



## (12) 发明专利

(10) 授权公告号 CN 1525436 B

(45) 授权公告日 2010.05.26

(21) 申请号 03165035.X

CN 1218334 A, 1999.06.02, 全文.

(22) 申请日 2003.09.17

PARK S-H ET AL. "MULTI-LAYER BIT-SLICED BIT-RATE SCALABLE AUDIO CODING". PREPRINTS OF PAPERS PRESENTED AT THE AES103. 1999, 1-16.

(30) 优先权数据

10-2002-0080320 2002.12.16 KR

DOBSON W K ET AL. "High quality low complexity scalable wavelet audio coding". ACOUSTICS, SPEECH, AND SIGNAL. 1997, 327-330.

(73) 专利权人 三星电子株式会社

地址 韩国京畿道

(72) 发明人 金重会 金尚煜 吴殷美

(74) 专利代理机构 北京铭硕知识产权代理有限公司 11286

代理人 韩明星

S-W KIM, R SPERSCHNEIDER, H PURNHAGEN, Y-B TKIM, J HERRE, M DIETZ, S HOTANI, T MORIYA, MNISHIGUCHI, T MLASKO, B GRILL. "ISO/IEC 14496-3 (MPEG-4 Audio) Amd. 1 FPDAM Annex A(informative)". ISO/IEC JTC1/SC29/WG11 N2803. 1999, 1-53.

(51) Int. Cl.

G10L 19/00 (2006.01)

H04B 14/04 (2006.01)

H03M 7/30 (2006.01)

审查员 刘红梅

(56) 对比文件

EP 0918407 A2, 1999.05.26, 全文.

CN 1218334 A, 1999.06.02, 说明书第 29 — 30 页第 1.8.4 节.

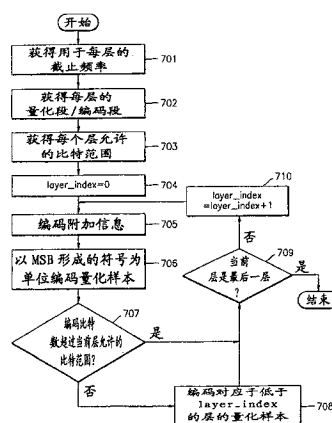
权利要求书 2 页 说明书 11 页 附图 6 页

(54) 发明名称

可伸缩地编解码音频数据的方法和装置

(57) 摘要

提供用于可伸缩地编解码音频数据的方法和装置。所述方法切片音频数据,使得切片的音频数据对应于多个层;获得对应于多个层的每个的定标段信息和编码段信息;基于对应于第一层的定标段信息和编码段信息,编码包含定标因子信息和编码模型信息的附加信息;通过参照定标因子信息量化对应于第一层的音频数据,获得量化样本;通过参照编码模型信息,按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的符号的顺序,以符号为单位编码所获得的多个量化样本;和随着每次逐层增加层的序数,重复执行所述步骤,直到完成多个层的编码。按照该方法,精细粒度可伸缩性 (FGS) 可以被提供,并可以提供具有较低的复杂性和良好的音频质量,即使是在低层。



1. 一种编码方法,包括:

切片音频数据,使得切片的音频数据对应于多个层;

获得定标段信息和编码段信息,定标段信息限定多个定标频段的每个频段的定标因子,编码段信息限定多个编码频段的编码模型,定标段和编码段对应于所述多个层;

基于对应于第一层的定标段信息和编码段信息,编码包含定标因子信息和编码模型信息的附加信息;

通过参照定标因子信息量化对应于第一层的音频数据,获得量化样本;

通过参照编码模型信息,按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的符号的顺序,以符号为单位赫夫曼编码所获得的多个量化样本;和

随着每次逐层增加层的序数,重复执行所述步骤,直到完成多个层的编码。

2. 根据权利要求 1 的方法,进一步包括,在编码附加的信息之前,

获得在多个层的每个中所允许的比特范围,其中在编码所获得的多个量化样本时,编码的比特的数目被计数,并且如果计数的比特的数目超过相应于该比特的比特范围,编码停止,而即使是在量化样本全被编码之后,如果计数的比特的数目小于相应于该比特的比特范围,在低层编码被完成之后仍然未编码的比特被编码到比特范围允许的范围。

3. 根据权利要求 1 的方法,其中音频数据的切片包括:

执行音频数据的子波变换;和

通过参照截止频率,切片子波变换的数据,使得切片的数据对应于多个层。

4. 根据权利要求 1 的方法,其中多个量化样本的编码包括:

在比特平面上映射多个量化样本;和

按照从由 MSB 比特形成的符号直到由 LSB 比特形成的符号的顺序,在对应于样本的层中所允许的比特范围内以符号为单位编码样本。

5. 根据权利要求 4 的方法,其中在映射多个量化样本时,K 个量化样本被映射在比特平面上,并且在样本的编码中,获得对应于由 K 比特二进制数据形成的符号的标量值,并通过参照 K 比特二进制数据,获得的标量值,和对应于高于比特平面上当前符号的一个符号的标量值,执行赫夫曼编码,其中 K 是一个整数。

6. 一种用于可伸缩地解码以分层结构编码的音频数据的方法,包括:

基于对应于第一层的定标段信息和编码段信息差分解码包含定标因子信息和编码模型信息的附加信息;

通过参照编码模型信息,按照从由 MSB 比特形成的符号直到由 LSB 比特形成的符号的顺序,以符号为单位赫夫曼解码音频数据,并获得量化样本;

通过参照定标因子信息反向量化所获得的量化样本;

反向 MDCT 变换该反向量化的样本;和

随着每次逐层增加层的序数,重复执行所述步骤,直到完成预定多个层的解码。

7. 根据权利要求 6 的方法,其中音频数据的赫夫曼解码进一步包括:

按照从由 MSB 比特形成的符号直到由 LSB 比特形成的符号的顺序,在对应于音频数据的层中所允许的比特范围内以符号为单位解码音频数据;和

从其上安排了解码符号的比特平面获得量化样本。

8. 根据权利要求 7 的方法,其中在解码音频数据时,获得由解码符号形成的  $4 \times K$  比特平

面,并且从获得的量化样本中,从  $4 \times K$  比特平面获得  $K$  个量化样本,其中  $K$  是一个整数。

9. 一种用于可伸缩地解码以分层结构编码的音频数据的装置,包括:

解封单元,其基于对应于第一层的定标段信息和编码段信息解码包含定标因子信息和编码模型信息的附加信息,并且通过参照编码模型信息,按照从由 MSB 比特形成的符号直到由 LSB 比特形成的符号的顺序,以符号为单位赫夫曼解码音频数据,并获得量化样本;

反向量化单元,其通过参照定标因子信息反向量化所获得的量化样本;和

反向变换单元,其反向变换该反向量化样本。

10. 根据权利要求 9 的装置,其中解封单元按照从由 MSB 比特形成的符号直到由 LSB 比特形成的符号的顺序,在对应于音频数据的层中所允许的比特范围内以符号为单位解码音频数据,从其上安排了解码符号的比特平面获得量化样本。

11. 根据权利要求 10 的装置,其中解封单元获得由解码符号形成的  $4 \times K$  比特平面,并且从  $4 \times K$  比特平面获得  $K$  个量化样本,其中  $K$  是一个整数。

12. 一种用于可伸缩地编码音频数据的装置,包括:

MDCT 变换音频数据的变换单元;

量化单元,其通过参照定标因子信息,量化对应于每层的 MDCT 变换音频数据,并输出量化样本;和

封装单元,其差分编码包含对应于第一层的定标因子信息和编码模型信息的附加信息,通过参照编码模型信息,按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的符号的顺序,以符号为单位赫夫曼编码来自量化单元的多个量化样本。

13. 根据权利要求 12 的装置,其中封装单元获得定标段信息和编码段信息,并基于对应于每层的定标段信息和编码段信息编码包含定标因子信息和编码模型信息的附加信息,其中,定标段信息限定多个定标频段的每个频段的定标因子,编码段信息限定多个编码频段的编码模型,定标段和编码段对应于所述多个层的每个层。

14. 根据权利要求 12 的装置,其中封装单元计数编码的比特的数目,并且如果计数的比特的数目超过相应于该比特的比特范围,编码停止,而即使是在量化样本全被编码之后,如果计数的比特的数目小于相应于该比特的比特范围,在低层的编码完成之后仍然未编码的比特被编码到比特范围允许的范围。

15. 根据权利要求 12 的装置,其中封装单元参照截止频率切片 MDCT 变换的数据,使得切片的数据对应于多个层。

16. 根据权利要求 12 的装置,其中封装单元在比特平面上映射多个量化样本,并且按照从由 MSB 比特形成的符号直到由 LSB 比特形成的符号的顺序,在对应于样本的层中所允许的比特范围内以符号为单位编码样本。

17. 根据权利要求 16 的装置,其中封装单元在比特平面上映射  $K$  个量化样本,获得对应于由  $K$  比特二进制数据形成的符号的标量值,并通过参照  $K$  比特二进制数据,获得的标量值,和对应于高于比特平面上当前符号的一个符号的标量值执行赫夫曼编码,其中  $K$  是一个整数。

## 可伸缩地编解码音频数据的方法和装置

### 技术领域

[0001] 本发明涉及编码和解码音频数据,尤其特别的是,涉及用于编码音频数据,以便编码的音频比特流具有可伸缩的比特率的方法和装置,以及用于解码音频数据的方法和装置。

### 背景技术

[0002] 由于近来数字信号处理技术的发展,多数情况下音频信号通常被存储成数字信号并且再现。数字音频存储/恢复装置通过采样和量化把音频信号变换成脉冲编码调制(PCM),也就是数字信号。通过这样的操作,数字音频存储/再现装置在信息存储介质,比如光盘(CD)和数字化视频光盘(DVD)中存储PCM音频数据,并响应用户的命令再现存储的信号以便用户能听音频数据。相对于使用密纹(LP)记录或磁带的模拟方法来说,数字存储/恢复方法大大地提高了音频质量,并显著减少了由长的存储周期引起的恶化。然而,由于大量的数字数据,数字方法在存储和传输方面存在问题。

[0003] 为解决该问题,各种压缩方法被用于压缩数字音频信号。

[0004] 在由国际标准化组织标准化的运动图象专家组(MPEG)中,或者由Dolby开发的AC-2/AC-3中,使用音质模型减少了数据量。作为结果,数据量能被有效地减少而不管信号的特性如何。就是说,MPEG/音频标准或AC-2/AV-3方法能以仅仅64~384Kbps的比特率提供几乎与CD相同的音频质量,该比特率是先前数字编码方法的比特率的1/6到1/8。

[0005] 然而,在这些方法中,搜索适用于固定比特率的最佳状态并且接着执行量化和编码。因此,如果在通过网络发送比特流时由于网络条件很差而使得传输带宽被降低,可能会出现断开和适当的服务不能再提供给用户。此外,当比特流期望被变换成较小尺寸的比特流以更适用于具有有限存储容量的移动装置时,应该执行再编码处理以减少比特流的尺寸,并增加了所需的计算量。

[0006] 为解决该问题,本发明的的申请人提出了韩国专利申请 No. 97-61298,1997年11月19日,标题“使用位切片算法编码(BSAC)的可伸缩比特率音频编码/解码的方法和装置”,此专利在2000年4月17日被授权,韩国专利号 No. 261253。根据BSAC技术,具有高比特率编码的比特流能被变成具有低比特率的比特流,并能够只用部分的比特流进行恢复。因此,当网络过载时,或者解码器的性能很差时,或用户请求低比特率时,通过只使用部分的比特流,可以把具有一定音频质量的服务提供给用户,尽管随着比特率的下降,质量会不可避免地成比例地下降。

[0007] 然而,由于BSAC技术采用算术编码,复杂性很高,并且当BSAC技术在实际的装置中被实现时,成本增加。此外,由于BSAC技术利用修正离散余弦变换(MDCT)来进行音频信号的变换,低层中的音频质量会被严重的损坏。

### 发明内容

[0008] 发明提供了用于可伸缩地(with scalability)编解码音频数据的一种方法和装

置,通过该方法和装置为精细粒度可伸缩性 (FGS) 提供较低的复杂度。

[0009] 根据本发明的另一个方面,提供了一种编码方法,包括切片 (slicing) 音频数据以使切片音频数据对应于多个层,获得对应于多个层的每个的定标段信息和编码段信息,基于对应于第一层的定标段信息和编码段信息编码包含定标因子信息和编码模型信息的附加信息,通过参照定标因子信息量化对应于第一层的音频数据,获得量化样本,通过参照编码模型信息,以从利用最高有效位 (MSB) 形成的符号直到利用最低有效位 (LSB) 形成的符号的方式,按顺序以符号为单位编码所得到的多个量化样本,以及随着每次逐层增加层的序数,重复执行步骤,直到完成预定的多个层的编码。

[0010] 在编码附加信息之前,该方法可以进一步包括,获得多个层的每个中所允许的比特范围,其中在获得的多个量化样本的编码中,编码比特的数目被计数,和如果计数的比特的数目超过相应于比特的比特范围,编码停止,以及即使在量化样本全被编码之后,如果计数的比特的数目小于相应于比特的比特范围,在低层编码被完成之后仍然未编码的比特被编码到比特范围允许的范围。

[0011] 音频数据的切片包括执行音频数据的子波变换,并且通过参照截止频率切片子波变换的数据,使得切片的数据对应于多个层。

[0012] 多个量化样本的编码包括在比特平面上映射多个量化样本,并按照从利用最高有效位 (MSB) 形成的符号直到利用最低有效位 (LSB) 形成的符号的顺序,在对应于样本的一个层中允许的比特范围内以符号为单位编码样本。在多个量化样本的映射中, K 个量化样本被映射在比特平面上,并且在样本的编码中,获得对应于由 K 比特二进制数据形成的符号的标量值,并通过参照 K 比特二进制数据,获得的标量值,和对应于高于比特平面上当前符号的一个符号的标量值执行赫夫曼编码 (Huffman coding),其中 K 是一个整数。

[0013] 根据本发明的另一个方面,提供了一种编码方法,包括差分解码包含对应于第一层的定标因子信息和编码模型信息的附加信息,按照从由最高有效位 (MSB) 形成的符号到由最低有效位 (LSB) 形成的符号的顺序,以符号为单位赫夫曼解码音频数据,并通过参照编码模型信息获得量化样本,通过参照定标因子信息反向量化所获得的量化样本,反向 MDCT 变换该反向量化样本,以及随着每次逐层增加层的序数,重复执行步骤,直到完成预定的多个层的解码。

[0014] 附加信息的赫夫曼解码包括按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的符号的顺序,在对应于音频数据的层中所允许的比特范围内以符号为单位解码音频数据,以及从其上安排了解码的符号的比特平面获得量化样本。

[0015] 在解码音频数据时,获得由解码符号形成的  $4 \times K$  比特平面,并且在获得量化样本时,从  $4 \times K$  比特平面获得 K 个量化样本,其中 K 是一个整数。

[0016] 根据本发明的另一个方面,提供了一种装置,用于可伸缩地解码在分层结构中被编码的音频数据,包括解封单元,其解码包含对应于第一层的定标因子信息和编码模型信息的附加信息,并且通过参照编码模型信息,按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的符号的顺序以符号为单位解码音频数据和获得量化的样本;反向量化单元,其通过参照定标因子信息反向量化所获得的量化样本;和反向变换单元,其反向变换该反向量化样本。

[0017] 解封单元按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的

符号的顺序,在对应于音频数据的层中所允许的比特范围内以符号为单位解码音频数据,并且从其上安排了解码的符号的比特平面获得量化样本。

[0018] 解封单元获得由解码的符号形成的  $4 \times K$  比特平面,并且接着从  $4 \times K$  比特平面获得  $K$  个量化样本,其中  $K$  是一个整数。

[0019] 根据本发明的另一个方面,提供了一种装置,用于可伸缩地解码音频数据,包括 MDCT 变换音频数据的变换单元;量化单元,其通过参照定标因子信息量化对应于每层的变换音频数据,并输出量化样本;和封装单元,其差分编码包含对应于每层的定标因子信息和编码模型信息的附加信息,并且通过参照编码模型信息,按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的符号的顺序,以符号为单位赫夫曼编码来自量化单元的多个量化样本。

[0020] 封装单元获得对应于多个层的每个的定标段信息和编码段信息,并基于对应于每层的定标段信息和编码段信息编码包含定标因子信息和编码模型信息的附加信息。

[0021] 封装单元计数编码的比特的数目,如果计数的比特数超过相应于比特的比特范围,编码停止,并且即使在量化样本全被编码之后,如果计数的比特数小于相应于比特的比特范围,将低层编码完成之后仍然未编码的比特编码到比特范围允许的范围。

[0022] 封装单元通过参照一个截止频率切片 MDCT 变换数据,使得切片的数据对应于多个层。

[0023] 封装单元在比特平面上映射多个量化样本,并按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的符号的顺序,在对应于符号的层中所允许的比特范围内以符号为单位解码符号。

[0024] 封装单元在比特平面上映射  $K$  个量化样本,获得对应于由  $K$  比特二进制数据形成的符号的标量值,并通过参照  $K$  比特二进制数据,获得的标量值,和对应于高于比特平面上当前符号的一个符号的标量值执行赫夫曼编码,其中  $K$  是一个整数。

## 附图说明

[0025] 通过结合参考附图详细描述本发明的优选实施例,本发明的上述目的和优点将变得更加清楚,其中:

[0026] 图 1 是根据本发明的一个优选实施例的编码装置的方框图;

[0027] 图 2 是根据本发明的优选实施例的解码装置的方框图;

[0028] 图 3 是帧的结构图,所述的帧形成在分层结构中编码的比特流以便能够控制比特率;

[0029] 图 4 是附加信息的结构的详细图;

[0030] 图 5 是参考图,用以解释按照本发明的一种编码方法;

[0031] 图 6 是参考图,用以更加具体地解释按照本发明的编码方法;

[0032] 图 7 是流程图,用于解释按照本发明优选实施例的编码方法;

[0033] 图 8 是流程图,用于解释按照本发明的优选实施例的解码方法;而

[0034] 图 9 是流程图,用于解释按照本发明的另一个优选实施例的解码方法。

## 具体实施方式

[0035] 参考图 1,按照本发明,编码装置以分层结构编码音频数据,以便能控制编码的比

特流的比特率,并且包括变换单元 11,音质单元 12,量化单元 13,和比特封装单元 14。

[0036] 变换单元 11 接收作为时域音频信号的脉冲编码调制 (PCM) 音频数据,并把信号转换成频域信号,其中参照由音质单元 12 提供的有关音质模型的信息。当人能感知的音频信号的特性之间的差在时域中不是很大时,在通过变换获得的频域音频信号中,人能感知的信号和不能被人所感知的信号的特性之间具有大的差别。所以,通过差分分配到各个频段的比特的数目,压缩效率可以被提高。在本发明实施例中,变换单元 11 执行修正离散余弦变换 (MDCT)。

[0037] 音质单元 12 提供音质模型的信息,比如冲击感信息给变换单元 11,并把变换单元 11 变换的音频信号组合成适当子频段的信号。此外,音质单元 12 通过使用各个信号之间的交互作用所引起的屏蔽效应计算每个子频段中的屏蔽门限,并提供该门限值给量化单元 13。屏蔽门限是由于信号间的交互作用而不能被人所感觉到的信号的最大值。在本实施例中,音质单元 12 通过两耳屏蔽电平降低 (binaural masking level depression) (BMLD) 来计算立体声分量的屏蔽门限。

[0038] 量化单元 13 根据相应于音频信号的定标因子信息在每个频段标量量化音频信号,使得频段中量化噪声的水平小于音质单元 12 所提供的屏蔽门限,以致人不能感知到噪声。接着,量化单元 13 输出量化的样本。就是说,通过使用音质单元 12 中计算的屏蔽门限和每个频段产生的作为噪声比率的噪声-屏蔽比率 (NMR),量化单元 13 执行量化,使得全频段中的 NMR 值是 0dB 或更小。0dB 或更小的 NMR 值意味着人不能感知量化噪声。

[0039] 比特封装单元 14 编码属于每层的量化样本和附加信息,并以分层结构封装编码信号。附加信息包括每层中的定标段信息,编码段信息,它们的定标因子信息,和编码模型信息。定标段信息和编码段信息可以被封装成首部信息,并且接着被发送到解码装置。否则,定标段信息和编码段信息可以被编码和封装成每层的附加信息,并接着发送到解码装置。定标段信息和编码段信息可以不被发送到解码装置,因为在一些情况下它们被预存在解码装置中。

[0040] 更为特别的是,当编码包含对应于第一层的定标因子信息和编码模型信息的附加信息时,比特封装单元 14 参照对应于第一层的编码模型信息,按照从由最高有效位 (MSB) 形成的符号直到由最低有效位 (LSB) 形成的符号的顺序,以符号为单位执行样本和信息的编码。接着,在第二层中,相同的处理被重复执行。就是说,随着层数的增加而执行编码,直到多个预定层的编码被完成。在本实施例中,比特封装单元 14 差分编码定标因子信息和编码模型信息,并赫夫曼编码量化样本。后面将解释根据本发明编码的比特流的分层结构。

[0041] 定标段信息是指用于按照音频信号的频率特性更合适地执行量化的信息。当频率区域被分成多个频段并且一个合适的定标因子被分配到每个频段时,定标段信息指示相应于每层的定标段。这样,每层属于至少一个定标段。每个定标段具有一个分配的定标因子。此外,编码段信息是指用于根据音频信号的频率特性更合适地执行编码的信息。当频率区域被分成多个频段并且适当的编码模型被分配到每个频段时,编码段信息指示对应于每层的编码段。定标段和编码段以经验为主进行划分,并且分别与之对应的定标因子和编码模型基于相同的方式被确定。

[0042] 图 2 是按照本发明的优选实施例的解码装置的方框图。

[0043] 参考图 2,解码装置解码比特流到由网络条件,解码装置的性能和用户的选择所确

定的目标层,使得比特流的比特率能被控制。解码装置包括解封单元 21,反向量化单元 22,和反向变换单元 23。

[0044] 解封单元 21 解封比特流到目标层,并解码每层中的比特流。就是说,包含相应于每层的定标因子信息和编码模型信息的附加信息被解码,并接着基于获得的编码模型信息,属于该层的编码量化样本被解码,并且量化样本被恢复。在本实施例中,解封单元 21 差分解码定标因子信息和编码模型信息,并赫夫曼解码所编码的量化样本。

[0045] 同时,从比特流的首部信息,或通过解码每层中的附加信息,获得定标段信息和编码段信息。可替换的,解码装置可以提前存储定标段信息和编码段信息。按照相应于样本的定标因子信息,反向量化单元 22 反向量化和恢复每层中的量化样本。反向变换单元 23 频率/时间映射所恢复的样本,以便把样本变换成时域的 PCM 音频数据,并输出它。在当前实施例中,反向变换单元 23 执行基于 MDCT 的反向变换。

[0046] 图 3 是帧的结构图,所述帧形成以分层结构编码的比特流,使得可以控制比特率。

[0047] 参考图 3,按照本发明的比特流的帧通过映射量化样本和附加信息被编码到分层结构,以获得精细粒度可伸缩性 (FGS)。换句话说,低层比特流被包括在分层结构的增强层比特流中。每层中需要的附加信息被分配到每层并接着被编码。

[0048] 用于存储首部信息的首部区域被放在比特流的前面,然后有关层 0 的信息在首部区域之后被封装,接着,属于作为增强层的层 1-N 的信息按顺序被封装。从首部区域至层 0 信息的层被称作基层,从首部区域至层 1 信息的层被称作层 1,和从首部区域至层 2 信息的层被称作层 2。同样,最上层表示从首部区域至层 N 信息,就是说,从基层到作为增强层的层 N。附加信息和编码音频数据被存储成每个层信息。例如,附加信息 2 和编码量化样本被存储成层 2 信息。这里,N 是大于或等于 1 的一个整数。

[0049] 图 4 是附加信息的结构的详细图。

[0050] 参考图 4,附加信息和编码量化样本被存储成任意的附加信息,并在本实施例中,附加信息包括赫夫曼编码模型信息,量化因子信息,有关信道的附加信息,和其它附加信息。赫夫曼编码模型信息是赫夫曼编码模型的索引信息,应该被用于编码或解码属于相应于该信息的层的量化样本。量化因子信息指示量化步长,该步长用于量化或反向量化属于相应于信息的层的音频数据。有关信道的附加信息是有关信道的诸如 M/S 立体声的信息。其它附加信息是有关是否使用 M/S 立体声的标志信息。

[0051] 在本实施例中,比特封装单元 14 对赫夫曼编码模型信息和量化因子信息执行差分编码。在差分编码中,一个直接在先频段的值的差分值被编码。有关信道的附加信息被赫夫曼编码。

[0052] 图 5 是参考图,用于更具体地解释根据本发明的编码方法。

[0053] 参考图 5,要编码的量化样本具有 3- 层结构。斜线矩形表示包括量化样本的频谱线,实线表示定标段,虚线表示编码段。定标段 (1), (2), (3), (4) 和 (5) 以及编码段 (1), (2), (3), (4) 和 (5) 属于层 0。定标段 (5) 和 (6) 以及编码段 (6), (7), (8), (9) 和 (10) 属于层 1。定标段 (6) 和 (7) 以及编码段 (11), (12), (13), (14) 和 (15) 属于层 2。同时,定义层 0,使得执行编码直到频段 (a),定义层 1,使得执行编码直到频段 (b),并且定义层 2,使得执行编码直到频段 (c)。

[0054] 首先,使用相应编码模型在 100 的比特范围内编码属于层 0 的量化样本。此外,作

为层 0 的附加信息,属于层 0 的定标段 (1), (2), (3), (4) 和 (5) 以及编码段 (1), (2), (3), (4) 和 (5) 被编码。在以符号为单位编码量化样本时,比特的数目被计数。如果计数的比特数超过允许的比特范围,层 0 的编码被停止,并且层 1 被算术编码。在属于层 0 的量化样本中,当层 0 和 1 中的允许比特数仍然有空间时,未编码的量化样本下一次被编码。

[0055] 接下来,属于层 1 的量化样本被编码,其中使用属于层 1 的编码段,就是说,编码段 (6), (7), (8), (9) 和 (10) 中要编码的量化样本所属的一个编码段的编码模型。此外,作为层 1 的附加的信息,属于层 1 的定标段 (5) 和 (6) 以及编码段 (6), (7), (8), (9) 和 (10) 被编码。甚至在编码相应于层 1 的所有样本之后,如果在允许的比特范围,即 100 比特中仍然有空间,层 0 中剩余的未编码比特被编码,直到允许比特,即 100 比特被计数到。如果针对编码而计数的比特数超过允许比特范围,层 1 的编码被停止,并且层 2 的编码开始。

[0056] 最后,属于层 2 的量化样本被编码,其中使用属于层 2 的编码段,即编码段 (11), (12), (13), (14) 和 (15) 中要编码的量化样本所属的一个编码段的编码模型。此外,作为层 2 的附加信息,属于层 2 的定标段 (6) 和 (7) 以及编码段 (11), (12), (13), (14) 和 (15) 被编码。甚至在编码相应于层 2 的所有样本之后,如果在允许的比特范围,即 100 比特中仍然有空间,层 0 中剩余的未编码比特被编码,直到允许的比特,即 100 比特被计数到。

[0057] 如果所有的量化样本被编码而不考虑层 0 的允许比特范围,即如果所有的量化样本被编码,甚至在编码比特数超过允许比特范围,即 100 之后(这意味着下一层,即层 1 的允许比特范围中的一些比特被用于编码当前层),通常的情况是,属于层 1 的量化样本不能被编码。因此,在可伸缩解码的情况下,如果在范围到层 1 的层上执行解码,由于相应于层 1 的范围到预定频段 (b) 的所有量化样本不被编码,在低于 (b) 的频率上解码的量化样本会波动,导致“Birdy”效应,其中音频质量会恶化。

[0058] 在确定多个层中(目标层)时,分配比特范围,其中考虑到被编码的所有音频数据的整体大小。这样,没有可能因为其中安排被编码的比特的比特范围的缺陷而不执行编码。

[0059] 在以和编码处理相反的方式执行解码时,按照允许的比特范围计数比特数。因此,预定层的解码定时点能被识别。

[0060] 图 6 是参考图,用以更具体地解释按照本发明的编码方法。

[0061] 按照本发明,比特封装单元 14 通过比特明码 (bit-plain) 编码和赫夫曼编码在对相应于每个层的量化样本执行编码。多个量化样本被映射在比特平面上,以便接着以二进制形式表示,并在每层的允许的比特范围内,按照从由 MSB 形成的符号直到由 LSB 形成的符号的顺序被编码。比特平面上重要的信息首先被编码,而相对不太重要的信息随后被编码。通过这样操作,相应于每层的比特率和频段在编码处理中被固定,使得能够减少被称作“Birdy 效应”的失真。

[0062] 图 6 示例了在此情况下的编码例子,其中包括 MSB 的符号的比特数是 4 或更少。当量化样本 9, 2, 4, 和 0 被映射在比特平面上时,它们以二进制形式被表示,也就是,分别为 1001b, 0010b, 0100b 和 0000b。就是说,在本实施例中,作为比特平面上的编码单元的编码块的大小是 4\*4。

[0063] 由 MSB 形成的符号 msb 是“1001b”,由下一 MSB 形成的符号 msb-1 是“0010b”,由下一 MSB 形成的符号 msb-2 是“0100b”,由 LSB 形成的符号 msb-3 是“1000b”。

[0064] 用于赫夫曼编码的赫夫曼模型信息,即码本索引被示于表 1:

[0065] 表 1  
[0066]

| 附加信息 | 有效性 | 赫夫曼模型 |
|------|-----|-------|
| 0    | 0   | 0     |
| 1    | 1   | 1     |
| 2    | 1   | 2     |
| 3    | 2   | 3     |
|      |     | 4     |
| 4    | 2   | 5     |
|      |     | 6     |
| 5    | 3   | 7     |
|      |     | 8     |
|      |     | 9     |
| 6    | 3   | 10    |
|      |     | 11    |
|      |     | 12    |
| 7    | 4   | 13    |
|      |     | 14    |
|      |     | 15    |
|      |     | 16    |
| 8    | 4   | 17    |
|      |     | 18    |
|      |     | 19    |
|      |     | 20    |

| 附加信息 | 有效性 | 赫夫曼模型 |
|------|-----|-------|
| 9    | 5   | *     |
| 10   | 6   | *     |
| 11   | 7   | *     |
| 12   | 8   | *     |
| 13   | 9   | *     |
| 14   | 10  | *     |
| 15   | 11  | *     |
| 16   | 12  | *     |
| 17   | 13  | *     |
| 18   | 14  | *     |
| *    | *   | *     |

[0067] 根据表 1, 对于相同有效水平 (本实施例中的 msb) 甚至存在两个模型。这是因为两个模型是针对显示出不同分布的量化样本而产生的。

[0068] 现在将更详细地解释按照表 1 的图 6 例子的用于编码的处理过程。

[0069] 在一个符号的比特数是 4 或更小的情况下, 按照本发明的赫夫曼编码如公式 1 所示:

[0070] 赫夫曼码值 = 赫夫曼码本 [ 码本索引 ] [ 更高比特平面 ] [ 符号 ] ..... (1)

[0071] 就是说, 赫夫曼编码使用 3 个输入变量, 包括码本索引, 更高比特平面, 和符号。码本索引表示从表 1 获得的值, 更高比特平面表示比特平面上紧临当前期望编码的符号之上的符号。符号表示目前期望编码的符号。

[0072] 由于在图 6 的例子中赫夫曼模型的 msb 是 4, 选择 13-16 或 17-20。如果被编码的附加信息是 8,

[0073] 由 msb 比特形成的符号的码本索引是 16,

[0074] 由 msb-1 比特形成的符号的码本索引是 15,

[0075] 由 msb-2 比特形成的符号的码本索引是 14, 并且

[0076] 由 msb-3 比特形成的符号的码本索引是 13。

[0077] 同时, 由于收 msb 比特形成的符号不具有更高比特平面的数据, 如果更高比特平面的值是 0, 用赫夫曼码本 [16] [0b] [1000b] 的码执行编码。由于由 msb-1 比特形成的符号的更高比特平面是 1000b, 用赫夫曼码本 [15] [1000b] [0010b] 的码执行编码。由于由 msb-2 比特形成的符号的更高比特平面是 0010b, 用赫夫曼码本 [14] [0010b] [0100b] 的码执行编码。由于由 msb-3 比特形成的符号的更高比特平面是 0100b, 用赫夫曼码本 [13] [0100b]

[1000b] 的码执行编码。

[0078] 比特封装单元 14 计数编码比特的数目,用层中允许使用的比特的数目比较该计数,如果计数大于允许数目,停止编码。当在下一层中允许空间时,没有被编码的剩余比特被编码并且被放进下一层。在分配到相应层的量化样本被全部编码之后,如果层中允许的比特的数目中仍然有空间,即如果层中有空间,则低层中编码完成之后仍然未编码的量化样本被编码。

[0079] 同时,如果由 msb 形成的符号的比特数大于或等于 5,使用当前比特平面上的位置确定赫夫曼码值。换句话说,如果有效性大于或等于 5,每个比特平面上的数据中只具有较小统计差,使用相同的赫夫曼模型对数据进行赫夫曼编码。就是说,每个比特平面均存在赫夫曼模式。

[0080] 如果有效性大于或等于 5,就是说,符号的比特数大于或等于 5,本发明的赫夫曼编码满足公式 2:

[0081] 赫夫曼码 =  $20 + \text{bpl} \times 2$

[0082] 其中 ‘bpl’ 表示期望当前被编码的比特平面的索引,并且是大于或等于 1 的整数。如表 2 所列,常数 20 是增加的一个值,用于表示索引从 21 开始,因为赫夫曼模型的最后索引(对应于附加数 8)是 20。因此,用于一个编码段的附加信息简单地表示了有效性。在表 2 中,按照期望被当前编码的比特平面的索引确定赫夫曼模型。

[0083] 表 2

[0084]

| 附加信息 | 有效性 | 赫夫曼模型 |
|------|-----|-------|
| 9    | 5   | 21-25 |
| 10   | 6   | 21-26 |
| 11   | 7   | 21-27 |
| 12   | 8   | 21-28 |
| 13   | 9   | 21-29 |
| 14   | 10  | 21-30 |
| 15   | 11  | 21-31 |
| 16   | 12  | 21-32 |
| 17   | 13  | 21-33 |
| 18   | 14  | 21-34 |
| 19   | 15  | 21-35 |

[0085] 对于附加信息中的量化因子信息和赫夫曼模型信息,在相应于信息的编码段上执行 DPCM。当量化因子信息被编码时,在帧的首部信息中用 8 个比特表示初始的 DPCM 值。用于赫夫曼模型信息的 DPCM 的初始值被设置为 0。

[0086] 下面列出根据本发明和现有技术的 BSAC 技术的编码方法之间的差别。首先,在 BSAC 技术中,编码以比特为单位来执行,而在本发明中是以符号为单位来执行编码。第二,在 BSAC 技术中,使用了算术编码,而在本发明中使用赫夫曼编码。算术编码提供了较高的压缩增益,但增加了复杂性和成本。因此,在本发明中,数据不是以比特单位被编码,而是以符号为单位通过赫夫曼被编码,以便降低复杂性和成本。

[0087] 为了控制比特率,就是说为了提供可伸缩性,相应于一个帧的比特流被截止,考虑到每层中允许使用的比特的数目,使得只利用较小的数据量但可进行解码。例如,如果只有对应于 48kbps 的比特流期望被解码,只使用比特流的 1048 比特,使得能够获得对应于 48kbps 的解码音频数据。

[0088] 现在将解释基于上述结构的按照本发明的编码和解码方法。

[0089] 编码装置读取 PCM 音频数据,在存储器中存储数据(未显示),并且通过音质建模从存储的 PCM 音频数据中获得屏蔽门限和附加信息。由于 PCM 音频数据是时域信号,PCM 音频数据被子波变换成频域信号。接着,编码装置根据量化段信息和量化因子信息通过量子波变换的信号来获得量化样本。如上所述,编码量化样本并通过比特切片编码,基于符号单位的编码和赫夫曼编码来封装。

[0090] 图 7 是流程图,用于解释按照本发明优选实施例的编码方法。

[0091] 参考图 7,现在将解释编码装置的比特封装单元 14 编码和封装量化的样本的处理过程。

[0092] 首先,比特封装单元 14 根据所提供的目标比特率和附加信息提取相应于每层的信息。该处理在步骤 701 至 703 中执行。更具体的,在步骤 701 获得作为用于每层的截止的基础的截止频率,在步骤 702 获得对应于每层的量化段信息和编码段信息,并且在步骤 703 分配比特范围,在该范围内,应当编码的比特在每个层中能够被编码。

[0093] 接着,在步骤 704 中,层索引被确定为基层,并且附加信息(包括量化段信息和编码段信息)在步骤 705 被编码。

[0094] 接下来,相应于基层的量化样本被映射在比特平面上,并在步骤 706 根据由 msb 比特形成的符号以  $4 \times 4$  块为单位进行编码。在步骤 707,编码的比特数被计数,并且如果该计数超过当前层的比特范围,则当前层的编码被停止,并且在下一层开始编码。如果在步骤 707 计数的比特数没有超出比特范围,则在步骤 709,过程返回到步骤 705 以处理下一层。由于基层没有更低的层,步骤 708 不被执行,但针对基层之后跟着的层执行步骤 708。通过上述步骤,直到目标层的所有范围的层均被编码。

[0095] 步骤 706,也就是用于编码量化的样本的步骤如下所述:

[0096] 1. 相应于一个层的量化样本被以 N 样本为单位分组和映射在比特平面上。

[0097] 2. 根据由映射的二进制数据的 msb 比特形成的符号执行赫夫曼编码。

[0098] 子步骤 2 可以如下进行详细解释:

[0099] 2.1 相应于期望编码的符号的标量值(curVal) 被获得。

[0100] 2.2 相应于标量值(upperVal) 的赫夫曼代码被获得,该标量值对应于更高比特平

面中的符号,就是说,比特流中处于比期望当前编码的符号更高的位置的符号。

[0101] 对于附加信息中的量化因子信息和赫夫曼模型信息,在相应于信息的编码段上执行 DPCM。当量化因子信息被编码时,DPCM 的初始值在帧的首部信息中由 8 个比特表示。用于赫夫曼模型信息的 DPCM 的初始值被设置到 0。

[0102] 图 8 是流程图,用于解释按照本发明的优选实施例的解码方法。

[0103] 参考图 8,解码装置接收由在分层结构中编码的音频数据形成的比特流,并解码每帧中的首部信息。接着,在步骤 801,包括相应于第一层的定标因子信息和编码模型信息的附加信息被解码。在步骤 802,参照编码模型信息,通过按照从由 MSB 比特形成的符号直到由 LSB 比特形成的符号的顺序以符号为单位地解码比特流,获得量化样本。在步骤 803,通过参考定标因子信息,获得的量化样本被反向量化,并在步骤 804 中,反向量化样本被反向变换。随着每次逐层增加层的序数,重复执行步骤 801-804,直到完成预定的多个层的编码。

[0104] 图 9 是流程图,用于解释按照本发明另一个优选实施例的解码方法。

[0105] 参考图 9,接收由在分层结构中编码的音频数据形成的比特流,并在步骤 901,根据每帧中的首部信息解码相应于每层的截止频率。在步骤 902,通过解码,根据首部信息识别相应于每层的量化段信息和编码段信息。在步骤 903,每层的允许使用比特范围被识别。在步骤 904,层索引被设置为基层。步骤 905 解码基层上的附加信息,在步骤 906,通过按照从由 MSB 比特形成的符号直到由 LSB 比特形成的符号的顺序以符号单位地将比特流解码为每层允许的比特范围,获得量化样本。在步骤 907,检查当前层是否最后一个。随着层数的逐个增加,步骤 905 和 906 在各层上重复执行,直到到达预定的目标层。在步骤 901-903 中,解码装置可以提前具有截止频率,量化段信息,编码段信息和比特范围,而不是根据存储在接收的比特流的每帧中的首部信息获得这些信息。在此情况下,通过读取存储的信息,解码装置获得信息。

[0106] 如上所述,根据本发明,通过在执行比特切片之后以符号为单位编码比特,提供借以能够通过自顶向下方式控制比特率的可伸缩性,使得编码装置的计算量不太大于没有提供可伸缩性的装置。就是说,根据本发明,提供了一种用于编解码带有可伸缩性的音频数据的方法和装置,其中复杂性较低,同时可以提供 FGS,以及良好的音频质量,即使是在低层。

[0107] 此外,比较于使用算术编码的 MPEG-4 音频 BSAC 技术,使用赫夫曼编码的本发明的编解码装置减少了用于封装/解封处理的计算量,其降至 BSAC 技术的八分之一。即使当按照本发明的比特封装被执行以便提供 FGS 时,开销是小的,使得编码增益与没有提供可伸缩性时相同。

[0108] 此外,由于按照本发明的装置具有分层结构,为使得服务器端能够控制比特率而再产生比特流的处理是很简单的,因此,用于变换编码的装置的复杂性是低的。

[0109] 当通过网络发送音频流时,能根据用户的选择或网络条件控制传输比特率,以便可以提供不停断的服务。

[0110] 此外,当音频流被存储在具有有限容量的信息存储介质中时,文件的大小能被任意控制和被存储。如果比特率变低,频段被约束。因此,是编/解码器中最复杂装置的滤波器的复杂性被大大降低,并且与比特率成反比,编/解码器装置的实际复杂性降低。

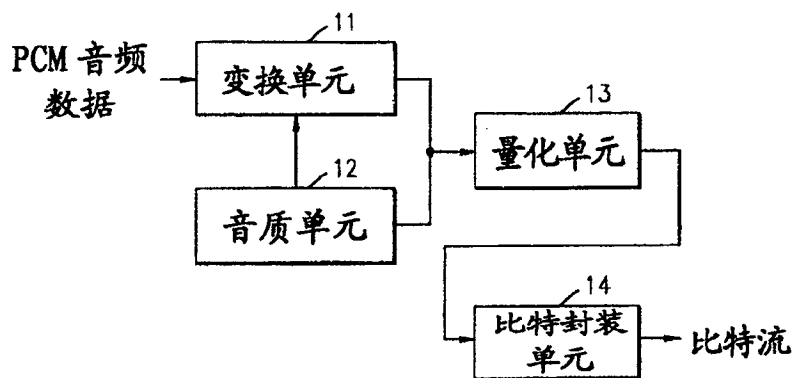


图 1

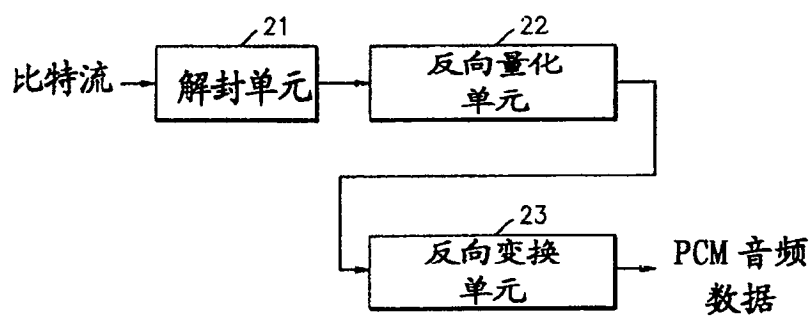


图 2

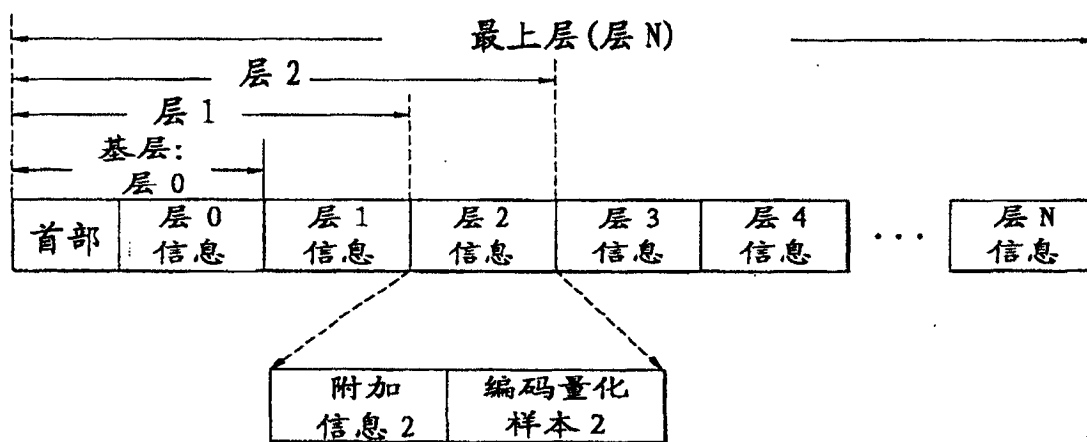


图 3

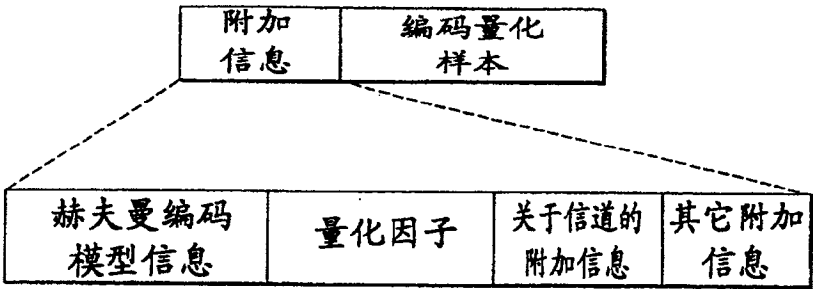


图 4

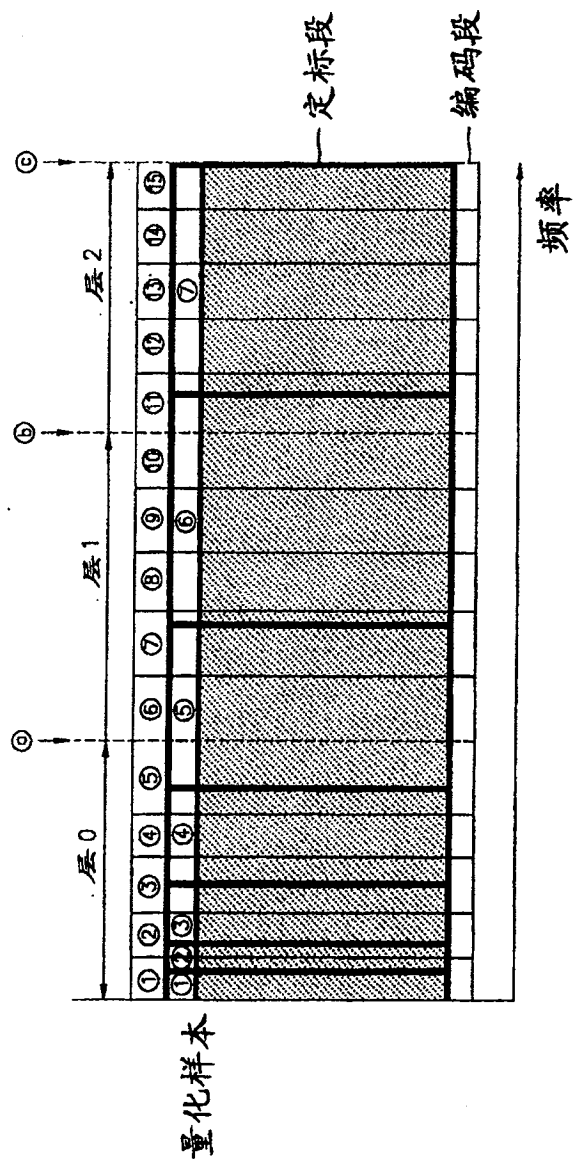


图 5

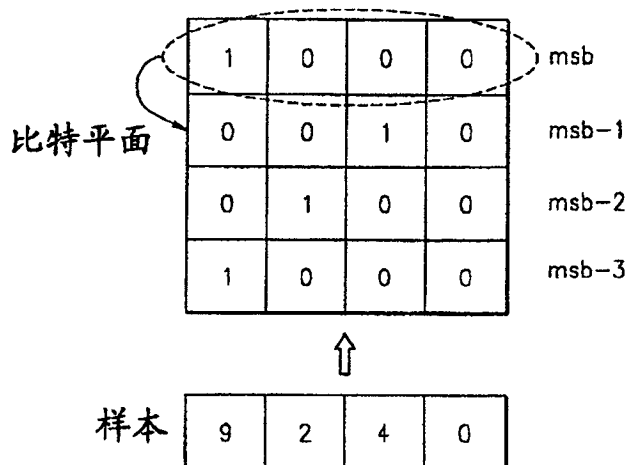


图 6

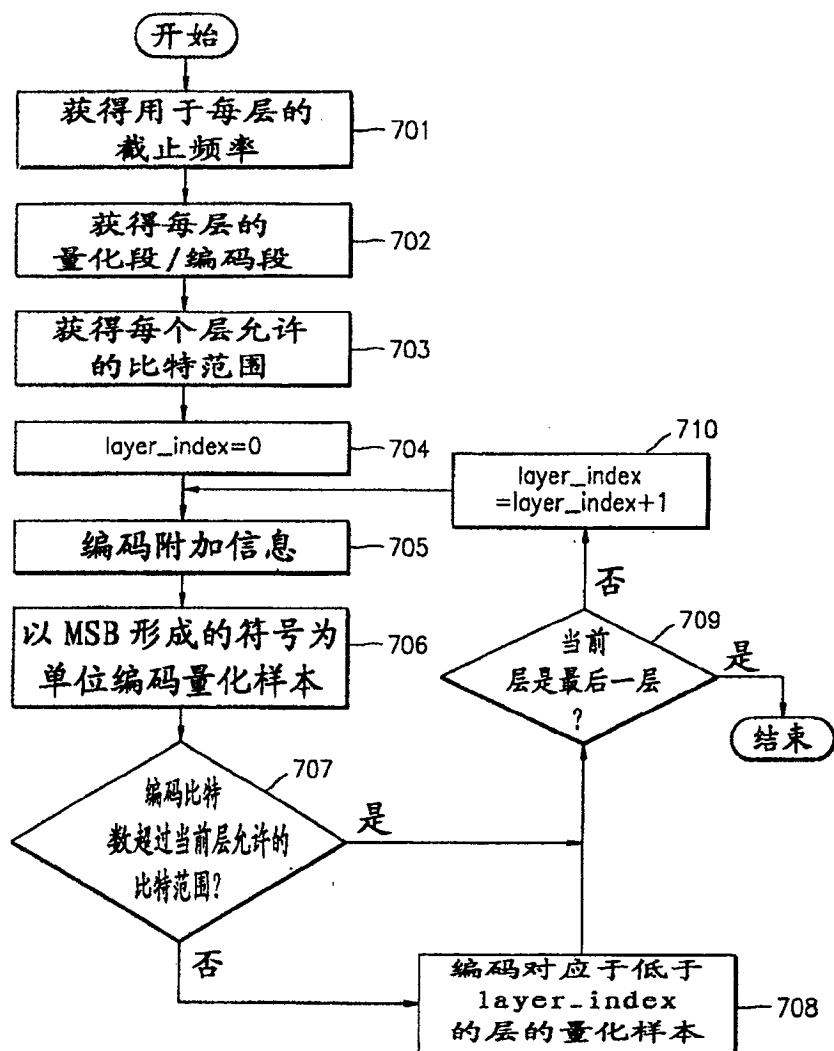


图 7

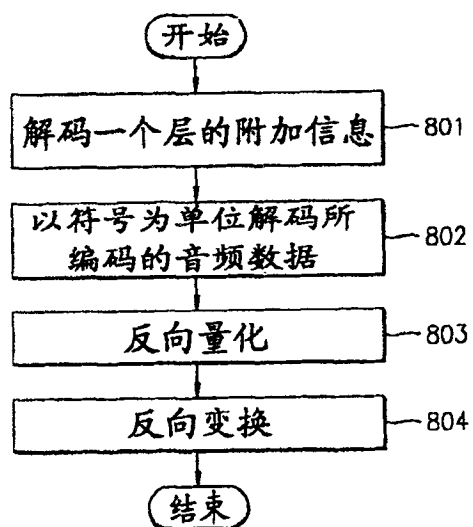


图 8

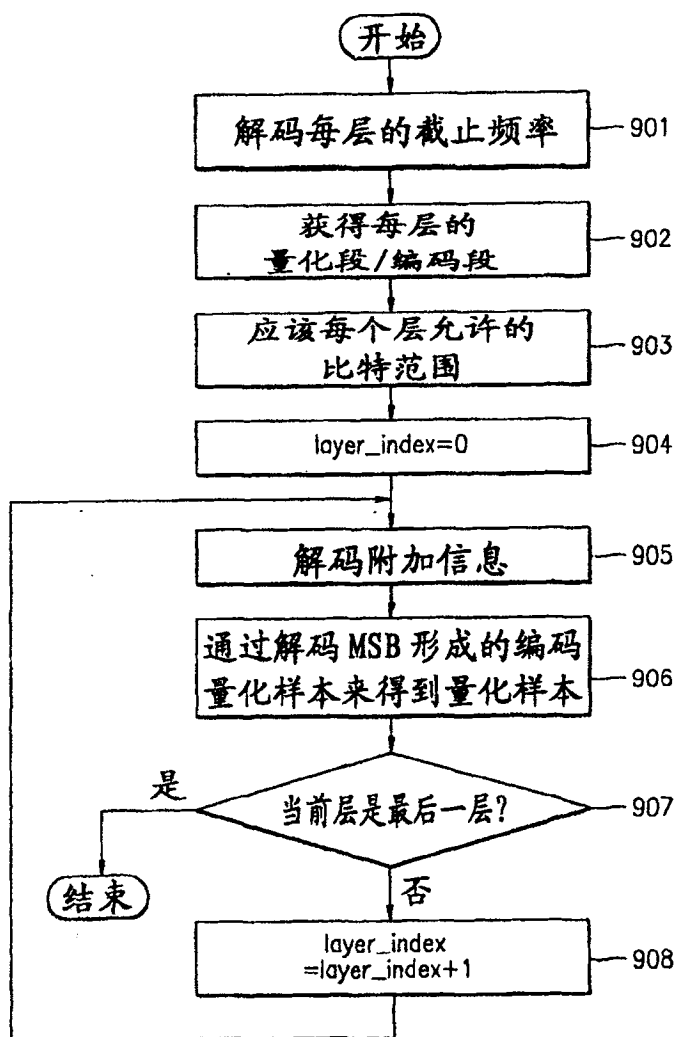


图 9