



(12) 发明专利申请

(10) 申请公布号 CN 102160378 A

(43) 申请公布日 2011. 08. 17

(21) 申请号 200980136191. 1

代理人 吕晓章

(22) 申请日 2009. 07. 14

(51) Int. Cl.

(30) 优先权数据

H04N 7/26 (2006. 01)

08160549. 5 2008. 07. 16 EP

(85) PCT申请进入国家阶段日

2011. 03. 16

(86) PCT申请的申请数据

PCT/EP2009/058944 2009. 07. 14

(87) PCT申请的公布数据

W02010/007037 EN 2010. 01. 21

(71) 申请人 汤姆森特许公司

地址 法国伊西莱穆利诺

(72) 发明人 彼得·贾克斯 斯文·科唐

(74) 专利代理机构 北京市柳沈律师事务所

11105

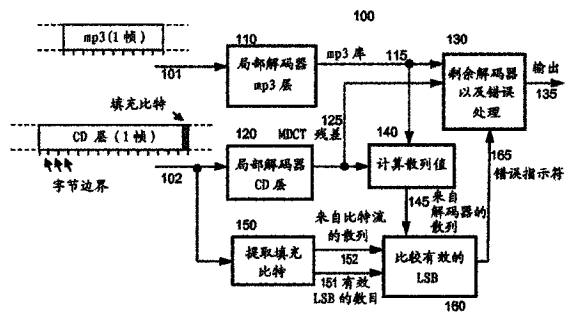
权利要求书 3 页 说明书 10 页 附图 3 页

(54) 发明名称

对被高度压缩的增强层数据进行同步的方法和装置

(57) 摘要

多媒体应用数据格式经常使用数据分组并且可能包含分组化的具有基本层 (BL) 以及一个或多个增强层 (EL) 的分层级的层。包括变长编码的数据的分组典型地提供填充比特, 所述填充比特可以用于添加校验和信息。该校验和信息参考解码的基本层和解码的增强层分组数据这两者, 并且用于检测各层之间的同步丢失。尽管各分组提供不同的单独的填充比特量, 但每个分组的该可变比特量对于添加用于与同步有关的目的的校验和信息是足够的。即使增强层编码没有提供用于同步的开销信息, 具体地, 没有分组首标和分组同步字, 这也允许例如进行失去同步检测。



1. 一种用于编码具有基本层比特流 (215) 和与基本层比特流有关的增强层比特流 (255) 的音频或视频信号的方法, 其中以分组构成基本层数据和增强层数据, 并且其中基本层比特流的分组具有对应的增强层比特流的分组, 所述方法包括以下步骤:

- 计算 (260) 关于基本层比特流的分组和对应的增强层比特流的分组的校验和;
- 熵编码 (220) 基本层比特流的分组, 其中获得以同步字开始的熵编码的字节对齐的基本层分组;
- 使用变长编码方法来熵编码 (270) 增强层比特流的分组, 其中获得不具有首标并且不以同步字开始的熵编码的增强层分组 (273);
- 将编码的增强层分组映射 (271) 到字节, 每个字节具有固定数目的比特, 其中在编码的增强层分组的最后字节中 N 个数目的比特是填充比特;
- 确定 (272) 编码的增强层分组的最后字节内的填充比特量 N;
- 提取 (261) 所述所计算的校验和的 N 个最低有效比特, N 等于所确定的填充比特量;
- 将所提取的所计算的校验和的 N 个最低有效比特插入 (280) 到编码分组的结尾处的填充比特中; 以及
- 提供或输出包括所插入的校验和比特的编码的增强层分组、以及编码的基本层分组。

2. 如权利要求 1 所述的方法, 其中在没有像分组或者帧首标那样的附加结构数据的情况下, 由单一的开头同步字和多个所述编码的增强层分组来形成 (290) 编码的增强层比特流 (295)。

3. 如权利要求 1 或 2 所述的方法, 其中, 在所述熵编码之前, 基本层比特流 (215) 和增强层比特流 (255) 包括整数值, 并且计算校验和的步骤包括在熵编码之前对所述整数值求和 (260)。

4. 如权利要求 1-3 中的任一项所述的方法, 其中使用 Golomb 码来编码增强层比特流的分组。

5. 如权利要求 1-4 中的任一项所述的方法, 其中每个编码的增强层分组的填充比特的数目 N 是可变的, 小于字节的比特数目, 并且在舍入或者舍去的情况下平均为字节的比特数目的一半。

6. 如权利要求 1-5 中的任一项所述的方法, 其中基本层比特流 (215) 包括 MP3 格式化的音频数据, 而增强层比特流 (255) 包括适合于对基本层比特流的 MP3 格式化的音频数据进行增强解码的 MDCT 残差。

7. 如权利要求 1-6 中的任一项所述的方法, 还包括在文件中存储编码的基本层比特流 (225) 和编码的增强层比特流 (295) 的步骤, 其中在该文件的第一部分 (10) 中彼此相邻地存储基本层分组, 而在该文件的随后的第二部分 (11) 中彼此相邻地存储增强层分组。

8. 一种用于解码具有基本层比特流 (101) 和对应的增强层比特流 (102) 的音频或视频信号的方法, 其中增强层比特流包括用于对基本层比特流进行增强解码的变长编码的参数, 所述方法包括以下步骤:

- 局部解码 (110) 基本层比特流的一部分, 其中获得解码的第一基本层分组 (115);
- 局部解码 (120) 增强层比特流的一部分, 其中获得用于对基本层比特流进行增强解码的解码参数 (125);
- 确定 (150) 增强层比特流的所述部分中 N 个数目的填充比特, 以及提取和 / 或存储填

充比特；

- 计算 (140) 关于解码的第一基本层分组 (115) 和来自增强层比特流的局部解码部分的所述解码参数 (125) 的校验和；

- 将所提取的所计算的校验和的 N 个最低有效比特与所提取的 N 个填充比特进行比较 (160), 其中 N 等于所确定的填充比特量；

- 如果两者相等, 则使用用于对基本层比特流进行增强解码的所述解码参数 (125) 来解码 (130) 被局部解码的基本层比特流 (115), 否则, 利用其中不使用来自增强层比特流的数据的简化解码来解码 (130) 被局部解码的基本层比特流 (115)。

9. 如权利要求 8 所述的方法, 其中所述确定 (150) N 个数目的填充比特的步骤包括以下步骤: 确定基本层比特流 (115) 的被局部解码部分中码字的数目, 从变长编码的增强层比特流的部分中解码对应数目的码字 (125), 以及将增强层比特流的该部分的最后字节的剩余比特确定 (150) 为所述 N 个填充比特。

10. 如权利要求 8 或 9 所述的方法, 其中对于每个基本层分组, 增强层比特流包括预定义数目的解码参数。

11. 一种包括第一部分 (10) 和第二部分 (11) 的编码的音频 / 视频信号, 所述第一部分包括基本层音频 / 视频比特流的字节对齐的熵编码的分组 (12), 而第二部分包括包含变长编码的增强层数据的字节对齐的熵编码的增强层分组 (13),

其中, 基本层分组包括具有同步信息 (23) 的首标, 而增强层分组不包括具有同步信息的首标,

以及其中, 每个增强层分组包括适合于对对应的基本层分组进行增强解码的编码参数,

以及其中, 增强层分组的最后字节包括 N 个数目的填充比特, 该填充比特包括关于熵编码之前的增强层分组和对应的基本层分组的校验和的 N 个 LSB。

12. 如权利要求 11 所述的信号, 其中基本层音频 / 视频比特流和增强层比特流包括整数, 并且校验和是所述整数之和。

13. 一种用于编码具有基本层比特流 (215) 和与基本层比特流有关的增强层比特流 (255) 的音频或视频信号的装置, 其中以分组构成基本层数据和增强层数据, 并且其中基本层比特流的分组具有对应的增强层比特流的分组, 所述装置包括:

- 计算部件, 计算 (260) 关于基本层比特流的分组和对应的增强层比特流的分组的校验和;

- 基本层熵编码器 (220), 熵编码基本层比特流的分组, 其中生成以同步字起始的熵编码的字节对齐的基本层分组;

- 增强层熵编码器 (270), 使用变长编码方法来熵编码增强层比特流的分组, 其中生成不具有首标并且不以同步字起始的熵编码的增强层分组 (273);

- 映射部件 (271), 将编码的增强层分组映射到字节, 每个字节具有固定数目的比特, 其中在编码的增强层分组的最后字节中 N 个数目的比特是填充比特;

- 确定部件 (272), 确定编码的增强层分组的最后字节内的填充比特量 N;

- 提取部件 (261), 提取所述所计算的校验和的 N 个最低有效比特, N 等于所确定的填充比特量;

- 插入部件 (280), 将所提取的所计算的校验和的 N 个最低有效比特插入到编码分组的结尾处的填充比特中; 以及

- 用于提供或输出包括所插入的校验和比特的编码的增强层分组、以及编码的基本层分组的部件。

14. 如权利要求 13 所述的装置, 还包括用于在文件中存储编码的基本层比特流 (225) 和编码的增强层比特流 (295) 的部件, 其中, 在该文件的第一部分 (10) 中彼此相邻地存储基本层分组, 而在该文件的随后的第二部分 (11) 中彼此相邻地存储增强层分组。

15. 一种用于解码具有基本层比特流 (101) 和对应的增强层比特流 (102) 的音频或视频信号的装置, 其中增强层比特流包括用于对该基本层比特流进行增强解码的变长编码的参数, 所述装置包括:

- 基本层局部解码器 (110), 局部解码基本层比特流的一部分, 其中生成解码的第一基本层分组 (115);

- 增强层局部解码器 (120), 局部解码增强层比特流的一部分, 其中生成用于对基本层比特流进行增强解码的解码参数 (125);

- 用于确定 (150) 增强层比特流的所述部分中 N 个数目的填充比特的部件, 以及用于提取和 / 或存储填充比特的部件;

- 计算部件 (140), 计算关于解码的第一基本层分组 (115) 和来自增强层比特流的被局部解码部分的所述解码参数 (125) 的校验和的部件;

- 比较部件 (160), 将所提取的所计算的校验和的 N 个最低有效比特与所提取的 N 个填充比特进行比较, 并且如果两者不相等则提供错误指示 (165), 其中 N 等于所确定的填充比特量; 以及

- 解码部件 (130), 如果根据所述错误指示 (165) 所提取的所计算的校验和的 N 个最低有效比特与所提取的 N 个填充比特相等, 则使用用于对基本层比特流进行增强解码的所述解码参数 (125) 来解码被局部解码的基本层比特流 (115), 否则, 利用其中不使用来自增强层比特流的数据的简化解码来解码 (130) 被局部解码的基本层比特流 (115)。

对被高度压缩的增强层数据进行同步的方法和装置

技术领域

[0001] 本发明涉及一种对与基本层数据有关的增强层数据进行高度压缩编码的方法和装置、以及一种对被高度压缩的增强层数据进行对应解码的方法和装置,其中确定增强层数据和对应的基本层数据之间的同步。

背景技术

[0002] 一些多媒体应用数据格式使用具有基本层 (BL) 以及一个或多个增强层 (EL) 的分层级的多层。经常对一层内的数据进行分组化,即,以分组或者帧来组织该数据。尽管 BL 信号单独可以被解码以获得可再现的多媒体数据并且其包括用于基本解码的所有信息,但是 EL 信号包括不能够被单独解码以获得有用的多媒体数据的附加信息。替代地,EL 数据紧密地与 BL 数据耦合,并且仅仅在与其一起使用时是有用的。通常,BL 数据和 EL 数据要么用于共同解码 (common decoding) 要么在它们各自解码之后,被彼此相加或者叠加。在任一种情况下,必须将 EL 数据与 BL 数据进行同步,这是因为如果不这样做则 EL 数据包含无用的信息。

[0003] 例如,常见的是通过单独的 CRC 校验和或者前向纠错 (FEC) 机制来单独地校验每个子比特流的完整性。通常,向子比特流的单独的、分离的块或者信号帧应用这样的 CRC 或者 FEC 方案。如果在子比特流中的同步位置处开始解码,并且此外如果对于任一子比特流均未检测到传送错误,则解码器将认为分层级的比特流处于同步中。然而,因为该技术仅仅给出了与比特流的被单独保护的块的接收有关的确定性 (certainty),所以例如如果一个子比特流的一个受保护的块被完全丢失,则仍然可能存在未检测到的错误。

[0004] 此外,期望将数据率保持得尽可能低,这导致了复杂的数据压缩方法。CRC 或者 FEC 方法的缺点典型地在于:必须对比特流的每个块传送相当量的附加数据。因此,通常将变长编码用于未均等分布的数据字。在变长编码 (VLC) 中,将较频繁(即,较高概率)出现的数据字编码为较短的码字,而将以较低概率出现的数据字编码为较长的码字。由此,编码后的消息中的平均比特量比使用恒定码字长度的情况要短。然而,诸如 VLC 之类的高度压缩方法对比特错误更敏感,这可能导致完全的数据丢失。例如,对于 VLC,当同步已经丢失时,可能不能够确定哪些比特属于一个码字。因此,限制可能的数据丢失的常见的解决方案是插入能够以非常高的概率被辨识的唯一的同步字。然而,同步字增加了数据率,并且使用的同步字越多,则数据率越高。另外,在子比特流的块上应用 CRC 或者 FEC 方法的另一缺点在于:非常难应付像 VBR 编码中具有可变长度并且没有块首标的块。

[0005] 为了使得能够检测数据传送或者存储中的比特错误,常见的方法是计算关于编码数据的校验和,并且还传送或者存储该校验和。对于检测或者对于校正某些比特错误,已知不同类型的校验和。然而,校验和中的信息量越高,则其需要的比特越多。从而,校验和越有用,则其增加的数据率越多。

发明内容

[0006] 鉴于上述内容,本发明要解决的问题是提供一种用于与分组化的基本层数据流有关的高度压缩的增强层数据流的编码方法和格式,使得在解码期间内至少可以检测增强层数据和基本层数据之间的同步丢失,同时保持增强层数据流中的同步(sync)信息和校验和信息量为低。

[0007] 此外,本发明要解决的另一问题是提供一种对应的解码方法和装置。具体地,该解码方法和装置应当在解码期间内至少能够检测增强层数据和基本层数据之间的同步丢失。进行这样的检测,应当至少能够辅助减少干扰(disturbance)。

[0008] 对于上述问题,本发明提供了以下解决方案。

[0009] 在原理上,本发明是基于对以下事实的认识,即,包括变长编码数据的分组典型地在它们的填充比特(padding bit)中提供了空闲空间,该空闲空间可以用于添加校验和信息,并且校验和信息还可以用于检测同步丢失或者其它同步目的。出于此原因,校验和信息参考解码的基本层和解码的增强层分组数据这两者。尽管不是所有的分组都提供空闲的填充比特,但至少大部分分组都提供空闲的填充比特,并且平均每个分组提供 $(n-1)/2$ 个填充比特,假定 n 比特形成一个字节。尽管每个分组的该可变的比特量可能对于高度复杂的纠错是不够的,但已经发现对于上述的与同步有关的目的,其是足够的。有利地,即使增强层编码提供了非常有限的开销(overhead)信息(例如,没有分组首标以及没有分组同步字),这也允许例如在几个分组之后进行失去同步(out-of-sync)检测。

[0010] 一方面,本发明提供了一种编码基本层数据流和有关的增强层数据流的方法,该方法包括以下步骤:计算关于基本层数据流和增强层数据流的共同的校验和;编码具有独立长度(individual length)的基本层分组中的基本层数据,其中分组内的数据字具有恒定长度(例如,面向字节(byte-oriented)的结构)并且每个分组具有至少包括同步字和独立分组长度的指示的首标部分;编码具有独立长度的增强层分组中的增强层数据,其中每个增强层分组对应于特定的基本层分组,并且其中增强层分组内的数据字具有可变长度(即,不是面向字节的结构)并且该分组不包括首标、开头的分组同步字和分组长度的指示,并且其中被编码为分组的增强层数据字的数目依赖于被编码为对应的基本层分组的基本层数据字的数目;将编码的增强层分组映射到面向字节的结构,其中一个字节具有 n 比特并且其中分组的最后字节中的多个填充比特保留为未使用,填充比特的数目在零和 $n-1$ 之间;确定用于增强层分组的填充比特的数目;从与所述增强层分组及其对应的基本层分组有关的校验和中提取一定量的最低有效比特(LSB),所述量与所述所确定的填充比特的数目相等;以及将所提取的校验和 LSB 插入到所述填充比特中。

[0011] 根据所采用的多媒体应用格式,独立基本层分组的长度可以由它们的内容定义。例如,音频编码格式 MP3(MPEG-1 层 III) 或者视频编码格式 AVC 提供了可变的、依赖于内容的分组长度。

[0012] 此外,应注意,分组中的值的数目(即,信息)在熵编码之前和熵编码之后是相同的,然而比特或者字节的数目通常是不同的。

[0013] 在一个实施例中,基本层(BL)数据流的每个 BL 分组具有增强层(EL)比特流中的一个对应 EL 分组,并且 BL 分组中的每个数据值在 EL 分组中具有对应的数据增强值。例如,如果 BL 分组具有特定数目 X_b 个值,则对应的 EL 分组具有相同数目的增强值,对于各 BL 值

中的每一个具有一个增强值。

[0014] 在另一实施例中,其中 BL 数据流的每个 BL 分组具有 EL 比特流的一个对应 EL 分组, BL 分组中的(仅仅一个/多个)特定类型的每个数据值在 EL 分组中具有对应的数据增强值。例如,如果 BL 分组具有第一类型的 X_{B1} 个值和第二类型的 X_{B2} 个值,则对应的 EL 分组可能具有仅仅与第一类型的 BL 值有关的相同数目 X_{B1} 个增强值。

[0015] 在一个实施例中,依序排列(例如,在单个文件内存储)BL 数据和 EL 数据,并且 EL 数据包括不多于一个同步字。

[0016] 根据本发明的另一方面,编码的多媒体信号包括第一部分和第二部分,第一部分包括基本层音频比特流的字节对齐(byte-aligned)的熵编码的分组,而第二部分包括包含变长编码的增强层数据的字节对齐的熵编码的增强层分组,其中基本层分组包括同步信息而增强层分组不包括同步信息(至少没有分组首标),并且其中增强层分组包括适合于对它们的对应基本层分组进行增强解码的编码参数,并且其中增强层分组的最后字节包括独立的可变数目 N 个填充比特(即,不同的增强层分组可能具有不同的独立数目的填充比特)。填充比特包括关于两个分组(即,来自音频信号的第二部分的增强层分组和来自音频信号的第一部分的对应的基本层分组)的校验和的 N 个 LSB。在熵编码之前计算该校验和,即,校验和不参考所接收的数据,而是参考熵解码的数据。

[0017] 根据本发明的又一方面,一种解码基本层比特流和对应的增强层比特流的方法,其中增强层比特流包括用于对基本层比特流进行增强解码的变长编码的增强数据,所述方法包括以下步骤:熵解码基本层比特流的一部分,其中获得熵解码的第一基本层分组;熵解码增强层比特流的一部分,其中获得用于对基本层比特流的一部分进行增强解码的增强数据;确定增强层比特流的最后字节中 N 个数目的填充比特,以及提取和/或存储填充比特;计算关于解码的第一基本层分组和来自增强层比特流的熵解码部分的所述增强数据的校验和;提取所计算的校验和的 N 个最低有效比特(LSB), N 等于所确定的填充比特的数目;将所提取的所计算的校验和的 N 个最低有效比特与所提取的 N 个填充比特进行比较;以及如果两者相等,则使用所述增强数据和基本层比特流数据来解码基本层比特流,否则,利用其中不使用或忽略来自增强层比特流的数据的简化解码来解码基本层比特流。

[0018] 基本层分组的长度(在解码之前或者解码之后)可以在其首标中预定义或者指示,而增强层分组的长度(在比特方面)不进行指示而必须在解码处理期间例如使用来自对应的基本层分组的信息而导出。在一个实施例中,增强层分组中编码值的数目等于对应的基本层分组中预定义类型的值的数目,并且解码包括以下步骤:从基本层分组确定要从增强层分组解码的值的数目,并且从增强层数据中解码所确定的数目的值。在另一实施例中,每个增强层分组的值的数目是预定义的。

[0019] 在从属权利要求、以下描述和附图中公开了本发明的有利实施例。

附图说明

[0020] 参照附图来描述本发明的示例实施例,在附图中:

[0021] 图 1 示出具有基本层和增强层的流格式的结构;

[0022] 图 2 示出分组化的基本层的结构;

[0023] 图 3 示出包括填充比特的分组化的增强层的结构;

- [0024] 图 4 示出示例性的用于失去同步检测的所接收的同步字与所计算的同步字的比较；
- [0025] 图 5 示出生成 MP3 基本层数据和音频增强层数据的音频编码器的框图；
- [0026] 图 6 示出 MP3 基本层数据和音频增强层数据的音频解码器的框图；以及
- [0027] 图 7 示出音频编码器中填充比特替换的详细框图。

具体实施方式

[0028] 图 1 示出包括基本层 10 和增强层 11 的数据流。该增强层在此也被称作扩展层。由于数据流是有限的,所以其也被称为数据组块 (data chunk)。在所描述并描绘的实施例中,扩展层在基本层之后,使得两者可以依序使用相同信道进行传送或存储。在另一实施例中,这两个层可以使用分离的信道,使得它们可以同时被传送、接收、存储或者从存储装置中读取。

[0029] 在该示例中,基本层 10 包含 MP3 音频数据 (如图 2 中所示),而增强层 11 包含用于对音频数据进行增强解码的数据。基本层数据组块被分离为帧 12、21、22。由唯一的同步字 23 指示基本层帧 (例如,MP3 帧) 的起始。对应的解码器 (例如 MP3 解码器) 针对该同步字 23 分解数据流直至它发现 MP3 帧的起始为止,并且解码所发现的 MP3 帧的频谱值 (spectral value)。如果在帧数据内存在比特错误,则解码的数据是错误的并且帧 21 被丢弃 (drop)。然而,通过下一同步字可以发现下一帧 22 并且解码可以继续,使得没有丢失其它的帧。

[0030] 基于用于增强的 MP3 音频数据的格式和解码处理来解释本发明。解码器被称为“hd3”解码器,对应的编码器被称为“hd3”编码器。可替换地,其可以被称作“MP3HD”。所谓的“hd3”比特流格式包括作为基本层 10 的 MP3 格式化层和至少一个附加增强层或扩展层 11。(多个)MP3 帧 12 形成 hd3 文件的一个相干 (coherent) 数据组块 10,扩展层 11 形成之后的相干增强数据组块。在一个实施例中,基本层 10 和扩展层 11 形成完全的音频轨道 (等等)。在另一实施例中,对于超长轨道,相干的基本层数据组块和之后的相干的增强层数据组块可能仅仅形成轨道的一部分,而完全的音频轨道由两对或者更多对数据组块组成。

[0031] 如图 3 所示,也以帧来构成扩展层数据组块 11。然而,为了减少带宽,将其进行高度压缩。因此使用变长编码来编码增强层值,并且扩展层帧不以 sync 字 (同步字) 为起始。替代地,只是依序添加 (append) 各帧。此外,不管字节结构如何,只是依序添加增强层帧内的码字。从而,扩展层的帧边界是不可直接检测的,而是必须根据所定义的解码算法 (其包括每帧解码所定义数目的值) 来发现。

[0032] 由于变长编码,所以比特错误将不仅仅干扰当前值,而是还将致使当前值的长度不可知,使得该帧中所有之后的值的开始不确定。此外,由于缺失帧同步字,常规地识别增强层数据组块内的其它帧将是不可能的。

[0033] 示例性地,GoIomb 码可以用作变长码。GoIomb 码可以用于表示有限范围的非负整数,这对于例如 MP3 数据和 MP3 增强数据是足够的。原理上,GoIomb 码还可以用于其它多媒体数据。通过商和除法余数来构建码字,并且商处于一元表示而余数处于二元表示,且两者参考同一除数。例如,在除数为 8 的情况下,值 21 被表示为“110101”:开头的“11”表示处于一元表示的商 (2),之后的零是分隔符,而最后的比特“101”表示余数 5。由于余数

的比特数目依赖于预定义的除数,所以其是固定的。然而,如果开头的一元码或者分隔符由于比特错误而失真,则不可能确定码字长度。从而,也没有发现之后的码字。这种问题类似地适用于所有变长码,使得本发明对于其它变长码也是有用的。

[0034] 如上所述,编码所导致的一个问题是:在比特错误之后,利用现有技术的方法不能使用增强层数据流的随后部分。另外,更糟糕的问题在于:利用现有技术的方法不能检测到该错误,使得解码将使用错误的增强层数据。例如,在音频数据的情况下,这将干扰音频再现,使得与被正确解码的单纯的基本层数据相比,由基本层数据和非同步的增强层数据产生的解码音频信号将更糟糕。在通常情况下,被错误同步的增强层数据的使用干扰了与此不同的正确基本层数据的使用。

[0035] 根据本发明,提供了对于该问题的一种如下解决方案:由于一帧内的增强层数据的变长编码,并且由于变长编码的数据被映射到用于随后处理的字节结构的事实,所以增强层帧的最后字节可能包含填充比特。这些填充比特经常被设置为零并且被忽略。此外,由于各个帧可能由于变长码而具有不同的比特长度,因此填充比特的量逐帧变化。本发明包括:在熵编码之前计算关于基本层帧及其对应的增强层帧的共同的校验和,并且将校验和的尽可能多的比特(优选地是 LSB)写入在熵编码之后剩余的填充比特中。尽管常规的解码器将不能够使用该信息并且因此忽略它,但根据本发明的解码方面的解码器使用该信息来至少检测基本层帧和增强层帧是否同步。

[0036] 根据本发明的一方面,解码器从所定义的起始位置开始,对每个分组解码来自增强层数据的所定义数目的值。然后,解码器从当前增强层帧的最后字节中读取剩余的备用比特(spare bit),该剩余的备用比特通常被设置为零。然而,在根据本发明的增强层帧中,这些比特包含校验和的 LSB。因此,解码器提取/存储这些比特,计算关于解码的基本层帧和当前解码的增强层帧的校验和,并且将所提取的/存储的校验和与所计算的校验和的适当量的 LSB 进行比较。如果两个校验和值相等,则假定基本层帧和增强层帧已经被正确解码并且属于彼此,即,各流是同步的。

[0037] 在改进的实施例中,在已经检测到同步错误之后重新进行同步也是可能的。例如通过以下步骤,这将是可能的,即,在同步丢失之后解码随后的 BL 帧,解码来自 EL 帧的所定义量的值,从 EL 帧的最后字节中提取剩余的比特,计算关于解码的 BL 和 EL 帧的校验和,并且将该校验和与所提取的填充比特进行比较。在不匹配的情况下,逐比特地对 EL 中的起始位置进行移位,并且重复与 EL 有关的过程。如果在对应的 BL 和 EL 帧中不存在比特错误,则可以利用该方法来发现流中的这些帧并且从而进行重新同步。在一个实施例中,使用所描述的处理在后处理步骤中修补损坏的增强层流。为此目的,如前所述地检测 EL 分组中的比特错误,如上所述地确定下一正确的 EL 分组,并且修改错误的 EL 分组以便满足格式要求。在另一实施例中,在增强层帧比基本层帧更可靠的情况下(例如,由于更高质量的接收条件而导致没有错误),则可以确定或者修补用于给定 EL 分组的对应 BL 分组。

[0038] 如上所述,增强层帧数据的正确解码确保可以同步基本层和增强层。在下述的与 MP3 有关的实施例中,解码器是局部(partial)扩展层解码器。在数据组块或文件中的所定义的位置处(即,在基本层数据的最后帧之后)开始增强层的解码。增强层数据的起始对于发现例如其是否处于单独的文件或者流中是无关紧要的。在单个文件或者信道用于基本层和增强层的情况下,可以在文件首标中提供指向该位置的指针。另一可能性在于,增强层

组块 11 以特殊的指示（例如，特定的同步字）起始。然而，增强层没有用于同步目的的另外的附加字节。

[0039] 在表示例如 hd3 编解码器的实施例中，使用比特精度的 (bit exact) 熵编码算法。该算法使用可变的码字长度，将码字的长度适配于统计的码元出现频率。因此，每个帧的帧长度（在比特方面）变化并且不是直接在字节边界处结束。为了允许逐字节地剪切扩展层或者跳转到特定的扩展帧的起始（其例如根据在文件首标中存储的表格可能是已知的），通过使用填充比特总是将帧长度舍入 (round up) 到下一字节边界。从而，根据本发明，每个帧出现并且可以使用平均量 3.5 个填充比特（假定每个字节具有 8 个比特）。通常，如果一个字节具有 k 个比特，则填充比特将是 $(k-1)/2$ 个比特。

[0040] 如上面所提到的，对于这种类型的比特流的问题在于：该比特流内的错误将不仅导致当前帧中的错误，而且还将干扰同步。此外，在上述的分层级的 MP3 比特流格式中，因为扩展层帧需要其特定的 MP3 帧的被局部解码的值以重构编码信号，所以扩展层不得与基本层进行同步。因此，为保持扩展层同步，必须保证 MP3 帧和扩展层帧同步，并且保证扩展层的局部解码是正确的。

[0041] 本发明的一个优点在于：其实现了两种同步问题而不使用任何额外的比特或者字节，即，不增加数据量。

[0042] 示例性地，以下描述 hd3 解码器中失去同步检测的实施例。首先，为确保解码的 MP3 数据（频谱 MP3 值）是有效的，计算被局部解码的 MP3 值的整数和。其次，为了验证扩展层数据的正确解码，计算被局部解码的扩展层数据的整数和，并且将其加到被局部解码的 MP3 值的整数和上。原理上，该结果是两个对应的 hd3 帧的散列值。由于在熵编码之前计算散列值，所以该散列值在每个编码设备和解码设备上具有相同的比特精度，而与实际设备无关并且与实际的熵编码方法无关。通过接收比特流内的散列值，解码器能够检测同步错误并且对其进行反应。尽管不同的熵编码方法可能导致不同的增强层帧长度，并且由此导致不同量的填充比特，但校验和是相同的。此外，由于在解码器内确定要比较的比特量，所以解码器仍能够检测同步丢失。

[0043] 本发明的一方面是仅仅使用散列值的 LSB 并且将这些比特存储在每个扩展层帧的填充比特中。然后，解码器可以在局部（即，熵）解码两个层之后，提取填充比特并且从局部解码的值来计算散列值。如果散列值的 LSB 不等于填充比特，则在局部解码处理中存在错误，并且扩展层可能通过进一步解码而脱离同步。结果，在一个实施例中，解码器可以停止整个解码处理。然而，在另一实施例中，因为基本层组块的解码仍然是可行的，所以解码器切换到单独解码基本层帧。然后，如上所述，可以继续单纯的基本层解码，并且可以抑制增强层解码或者尝试重新进行增强层的同步。该例程的具体优点在于：可以检测并且切掉被错误同步的增强层数据，并且由此不干扰基本层数据的再现。

[0044] 在下文中，如图 4 中所示，描述了示例性的所接收的同步字和所计算的同步字的比较。使用以下各项：

[0045] $X_{mp3}(k)$ \triangleq 被局部解码的 mp3 值，

[0046] $X_{ext}(k)$ \triangleq 被局部解码的扩展值

[0047] $H_{padding}$ \triangleq 所读取的填充比特

[0048] $N_{padding}$ \triangleq 所使用的填充比特的数目

[0049] $H_{frame} = \sum_K X_{mp3}(k) + \sum_K X_{ext}(k) \hat{=}$ 从局部解码的值中计算的散列值

[0050] 在该示例中,在解码器中计算关于被局部解码的基本层帧的值 $X_{mp3}(k)$ 及其对应的增强层帧的值 $X_{ext}(k)$ 的 16 比特校验和 H_{frame} 。检测到四个比特作为所接收的增强层帧的最后字节中的填充比特,并且从这些填充比特中提取所传送的 4 比特校验和 $H_{padding}$ 。为了检测正确的同步,所提取的填充比特 $H_{padding}$ 必须等于所计算的校验和 H_{frame} 的 LSB。如果它们不同,则这意味着帧失去同步(或者已经出现了比特错误,从而还可以对其进行校验)。

[0051] 本发明的一方面在于:不需要在增强层流中插入附加的数据字节。因此,与已知的错误检测方法不同,填充比特的数目是不同的,在 0 和 7 比特之间变化,使得错误检测概率也不同。因此,可能在无效的第一帧处没有立即检测到同步丢失。然而,因为之后的散列值也将是无效的,所以之后的每个帧增加了检测到错误的概率。因此,所提出的方法足以至少在几个帧内检测到失去同步的情况。在与音频有关的实施例中,考虑到 MP3 帧典型地表示几毫秒的声音,则将足够快地检测到这样的情况以防止对听众的较大干扰。因此,所提出的失去同步检测对于音频编码是特别有利的。

[0052] 在本发明的另一实施例中,如图 5 和图 6 中所示,所描述的失去同步检测机制被实施为基于 MP3 的分层级的无损音频编解码器。

[0053] 图 5 示出生成 MP3 基本层数据和音频增强层数据的示例性的音频编码器。将输入信号 200 输入常规的 MP3 滤波器组和量化单元 203,并输入整数 MDCT 计算单元 240。MP3 滤波器组提供 MP3 库 (bins) 215, MP3 库 215 被熵编码 220 并且然后表示作为输出的常规的 MP3 格式化的音频信号 225,即,MP3 比特流。通过 MP3 库到整数 MDCT 值的映射 230 获得增强层信号。这些与实际的整数 MDCT 240 的输出进行比较 250,其中计算 250 剩余残差 255。该残差 255 被熵编码 270,由此使用如上所述的并在图 3 中所示的变长编码。图 7 更详细地示出变长编码的分组 273 被映射 271 到字节。如上面所解释的,大多数被这样编码的分组 275 将具有可能被暂时设置为零的结尾的填充比特。在此确定 272 填充比特的数目 N 是明智的。

[0054] 在计算块 260 中,计算关于基本层帧的 MP3 库 215 和对应的增强层帧的残差 255 的散列值 265。这两个对应的帧都已经从输入信号的相同部分中生成,并且因此需要在高级解码器中将两者组合为输出信号的相同部分。然而,常规的解码器可能忽略增强层信号,并且仅仅将基本层信号解码为具有 MP3 质量的常规音频信号。此外,能够解码增强层信号的解码器不必一定执行失去同步检测,而是它们可以忽略散列值。然而,这些解码器冒着不能够检测失去同步情况的风险,并且由此可能比根据本发明的解码器提供更糟糕的输出信号质量。

[0055] 从 MP3 库 215 和整数残差 255 内的整数频谱 (integer spectra) 来计算 260 该散列值 265,就像稍后将在解码器中从解码的整数频谱来计算它一样。散列值 265 可以是关于这些整数值的简单的校验和,并且在其计算期间通常将具有 $M(M \geq N)$ 比特的固定宽度。原理上,将所确定的增强层帧结尾处可用填充比特量 N 274 提供给截短单元 261,该截短单元 261 从 M 比特散列值 265 中提取或者截出 N 个最低有效比特。然后将作为结果的截短的 N 比特散列值 263 插入 280 到增强层帧 275 的填充比特中,由此获得根据本发明的一方面的熵编码的增强层信号 285。由于增强层比特流允许从质量降低的 MP3 信号中无损地重构原始质量(如来自 CD 的原始质量),所以其在此也被称作“CD 层”比特流。

[0056] 在一个实施例中,将单一的初始同步字添加 290 到增强层比特流 295。在可能难以发现该流的开始的情形下这可能是有利的:例如,如果如图 1 所示在单个文件中存储两个层,并且文件首标中的指针指向第一增强层分组,则该指针中的比特错误将导致不能发现增强层。利用这样的单一的初始同步字,可以针对增强层的起始而分解该流。然而,其仅仅提供初始的同步,并且不适合于检测稍后的各层之间的同步丢失。

[0057] 很清楚,可以同时执行实施方案中的若干所描述的逻辑步骤,从而可能物理上得不到中间结果。

[0058] 图 6 示出解码基本层帧和增强层帧且同时执行失去同步检测的对应解码器。该解码器接收两个分离的比特流,例如,MP3 比特流 101 和增强层比特流 102。这两个比特流然后被局部解码 110、120,即,熵解码。术语“局部”解码指以下事实:该解码的结果不是想要的最终音频信号。替代地,局部解码的结果分别是用于 MP3 比特流的解码的 MP3 库 115 和用于增强层的解码的整数 MDCT 残差库 125。整数 MDCT 残差库 125 可以用于增强 MP3 库,使得获得如同在(有损的)MP3 编码之前的原始音频值。从而,利用增强层数据,可以基于有损的 MP3 数据获得音频数据的无损编码。

[0059] 两个局部解码的比特流 115、125 被表示为整数值。对于同步错误检测 160,从两个局部解码器 110、120 的结果 115、125 来计算 140 单个散列值 145。在当前的示例中,通过简单地添加来自 MP3 频谱的所有频率库加上来自整数 MDCT 残差频谱的所有频率库的整数和来计算散列值 145。如果信号频谱是非零的(如通常是该情形),则该和的几个最低有效比特(LSB)中被截短的散列值表示逐帧相当不相关的伪随机值。在其它实施例中,可以使用其它方法来计算散列,例如仅仅添加某些值,或者基于线性反馈移位寄存器使用类似 CRC 的方法。

[0060] 并行地,提取 150 使增强层表示的编码的子比特流结束(conclude)的填充比特。连同所提取的填充比特 152 一起,将所检测的填充比特的数目 151 转发给下一处理步骤。

[0061] 将从增强层子比特流 102 中取得的散列值 152 的 LSB 与从本地解码的频谱 115、125 中计算的散列值 145 的对应的 LSB 进行比较 160,以进行实际的错误校验。如果该比较指示差异,则生成错误指示标志 165 以便发信号向最终的解码块 130 通知失去同步。

[0062] 剩余的最终解码 130 组合两个局部解码的信号频谱,即,MP3 库 115 和 MDCT 残差 125,并且产生解码器的音频输出信号 135。如果如上所述已经生成错误指示标志 165,则解码器采取适当的措施来应对失去同步的情况。例如,可以将解码处理统统停止,或者解码器可以落入仅仅解码基本层的模式,从而忽略来自增强层的未同步的信息。

[0063] 优选地,本发明涉及分层级的或者可缩放的音频压缩。具体地,对于以下比特流格式,本发明是有利的,在所述比特流格式中,分层级的比特流的不同层经由不同的传输信道传送或者在分离的物理位置中存储。例如,如果分层级编码的音频信号的子比特流的至少大部分被存储在文件的不同位置中,则本发明是有益的;或者在不同的子比特流在传送期间或从存储装置中重新获得期间可能由于传送错误或者存储装置错误而失去同步的情况下,本发明是有益的。本发明使得能够确保在短时间内检测到这样的情况,并且使得能够采取适当的措施。

[0064] 本发明相对于针对同步问题的其它已知的解决方案具有优势。例如,不必与增强层子比特流的块一起连续地发送同步样式(pattern)或者时间戳,使得不需要传送附加的

数据。

[0065] 然而,对于基本层,可能需要连同比特流的每个块或者信号帧一起发送同步样式,如例如在 MPEG-1 层 III (MP3) 中,其中每个帧首标以 12 个数字 1 的固定样式起始。该样式与之后的首标信息的有效性校验一起允许将解码器保持为同步到 1 个比特流。然而,通过使用固定的同步样式不能实现不同的子比特流的同步。本发明通过从上层和下层的时间上对应的帧计算 CRC 或者散列值而总地解决了该问题。由此,在所述上层和对于共同解码有贡献的所有下层之间构建了牢固的链接。

[0066] 与计算/校验关于编码比特流的 CRC 值的常规方式相比,本发明的另一优点在于:从来自子比特流的单独的被局部解码的信息计算 CRC 或者散列值。这反映了不同的思路:典型地, CRC 被应用于保护比特的传送/存储,而在此应用 CRC/散列值以便保护整数参数的比特精度 (bit-exact) 的传送/存储/解码。本发明的另一优点在于:将传送所述 CRC/散列值所需要的额外的信息存储在附加于每个帧边界处的填充比特内以便在字节边界处开始之后的帧。尽管可用的填充比特量是变化的并且不可预测,但其对于向失去同步检测平均提供短的反应时间是足够的。

[0067] 另一优点在于:错误校验涵盖了部分解码处理。也就是说,除了只是校验比特流是对齐的之外,其还校验局部解码处理是比特精度的,即,比特级的同步。

[0068] 本发明还具有以下优点:它能够与可变比特率编码方案一起应用而不需要同步样式。因此,本发明获得非常高的压缩效率。由于对于每个帧在填充比特中存储 CRC/散列信息,所以不需要额外的比特率。利用本发明,不需要连续地跟踪所有子比特流的同步或者校正所有子比特流的接收。可以对于每个信号帧独立地计算错误校验。

[0069] 尽管上面的示例涉及音频信号,但本发明的总体构思还可一般地应用于表示各种类型的信号(例如,视频)的分层级的比特流。同样地,可以将本发明应用于需要与基本层流或者其它增强层流同步的另外的增强层流(例如,可以将第二 EL 流与第一 EL 流同步,或者它们中的每一个可以与 BL 流同步)。此外,可以将本发明应用于同时检测三个或者更多个流之间的同步。

[0070] 在本发明的一方面中,一种用于编码具有 BL 比特流和与 BL 比特流有关的增强层比特流的音频或视频信号的装置,其中以分组构成 BL 数据和增强层数据,并且其中 BL 比特流的分组具有对应的 EL 比特流的分组,所述装置包括:计算部件,计算关于 BL 比特流的分组和对应的 EL 比特流的分组的校验和;BL 熵编码器,熵编码 BL 比特流的分组,其中生成以同步字开始的熵编码的字节对齐的 BL 分组;EL 熵编码器,使用变长编码方法来熵编码 EL 比特流的分组,其中生成不具有首标并且不以同步字开始的熵编码的 EL 分组;映射部件,将编码的 EL 分组映射到字节,每个字节具有固定数目的比特,其中在编码的 EL 分组的最后字节中 N 个数目的比特是填充比特;确定部件,确定编码的 EL 分组的最后字节内的填充比特量 N;提取部件,提取所述所计算的校验和的 N 个最低有效比特,N 等于所确定的填充比特量;插入部件,将所提取的所计算的校验和的 N 个最低有效比特插入到编码分组的结尾处的填充比特中;以及用于提供或输出包括插入的校验和比特的编码 EL 分组以及编码 BL 分组的部件。

[0071] 在本发明的一方面中,所述用于编码的装置还包括用于在文件中存储编码 BL 比特流和编码 EL 比特流的部件,其中在文件的第一部分中彼此相邻地存储 BL 分组,并且在文

件的随后的第二部分中彼此相邻地存储 EL 分组。

[0072] 在本发明的另一方面中,一种用于解码具有 BL 比特流和对应的 EL 比特流的音频或视频信号的装置,其中 EL 比特流包括用于对 BL 比特流进行增强解码的变长编码的参数,所述装置包括:BL 局部解码器,局部解码 BL 比特流的一部分,其中生成解码的第一 BL 分组;EL 局部解码器,局部解码 EL 比特流的一部分,其中生成用于对 BL 比特流进行增强解码的解码参数;用于确定 EL 比特流的所述部分中 N 个数目的填充比特的部件;以及用于提取和/或存储填充比特的部件;计算部件,计算关于解码的第一 BL 分组和来自 EL 比特流的局部解码部分的所述解码参数的校验和;比较部件,将所提取的所计算的校验和的 N 个最低有效比特与所提取的 N 个填充比特进行比较并且如果两者不相等则提供错误指示,其中 N 等于所确定的填充比特量;以及解码部件,如果根据所述错误指示所提取的所计算的校验和的 N 个最低有效比特与所提取的 N 个填充比特相等,则使用用于对 BL 比特流进行增强解码的所述解码参数来解码被局部解码的 BL 比特流,否则,利用其中不使用来自 EL 比特流的数据的简化解码来解码被局部解码的 BL 比特流。

[0073] 将理解已经单纯地通过示例描述了本发明,并且可以做出对细节的修改而不脱离本发明的范围。

[0074] 可以独立地或者以任何适当的组合来提供在说明书和(在适当的情况下)权利要求书以及附图中公开的每个特征。在适当的情况下,可以以硬件、软件、或者两者的组合来实施这些特征。在可应用的情况下,可以将连接实施为无线连接或者有线连接,不一定是直接的或者专用的连接。出现在权利要求书中的参考标记仅仅是说明性的并且不应对于权利要求的范围有限制效果。

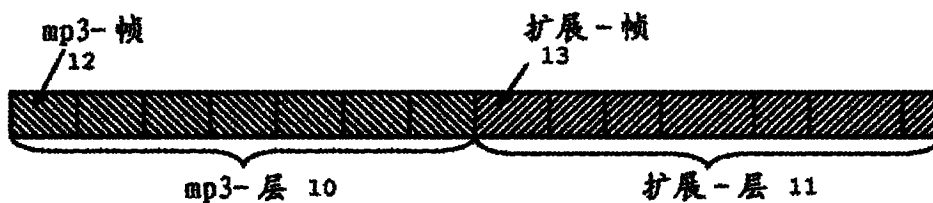


图 1

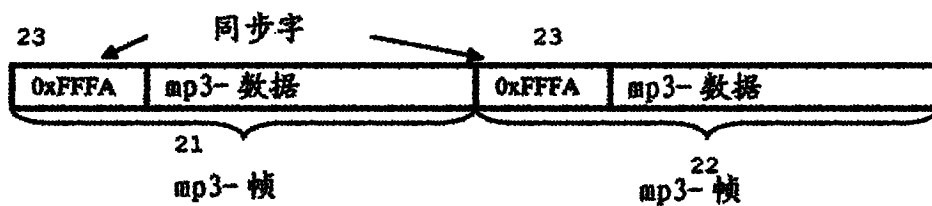


图 2

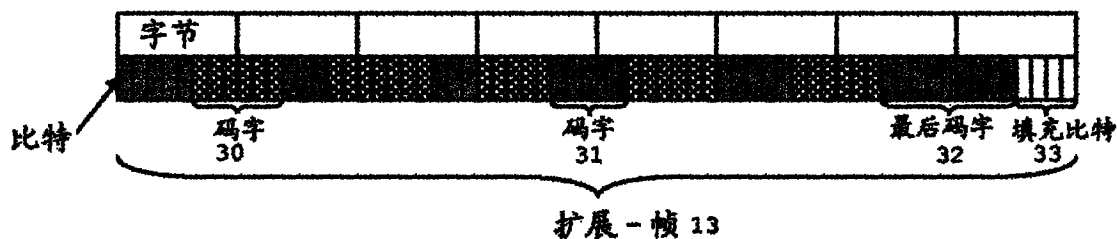


图 3

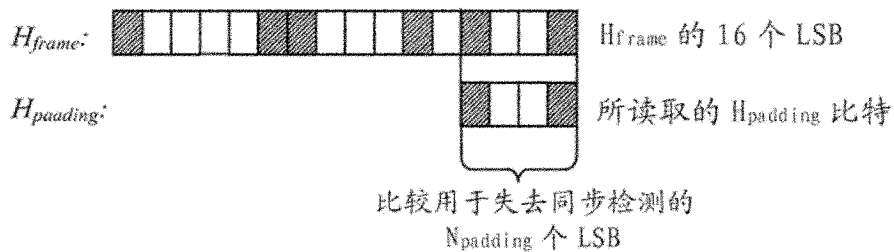


图 4

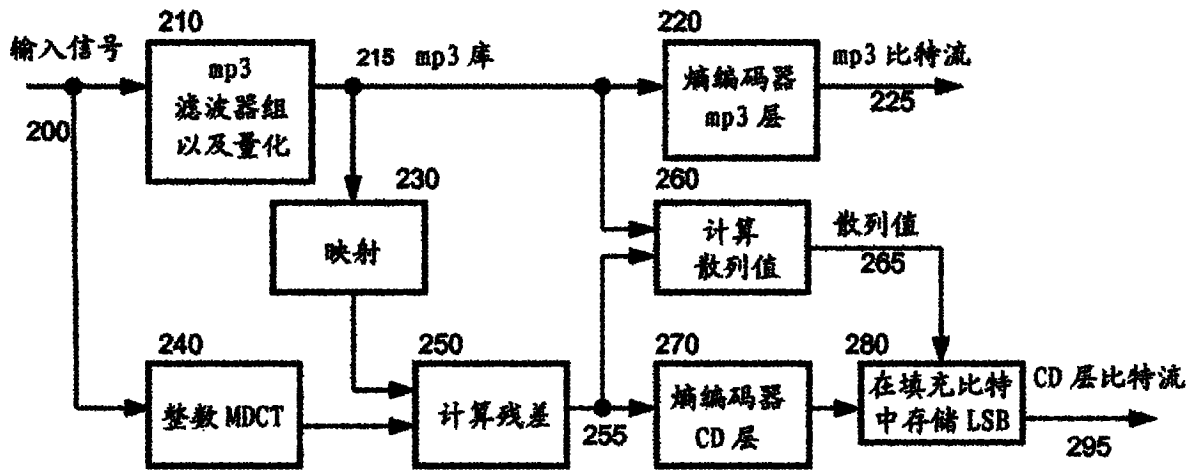


图 5

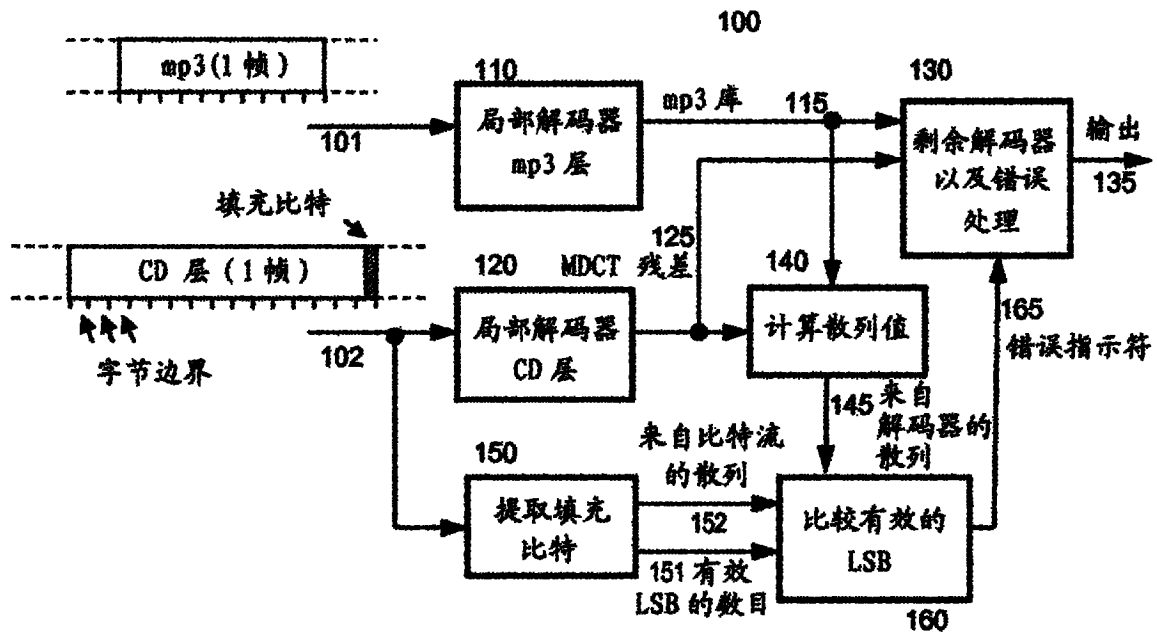


图 6

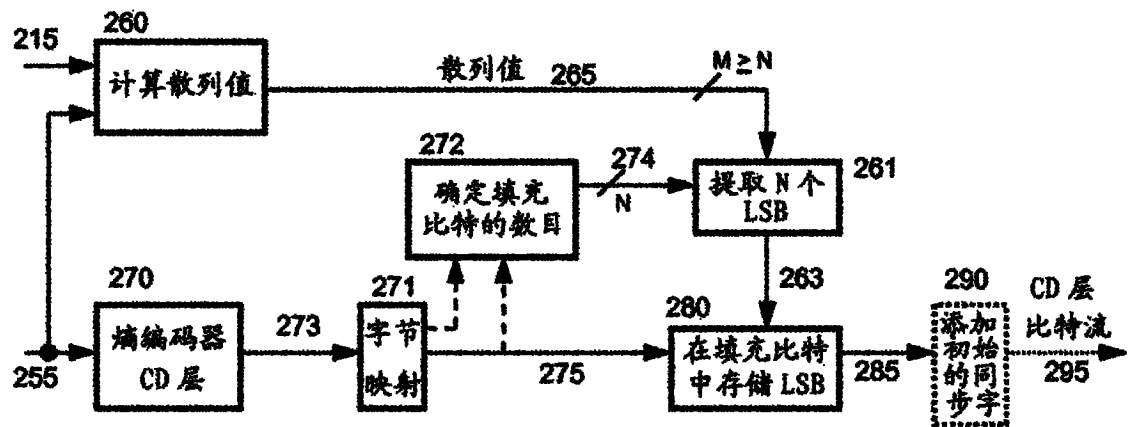


图 7