解析度功能部件所给出的时频解析度为参数按照下式计算而得:

$$E_{Curr} = \frac{\sum_{k=k_0}^{k_e} \sum_{l=l_0}^{l_e} |X_{High}(k, l)|^2}{(k_e - k_0 + 1)(l_e - l_0 + 1)}$$

(2)

其中:  $l_0$  为起始时间索引,  $l_e$  为结尾时间索引,  $k_0$  为起始频带索引,  $k_e$  为结束频带索引。这些索引值由高频解析表  $f_{TableHigh}$  和低频解析表  $f_{TableLow}$  给出。包络调节功能部件最后根据当前估计的包络参数  $E_{Curr}$  对高频子频带进行增益调节, 从而输出高频信号矩阵  $Y$ 。

核心 AAC 解码器输出的低频信号被分析滤波器组分成 32 个子频带经 SBR 算法重建高频后共有 64 道子频带, 包括延迟的低频频带  $X_{Low}$  和高频频带  $Y$ , 最后进入合成滤波器组合成为完整频宽的 AAC+ SBR 输出信号。

5 功能和性能评测

上述 AAC+ SBR 解码过程已经在计算机中通过软件实现, 中间过程的频谱分析图如图 4 所示。图 4(a) 表示原始信号的频谱, 图 4(b) 表示 AAC 核心解码器输出的音频信号低频频谱, 图 4(c) 表示高频产生功能部件复制低频后的音频信号频谱, 图 4(d) 由高频调节功能部件进行高频调整后的音频信号频谱。比较图 4(a) 和图 4(d) 的频谱可以看到非常相似, 而且从实际的听觉效果来说两者相差无几, 并且要比图 4(b) 仅由 AAC 核心解码器所输出的音频效果听起来显得明亮、悦耳。

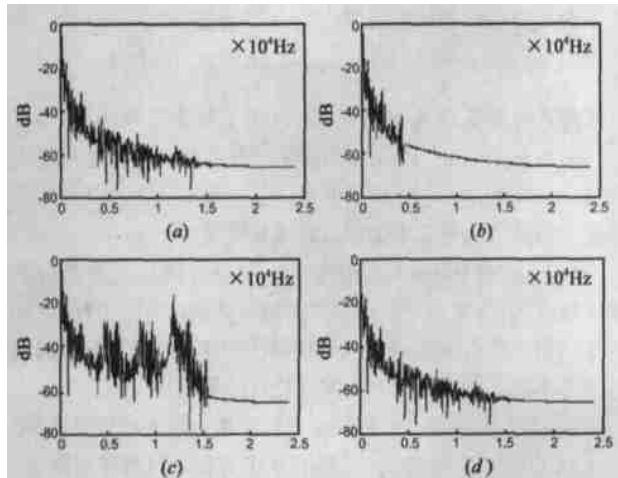


图 4 AAC + SBR 和 AAC 核心编码器输出频谱比较

由于 AAC+ SBR 解码算法非常复杂, 所占用的计算时间比较多, 这里对软件实现的解码器中的主要模块进行了性能分析。测试的环境为: P4 1.7G 微机, 内存 256M, 操作系统 Win

表 1 各功能单元性能测试参数表

| 模块<br>类别   | AAC 核心<br>解码器 | 分析滤<br>波器组 | 高频<br>产生 | 高频<br>调节 | 合成滤<br>波器组 | 其它       |
|------------|---------------|------------|----------|----------|------------|----------|
| 时间(秒)      | 1.742208      | 0.357731   | 0.093640 | 0.287600 | 0.587667   | 0.052312 |
| 所占时间比率     | 55.8193%      | 11.4615%   | 3.0002%  | 9.2145%  | 18.8285%   | 1.6761%  |
| 模块时间/帧(ms) | 1.893705      | 0.388838   | 0.101783 | 0.312608 | 0.638769   | 0.056861 |

注: 以上 920 帧总的解码时间为 3.121158 秒, 平均每帧解码约需 3.392563 毫秒。

dows XP, AAC+ SBR 解码器采用 Microsoft Visual C++ 6.0 实现。在测试过程中采用了如下测试参数: AAC 标准配置信道 (即单声道 AAC 核心解码器, 不使用 M/S 工具和强度立体声工具), AAC 核心解码器的采样率为 24KHz, 音频信号帧为 920 帧。通过分析, 可得如表 1 所示的测试结果。

从表 1 可以看到, 虽然 SBR 技术有效增强了解码器输出声音的品质, 但同时也占用了相当多的计算时间, 基本上与 AAC 核心解码所占用的时间相当。在我们所实现的 AAC+ SBR 解码系统中, 平均每解码一个音频帧所用的时间为 3.392563 毫秒 (各功能单元处理每帧音频数据所用时间相加), 计算性能能够满足实际应用的需要, 能进行实时解码与音频输出。

6 结论

本文在深入研究高频重建技术的基础上, 利用 SBR 高频重建技术, 设计实现了一种 AAC+ SBR 音频解码系统。该系统能够进行实时音频解码和输出, 在输出声音的音质和解码运算性能等方面达到了实际应用的要求, 能满足基于模拟/数字电话线路或无线系统和广播系统的应用。

参考文献:

[ 1 ] ISO/IEC 14496-3: Information Technology-Coding of Audior Visual Objects Part 3: Audio[ S ]. 2001.

[ 2 ] ISO/IEC 14496-3: Information Technology-Coding of Audior Visual Objects Part 3: Audio: AMENDMENT 1: Bandwidth Extension [ S ]. 2001.

[ 3 ] Jeongil Seo, Daeyoung Jang, Jinwoo Hong, Kyeoungok Kang. A Simple Method for Reproducing High Frequency Components at Low Bit Rate Audio Coding[ R ]. the 113th AES Convention, Los Angeles, 2002.

[ 4 ] 岳玲, 陈健, 徐盛. MPEG-4AAC 解码器在 TMS320C6416 上的实现[ J ]. 声电技术, 2003, ( 11 ): 35- 39.

[ 5 ] 王华明, 陈健. MPEG-AAC 音频编码技术及其软件解码器的实现[ J ]. 计算机工程, 2001, ( 6 ): 51- 53.

作者简介:



党 辰 男, 1971 年 6 月出生于四川省, 现为国防科大硕士研究生, 主要研究领域为 计算机系统结构、数字信号处理。E-mail: director\_2001@sohu.com.

戴 葵 男, 1968 年 11 月出生于湖北省, 博士, 副教授, 主要研究领域为系统结构、微处理器设计技术。