

⑫
45-48
MPEG 音频编码算法的研究与实时实现

王建昕, 董在望, 印昉强

清华大学 电子工程系, 北京 100084

TN 912.3

文 摘 为了改进 MPEG 音频编码算法的主要模块——子带滤波和掩蔽门限计算, 以利于实时实现, 提出了修正的高散余弦变换(MDCT)的快速算法, 使得子带滤波组的计算量降低约 30%, 并保持简明对称的结构。指出了一种简化的心理声学模型及其计算方法, 适合于低成本编码器的需要。采用双数字信号处理芯片实现了实时编码器, 能完成 MPEG 规定的多种抽样率、多种编码模式以及变速率输出等所有要求。

关键词 宽带声音; 压缩编码; 快速算法; 心理声学模型; 实时实现

分类号 TN 912

音频编码算法
MPEG, MDCT

宽带声音的压缩编码在近十年得到了迅速发展。宽带声音的带宽达 20 kHz, 一般采用 48 kHz 抽样和 16 bit 量化, 其压缩编码更重视保真度。

宽带声音一般采用基于人耳感觉特性的频域波形编码, 例如子带编码和变换编码, 其统一的框图可由图 1 表示。

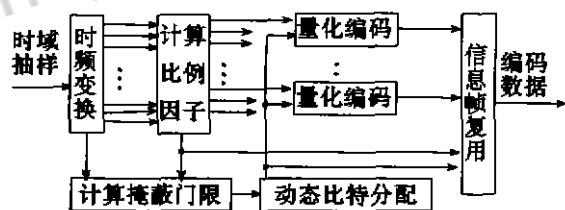


图1 宽带声音编码的统一框图

目前, 大多数压缩算法能在每声道 100 kbit/s 左右(压缩比 6:1~8:1)保证与 CD 相当的重建质量, 包括以子带编码为核心的 MPEG (moving pictures expert group) 音频编码标准 (ISO 11172-3)^[1] 和变换编码方案 PAC^[2]、Dolby AC-2^[3] 等。

收稿日期: 1996-10-14

第一作者: 男, 1968 年生, 博士研究生

MPEG 音频算法不仅具有很高的主观评价质量和适中的复杂度, 而且其解码复杂度低, 延时小, 独立性强, 因而被广泛地采用。

本文以 MPEG 音频算法第二层的设计为基础, 讨论其主要模块——子带滤波和掩蔽门限计算的改进算法, 并给出基于数字信号处理 (DSP) 芯片的实时编码器。

1 MPEG 音频编码的关键技术

1.1 子带滤波器组的设计

宽带声音的子带编码一般采用多相结构的正交镜像滤波器组 (QMF) (图 2), 抽取和内插保证系统工作在临界抽样状态。

MPEG 音频算法采用了准重构的多相正交滤波器组^[4], 子带数 $M=32$ 。首先设计一个有限冲激响应的原型低通滤波器, 保证其有 96 dB 以上的阻带衰减, 然后通过余弦函数的调制得到整个滤波器组。

分析和合成滤波器组的冲激响应分别为

$$h_i(n) = 2\cos\left[\frac{(2i+1)(n-\frac{M}{2})\pi}{2M}\right]h(n),$$

$$f_i(n) = 2M\cos\left[\frac{(2i+1)(n+\frac{M}{2})\pi}{2M}\right]h(n)$$
(1)

其中 $i=0, 1, \dots, M-1$; $h(n)$ 是原型低通滤波器的冲激响应, 其设计采用 Johnston 准则^[5]。

1.2 人耳感觉特性的应用及动态比特分配

在宽带声音压缩编码中, 利用同时掩蔽效应和动态比特分配对量化噪声整形, 使之不为人耳所察觉。图 3 给出了一个典型的掩蔽门限曲线。当采用感觉物理量临界频带率 (α , 单位 Bark^[6]) 代替客观物理量频率时, 曲线的形状将基本上不随掩蔽音的强弱和频域位置而改变^[6]。

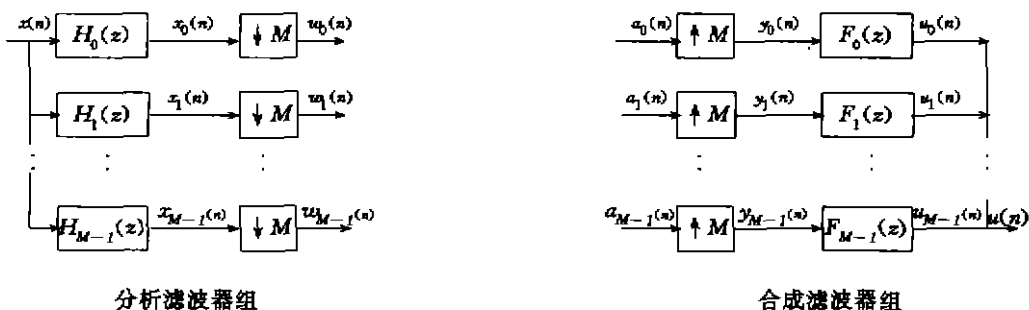
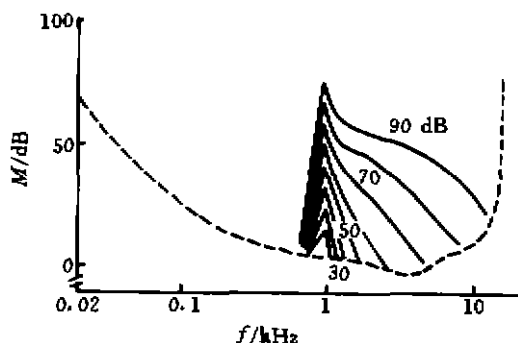


图2 子带滤波器组框图

图3 频率为1 kHz、不同声压级的纯音性掩蔽音对纯音的掩蔽曲线
(虚线是绝对可闻门限)

首先对声音序列进行傅里叶变换,以保证有较高的频域分辨率。然后,由频域抽样的声压级特性的不同提取纯音性和噪声性掩蔽音,利用统一的掩蔽门限模型(见图4)确定每个掩蔽音产生的掩蔽门限。最后,采用一定的叠加法则得到各被掩音上的总体掩蔽门限。

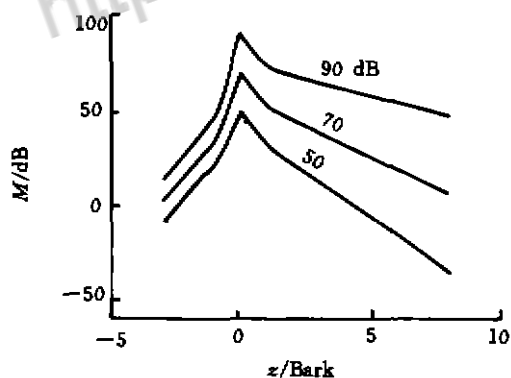


图4 MPEG 音频算法中采用的掩蔽门限模型

定义每个子带的信号-掩蔽门限之比(信掩比)

$$R_{S/N}(i)/\text{dB} = \min_p 10 \cdot \lg \frac{S(i)}{M_{\text{TH}}(i, p)} \quad (2)$$

式中 $M_{\text{TH}}(i, p)$ 是子带 i 中第 p 个被掩音上的总体掩蔽门限, $S(i)$ 为子带 i 的信号能量。

动态比特分配的策略是,依次增加各个子带所占用的比特数,保证各子带的掩噪比

$$R_{M/N}(i) = R_{S/N}(i) - R_{S/M}(i) \quad (\text{dB}) \quad (3)$$

大致相等。式(3)中, $R_{S/N}(i)$ 和 $R_{S/M}(i)$ 分别是子带 i 的信噪比和信号-掩蔽门限比。

掩噪比越大,说明噪声被掩蔽得越好。因此,在给定编码后比特率的情况下,可以通过动态比特分配充分利用比特资源;相反,如果对重建质量有固定要求,也可以通过合理的分配尽可能地减小比特率。

2 算法改进

2.1 子带滤波器组的快速算法

在图2中,设输入的时域抽样为 $x(n)$, 则第 i 个子带的输出样点为

$$w_i(n) = \sum_{l=0}^{L-1} 2 \cos \left[\frac{(2i+1)(l - \frac{M}{2})\pi}{2M} \right] h(l) x(nM-l) \quad (4)$$

用上式进行分析滤波计算每输出一组样值 $\{w_i\}$, $i=0, 1, \dots, M-1$, 需 $2ML$ 次实乘和 ML 次实加。

将余弦函数的周期性

$$\cos \frac{(2i+1)(k+2rM - \frac{M}{2})\pi}{2M} = (-1)^r \cos \frac{(2i+1)(k - \frac{M}{2})\pi}{2M}$$

$k=0, 1, \dots, 2M-1$, $r=0, 1, \dots, \frac{L}{M}-1$, 代入式(4)有

$$w_i(n) = \sum_{k=0}^{2M-1} \cos \left[\frac{(2i+1)(k - \frac{M}{2})\pi}{2M} \right] s(k, n) \quad (5)$$

其中

$$s(k, n) = \sum_{r=0}^{L/M-1} [(-1)^r \cdot 2 \cdot h(2rM+k)] \cdot x[nM - (2rM+k)] \quad (6)$$

而 $(-1)^r \cdot 2 \cdot h(2rM+k)$ 可以系数的形式给出。

用式(5)和(6)计算一组 M 个子带输出需实乘

$M(L/M+2M)$ 次,实加 $M(L/M+2M)$ 次。

实际上,式(5)是一个修正的离散余弦变换(MDCT)反变换。本文作者提出了一种快速 MDCT 算法,将式(5)简化成移位、求差等预处理和快速余弦变换(FCT)两部分,具体步骤如下^[7]:

1) 循环移位

$$s_1(k,n) = \begin{cases} s(k+\frac{M}{2},n), & k=0,\dots,\frac{3}{2}M-1 \\ -s(k+\frac{M}{2}-2M,n), & k=\frac{3}{2}M,\dots,2M-1 \end{cases} \quad (7)$$

2) 构造半序列

$$s_2(k,n) = s_1(k,n) - s_1(2M-k,n) \quad (8)$$

其中 $k=0,1,\dots,M-1$,并令 $s_1(2M)=0$ 。

3) 利用快速余弦变换(FCT)^[8]作 M 点 IDCT

$$w_i(n) = \sum_{k=0}^{M-1} s_2(k,n) \cos \frac{(2i+1)k\pi}{2M}, \quad i=0,1,\dots,M-1 \quad (9)$$

这样,实乘降为 $M(L/M+\frac{1}{2}\lg M)$ 次,实加降为 $M(L/M+\frac{2}{3}\lg M)$ 次,而且保持了 FCT 简明对称的结构。表 1 列出了当 $L=512$, $M=32$ 时,三种计算方法的运算量实乘次数 C_{mul} 和实加次数 C_{add} 。

表 1 子带分析滤波器组运算量
($L=512, M=32$, 每输出 32 个样值)

计算方法	C_{mul}	C_{add}
方法一 [式(4)]	32 768	16 384
方法二 [式(5)、(6)]	2 560	2 560
方法三 (快速 MDCT)	592	752

利用类似的方法可以得到合成滤波器组的快速算法,这里不再赘述。

2.2 心理声学模型算法的简化

为了降低计算掩蔽门限的复杂度,可以采用一些近似得到一个掩蔽门限简化算法。

首先考虑掩蔽门限模型的简化。当用临界频带率 z 表示信号的频率时,

近似一:掩蔽门限曲线与掩蔽音本身的频率位置和声压级无关。

近似二:掩蔽门限的大小与掩蔽音的声压级成正比。

这样,可以用两段直线近似表示图 3 中纯音的掩蔽门限曲线。临界频带率为 z_p 的掩蔽音 p 在被掩音 q 处产生的掩蔽门限 $M_{TH,p}(q)$ 为

$$M_{TH,p}(q) = L_p + v_p - 28 \cdot (z_p - z_q), \quad z_p \geq z_q \quad (10.a)$$

$$M_{TH,p}(q) = L_p + v_p + 10 \cdot (z_p - z_q), \quad z_p < z_q \quad (10.b)$$

其中 L_p 和 v_p 分别是第 p 个掩蔽音的声压级和掩蔽系数。

其次,采用子带滤波器组的结果直接计算掩蔽门限,从而省略运算量较大的傅里叶变换。为此,需采用另一近似。

近似三:子带信号构成频率在该子带内的一个纯音性掩蔽音。

近似四:只考虑最严格的掩蔽,即掩蔽效应最弱的情况,计算各子带的掩蔽门限。子带的掩蔽门限可以分成两类:带内掩蔽:由本子带的信号产生的掩蔽门限;带外掩蔽:由其它子带的信号在该子带内产生的掩蔽门限。

对于每个子带,分别计算其高、低端截止频率处的带内和带外掩蔽门限,再根据掩蔽门限的叠加关系,就可以求得各子带的最小掩蔽门限。

3 MPEG 音频编码器的实时实现

本文作者开发了基于 DSP 芯片 TMS320C31 的实时 MPEG 音频编码器。初步主观测试表明,重建质量很好。

MPEG 音频算法不仅具有很高的复杂度,而且有多种抽样率(48 kHz, 44.1 kHz, 32 kHz)、多种编码模式(单通道,双通道,立体声,联合立体声)以及变速率输出(32~384 kbit/s)等要求,为此,编码器采用双 DSP 芯片流水计算的结构。按照运算量,第一片完成数据采集、子带滤波和计算掩蔽门限的快速傅里叶变换部分,同时负责编码器的控制,第二片完成掩蔽门限的计算、动态比特分配和子带抽样的编码。

在表 2 的各模块中,分析滤波和心理声学模型的计算在编码器运算量中占有相当大的比重。在子带滤波中 MDCT 快速算法可减少 4.5 MI/s (I 为指令)。

表 2 MPEG 音频编码器复杂度分布

	Word 代码量	Word 数据量	最大运算量/ (MI·s ⁻¹)
分析滤波器组	111	2 560	10.36
心理声学模型	759	6 901	16.60
动态比特分配	214	4 260	6.26
量化编码	162	99	1.78
其它(控制、I/O)	739	390	3.48
合计	1 985	14 210	38.48

图5表明,简化算法得到的掩蔽门限一般小于MPEG的结果。简化算法得到的是一个相对严格的掩蔽曲线,以它为准进行的比特分配需要较多的比特数。实验证明,在同样的质量要求下,简化算法每通道的编码速率增加约20%,但是,简化算法的运算量减少 $11 \text{ MI} \cdot \text{s}^{-1}$,占整个编码器的29%,因而对于实现低成本编码器具有较高的价值。

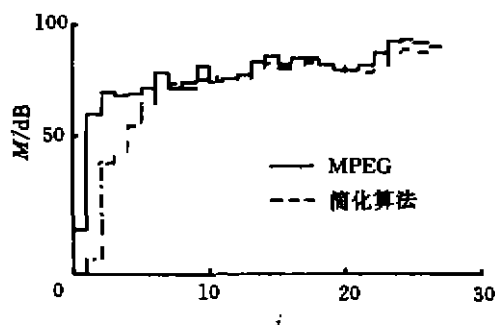
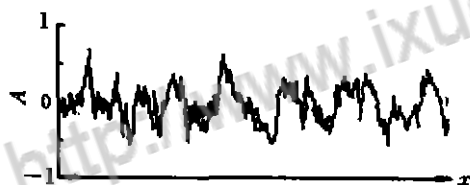


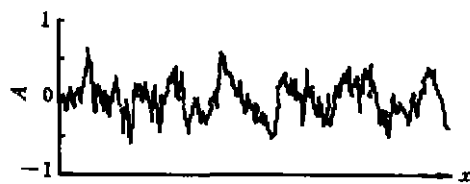
图5 MPEG和简化算法的子带掩蔽门限比较

从图6给出的乐段波形可以看出,编解码器在动态范围较大处也有较好的跟踪特性。

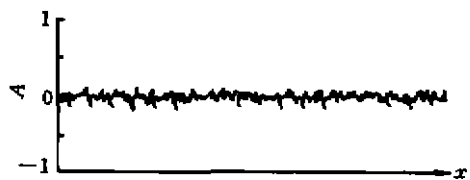
重建信号的分段信噪比(SNRSEG)一般在20~30 dB,与主观评价似乎不相称。这是因为宽带声音编码器虽然是波形编码,但利用掩蔽门限进行了噪声整形,保留了相当大的不可闻噪声。



(a) 原始波形



(b) 重建波形



(c) 误差波形

图6 声音波形比较(卡门序曲)

4 结 束 语

本文讨论了MPEG音频编码的原理与改进算法,并给出了基于DSP芯片的实时编码器。对两个主要模块——子带滤波和掩蔽门限计算的简化使得算法的复杂度大大降低。MPEG音频编码器可应用于数字音频广播(DAB)、数字音频工作站等许多领域。

参 考 文 献

- 1 ISO/IEC 11172-3. Information technology—Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s—Part 3: Audio. ISO, 1993-08
- 2 Jayant N S. The AT&T DAR system update. In: NAB 1994 Broadcasting Eng Conf Proc, 1994: 389 ~ 398
- 3 Fielder L D, Davidson G A. AC-2, A family of low complexity transform-based music coders. In: Proceedings of AES 10th International Conference, 1991. 57 ~ 70
- 4 Rothweiler J H. Polyphase quadrature filters, a new subband coding technique. In: IEEE ICASSP'83 Proceedings, 1983. 1 280 ~ 1 283
- 5 Johnston J D. A filter family designed for use in quadrature mirror filter banks. In: IEEE ICASSP'80 Proceedings, 1980. 291 ~ 294
- 6 Zwicker E, Fastl H. Psychoacoustics-facts and models. Berlin: Springer-Verlag, 1990
- 7 Wang Jianxin, Dong Zaiwang. A fast algorithm for modified discrete cosine transform. In: ICCT'96 Proceedings, 1996. 13.08.1 ~ 13.08.4
- 8 Lee B G. FCT-a fast cosine transform. In: IEEE ICASSP'84 Proceedings, 1984. 28A. 3.1 ~ 28A. 3.4

Study and real-time implementation of MPEG/Audio coding algorithm

Wang Jianxin, Dong Zaiwang, Yin Fangqiang
Department of Electronic Engineering,
Tsinghua University, Beijing 100084

Abstract This paper presents improvements on the subband filtering and masking-threshold calculation-two main parts of the MPEG/Audio coding algorithm. A new fast algorithm is proposed for calculating the modified discrete cosine transform(MDCT). The algorithm saves about 30% of the analysis filtering computation and still renders a simple and symmetric structure. A simplified psychoacoustic model especially suitable for low-cost implementation is introduced, which is proved to significantly reduce the coding complexity. The real-time encoder which utilizes the improvements can satisfy all the requirements of different sampling rates, bit rates and operating modes.

Key words wideband audio; compression coding; fast algorithm; psychoacoustic model; real-time implementation



知网查重限时 7折 最高可优惠 120元

本科定稿，硕博定稿，查重结果与学校一致

立即检测

免费论文查重: <http://www.paperyy.com>

3亿免费文献下载: <http://www.ixueshu.com>

超值论文自动降重: http://www.paperyy.com/reduce_repetition

PPT免费模版下载: <http://ppt.ixueshu.com>

阅读此文的还阅读了:

1. [MPEG-4 AAC实时音频编码算法的优化研究](#)
2. [MPEG音频实时编码优化设计](#)
3. [MPEG-4中的音频编码技术](#)
4. [MPEG-4可扩展编码算法的改进与实现](#)
5. [MPEG-4 AAC实时编码器在DAB+系统中的实现](#)
6. [15.2kb/sLD—CELP语音编码算法及实时实现](#)
7. [一种在MPEG-4编码过程中提高实时性的算法](#)
8. [基于定点DSP的MP3音频编码算法研究及实现](#)
9. [MPEG-2视频编码器的实时实现](#)
10. [VBR MPEG-2视频编码器的实时实现](#)
11. [MPEG-4视频压缩编码的算法优化](#)
12. [MPEG-4AAC音频编码综述](#)
13. [浅谈音频压缩编码技术——MPEG之AAC](#)
14. [MPEG-2音频压缩编码的一种快速算法](#)
15. [MPEG-4视频编码技术的研究及实现](#)
16. [单片ADSP2181实现MPEG音频层 II 实时编码](#)
17. [MPEG-2音频压缩编码的一种快速算法\(英文\)](#)
18. [MPEG-4通用音频编码技术及解码器实现](#)
19. [基于MPEG-4 AAC的音频编码算法研究](#)
20. [一种基于MPEG-2AAC编码的音频水印方法](#)
21. [MPEG编码算法的优化](#)
22. [DVB前端系统中的MPEG音频实时编码](#)
23. [ISO/IEC MPEG—2先进音频编码\(四\)](#)
24. [MPEG-4AAC实时音频编码器设计与实现研究](#)
25. [MPEG—II 实时编码系统中音频编码的实现](#)

[26. MPEG-2混合编码的实时实现](#)

[27. MPEG-4实时编码的Cache算法优化](#)

[28. 一种基于MPEG-4的实时编码系统的设计与实现](#)

[29. MPEG标准中的音频编码与应用](#)

[30. MPEG-4中DCT算法的研究及实现](#)

[31. 高效实时MPEG-2/4 AAC编码器的实现](#)

[32. MPEG-2 AAC与杜比AC-3音频编码](#)

[33. MPEG- II 实时编码系统中音频编码的实现](#)

[34. MPEG-4实时编码器在ADSP-BF533上的实现](#)

[35. MPEG-1音频部分的实现](#)

[36. MPEG-4对象分割与编码算法](#)

[37. 基于DSP的MPEG-2音频实时压缩编码算法](#)

[38. MPEG2音频编码中的辅助数据](#)

[39. MPEG-2AAC与杜比AC-33音频编码](#)

[40. MPEG-2/4 AAC音频编码模块的验证](#)

[41. MPEG-4 ALS无损音频编码](#)

[42. MPEG音频第Ⅲ层编码的算法研究与硬件实现](#)

[43. MPEG音频编码的高效算法研究及单片DSP实现](#)

[44. ISO/IEC MPEG-2先进音频编码\(四\)](#)

[45. MPEG2音频压缩及编码的硬件实现方案](#)

[46. 基于DSP的MPEG1音频第三层编码算法的硬件实现](#)

[47. MPEG 音频编码算法的研究与实时实现](#)

[48. MPEG-4音频编码标准](#)

[49. MP3音频编码算法研究](#)

[50. MPEG-2/4AAC音频编码模块的验证](#)