



(12)发明专利

(10)授权公告号 CN 107276552 B

(45)授权公告日 2020.09.11

(21)申请号 201710317299.7

(22)申请日 2013.01.29

(65)同一申请的已公布的文献号

申请公布号 CN 107276552 A

(43)申请公布日 2017.10.20

(30)优先权数据

61/754,882 2013.01.21 US

(62)分案原申请数据

201310034592.4 2013.01.29

(73)专利权人 杜比实验室特许公司

地址 美国加利福尼亚州

(72)发明人 杰弗里·里德米勒 迈克尔·沃德

(74)专利代理机构 北京集佳知识产权代理有限公司

公司 11227

代理人 吴琼 康建峰

(51)Int.Cl.

H03G 5/00(2006.01)

H03G 9/00(2006.01)

H03G 9/02(2006.01)

G10L 19/16(2013.01)

(56)对比文件

CN 101777370 A,2010.07.14

CN 1761308 A,2006.04.19

AU 2006223416 A1,2007.10.04

US 2009164473 A1,2009.06.25

WO 2006113062 A1,2006.10.26

US 2003103645 A1,2003.06.05

审查员 张瑞

权利要求书2页 说明书32页 附图6页

(54)发明名称

解码具有保留数据空间中的元数据容器的
编码音频比特流

(57)摘要

本发明公开了解码具有保留数据空间中的元数据容器的编码音频比特流。具体公开处理音频信号的设备,包括:输入缓冲器,存储编码音频比特流的至少一帧,编码音频比特流包括音频数据和元数据容器,元数据容器包括一或多个元数据有效载荷和保护数据;音频解码器,耦接至输出缓冲器并用于解码音频数据以产生解码音频数据;解析器,耦接至音频解码器或与之集成,解析器被配置成解析音频数据;以及所述输出缓冲器,存储解码音频数据,元数据容器以标识元数据容器开始的同步字开始,保护数据位于一或多个元数据有效载荷之后,并且每个元数据有效载荷以标识有效载荷中的元数据类型的有效载荷标识符开始,其后跟随标识有效载荷中的元数据量的有效载荷尺寸。



1. 一种用于处理音频信号的设备,所述设备包括:

输入缓冲器,所述输入缓冲器用于存储编码音频比特流的一帧的至少一部分,其中,所述编码音频比特流包括音频数据和元数据容器,并且所述元数据容器包括一个或更多个元数据有效载荷和保护数据,其中,所述一个或更多个元数据有效载荷包括节目响度有效载荷;

音频解码器,所述音频解码器被耦接至输出缓冲器,所述音频解码器用于解码所述音频数据以产生解码音频数据;

解析器,所述解析器被耦接至所述音频解码器或者与所述音频解码器集成,其中,所述解析器被配置成解析所述音频数据;以及

所述输出缓冲器,所述输出缓冲器用于存储所述解码音频数据,

其中,所述元数据容器以标识所述元数据容器的开始的同步字开始,所述保护数据位于所述一个或更多个元数据有效载荷之后,并且所述一个或更多个元数据有效载荷中的每个以标识存在于所述有效载荷中的元数据的类型的有效载荷标识符开始,跟随有效载荷标识符的是标识存在于所述有效载荷中的元数据的量的有效载荷尺寸。

2. 根据权利要求1所述的设备,其中,所述元数据容器被存储在选自skip域、auxdata域、addbsi域及其组合的AC-3或者E-AC-3保留数据空间中。

3. 根据权利要求1所述的设备,其中,所述一个或更多个元数据有效载荷包括表示连续音频节目之间的至少一个边界的元数据。

4. 根据权利要求1所述的设备,其中,所述节目响度有效载荷包含表示测量的音频节目的响度的数据。

5. 根据权利要求4所述的设备,其中,所述节目响度有效载荷包括指示音频通道是否包含口语会话的域。

6. 根据权利要求4所述的设备,其中,所述节目响度有效载荷包括指示已用于生成包含在所述节目响度有效载荷中的响度数据的响度测量方法的域。

7. 根据权利要求4所述的设备,其中,所述设备被配置成使用所述节目响度有效载荷来执行自适应响度处理。

8. 根据权利要求1所述的设备,其中,所述编码音频比特流是AC-3比特流或者E-AC-3比特流。

9. 根据权利要求1所述的设备,其中,每个有效载荷标识符是唯一的并且位于每个元数据有效载荷的开始处。

10. 根据权利要求1所述的设备,其中,所述同步字是值为0x5838的16比特同步字。

11. 一种音频解码器,包括:

输入缓冲器,所述输入缓冲器用于存储编码音频比特流的一帧的至少一部分,其中,所述编码音频比特流包括音频数据和元数据容器,并且所述元数据容器包括一个或更多个元数据有效载荷和保护数据,其中,所述一个或更多个元数据有效载荷包括节目响度有效载荷;

解析器,所述解析器被配置成解析所述音频数据,其中,所述元数据容器以标识所述元数据容器的开始的同步字开始,所述保护数据位于所述一个或更多个元数据有效载荷之后,并且所述一个或更多个元数据有效载荷中的每个以标识存在于所述有效载荷中的元数

据的类型的有效载荷标识符开始,跟随有效载荷标识符的是标识存在于所述有效载荷中的元数据的量的有效载荷尺寸;以及

输出缓冲器,所述输出缓冲器用于存储解码音频数据。

12.根据权利要求11所述的音频解码器,其中,所述元数据容器被存储在选自skip域、auxdata域、addbsi域及其组合的AC-3或者E-AC-3保留数据空间中。

13.根据权利要求11所述的音频解码器,其中,所述编码音频比特流是AC-3比特流或者E-AC-3比特流。

14.根据权利要求11所述的音频解码器,其中,所述节目响度有效载荷包含表示测量的音频节目的响度的数据。

15.根据权利要求11所述的音频解码器,其中,所述同步字是值为0x5838的16比特同步字。

16.一种非暂态计算机可读介质,包含编码音频比特流的至少一帧,所述至少一帧包括:

音频数据;以及

元数据容器,所述元数据容器包括一个或更多个元数据有效载荷和保护数据,其中,所述一个或更多个元数据有效载荷包括节目响度有效载荷;

其中,所述元数据容器以标识所述元数据容器的开始的同步字开始,所述保护数据位于所述一个或更多个元数据有效载荷之后,并且所述一个或更多个元数据有效载荷中的每个以标识存在于所述有效载荷中的元数据的类型的有效载荷标识符开始,跟随有效载荷标识符的是标识存在于所述有效载荷中的元数据的量的有效载荷尺寸。

17.根据权利要求16所述的非暂态计算机可读介质,其中,所述元数据容器被存储在选自skip域、auxdata域、addbsi域及其组合的AC-3或者E-AC-3保留数据空间中。

18.根据权利要求16所述的非暂态计算机可读介质,其中,所述编码音频比特流是AC-3比特流或者E-AC-3比特流。

19.根据权利要求16所述的非暂态计算机可读介质,其中,所述节目响度有效载荷包含表示测量的音频节目的响度的数据。

20.根据权利要求16所述的非暂态计算机可读介质,其中,所述同步字是值为0x5838的16比特同步字。

解码具有保留数据空间中的元数据容器的编码音频比特流

[0001] 本申请为2013年1月29日提交的申请号为201310034592.4、发明名称为“利用响度处理状态元数据的音频编码器和解码器”的发明专利申请的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请要求于2013年1月21日提交的、题为“利用响度处理状态元数据的音频编码器和解码器”、发明人为Michael Ward和Jeffrey Riedmiller的美国临时专利申请No.61/754,882的优先权。

技术领域

[0004] 本发明涉及音频信号处理,更具体地,本发明涉及使用表示音频内容的响度处理状态的元数据对音频数据比特流进行编码和解码。本发明的一些实施例生成或解码作为“杜比数字(Dolby Digital)(AC-3)”、“杜比数字+(Dolby Digital Plus)(增强型AC-3或E-AC-3)”或者“杜比E(Dolby E)”已知的格式之一的音频数据。

背景技术

[0005] “杜比”、“杜比数字”、“杜比数字+”和“杜比E”是杜比实验室特许公司的商标。杜比实验室提供分别作为“杜比数字”和“杜比数字+”已知的AC-3和E-AC-3的专有实施。

[0006] 音频数据处理单元通常以盲目的方式来操作,并且不注意在数据被接收之前出现的音频数据的处理历史。这可能在以下处理框架内起作用:其中,单个实体进行各种各样的目标媒体渲染装置的所有的音频数据处理和编码,同时,目标媒体渲染装置进行对编码音频数据的所有的解码和渲染。然而,该盲目的处理在以下情况下不能很好地起作用(或者一点都不起作用):其中,多个音频处理单元散布在多种多样的网络上或者串联放置(即,链式放置)并被期望最优地执行它们各自类型的音频处理。例如,某些音频数据可以被编码以用于高性能的媒体系统并且可能必须沿着媒体处理链被转换成适合移动装置的缩减形式。因而,音频处理单元可能不必对已经被执行了某种类型的处理的音频数据执行该类型的音频处理。例如,音量调节(volume leveling)单元可以对输入音频剪辑执行处理,而不管是否之前已经对输入音频剪辑执行了相同的或类似的音量调节。因此,音量调节单元可能在不需要时执行调节。该非必要的处理还可能引起在对音频数据的内容进行渲染时的特定特征的去除和/或降级。

[0007] 音频数据的典型的流包括音频内容(如,音频内容的一个或更多个通道)和表示音频内容的至少一个特征的元数据二者。例如,在AC-3比特流中,存在若干音频元数据参数,这些音频元数据参数具体地意在用于改变被递送到倾听环境的节目的声音。元数据参数之一是“DIALNORM”参数,其意在表示出现音频节目的会话的平均电平,并且用于确定音频回放信号电平。

[0008] 在包括不同的音频节目分段(每个音频节目分段具有不同的DIALNORM参数)序列的比特流的回放期间,AC-3解码器使用每个分段的DIALNORM参数来执行某种类型的响度处理,其中,其修改回放电平或响度,使得分段序列的会话的感知响度处于恒定的电平。编码

音频项的序列中的每个编码音频分段(项)会(通常)具有不同的DIALNORM参数,并且解码器可以对其中每个项的电平进行缩放,使得每个项的会话的回放电平或响度相同或者非常类似,虽然这可能需要在回放期间将不同量的增益应用于不同的项。

[0009] DIALNORM通常由用户来设置,并且不是自动生成的,虽然如果用户没有设置任何值时则存在默认的DIALNORM值。例如,内容产生器可以用AC-3编码器外部的装置来进行响度测量,并且接着将(表示音频节目的口语会话的响度的)结果传输给编码器以设置DIALNORM值。因此,存在为了正确地设置DIALNORM参数而对内容产生器的依赖。

[0010] AC-3比特流中的DIALNORM参数可能不正确的原因有若干不同的原因。首先,如果内容产生器没有设置DIALNORM值,则每个AC-3编码器具有在比特流的生成期间使用的默认的DIALNORM值。该默认值可能与音频的实际会话响度电平相当不同。第二,即使内容产生器测量响度并且相应地设置DIALNORM值,也可能使用不遵守推荐的AC-3响度测量方法的响度测量算法或仪表,从而导致错误的DIALNORM值。第三,即使已经利用内容产生器正确地测量的并且设置的DIALNORM值产生了AC-3比特流,其可能在比特流的传输和/或存储期间已经被改为错误的值。例如,在电视广播应用中,使用错误的DIALNORM元数据信息来对AC-3比特流进行解码、修改和接着重新编码不是不常见。因此,AC-3比特流中所包括的DIALNORM值可能是不正确的或者不准确的,因此可能对于倾听体验的质量有负面影响。

[0011] 此外,DIALNORM参数不指示相应的音频数据的响度处理状态(如,已经对音频数据执行了什么类型的响度处理)。在本发明之前,音频比特流一直没有以本公开中描述的类型格式包括元数据,所述元数据表示音频比特流的音频内容的响度处理状态(如,所应用的响度处理的类型)、或者比特流的音频内容的响度处理状态和响度。这样的格式的响度处理状态元数据用于以特别有效的方式便利对音频比特流的自适应响度处理和/或音频内容的响度处理状态和响度的有效性的验证。

[0012] 国际提交日为2011年12月1日并且被受让给本申请的受让人的公开号为WO 2012/075246A2的PCT国际申请公开了用于生成、解码和处理包括表示音频内容的处理状态(如,响度处理状态)和特性(如,响度)的元数据在内的音频比特流的方法和系统。该参考文献还描述了使用元数据执行的对比特流的音频内容的自适应处理、以及使用元数据执行的对比特流的音频内容的响度处理状态和响度的有效性的验证。然而,该参考文献没有描述以本公开内容中描述的类型格式在音频比特流中包括表示音频内容的响度处理状态和响度的元数据(LPSM)。如所注意到的,这样的格式的LPSM用于以特别有效的方式便利对流的自适应响度处理和/或音频内容的响度处理状态和响度的有效性的验证。

[0013] 虽然本发明不限于与AC-3比特流、E-AC-3比特流或者杜比E比特流一起使用,然而,为了方便,将在实施例中进行描述,在实施例,其生成、解码或者处理这样的包括响度处理状态元数据在内的比特流。

[0014] AC-3编码比特流包括元数据以及音频内容的一至六个通道。音频内容是已经使用感知音频编码被压缩的音频数据。元数据包括意在用于改变被递送给倾听环境的节目的声音的若干音频元数据参数。

[0015] AC-3(也称为杜比数字)编码的细节是公知的,并且在很多公开参考文献中被阐明,这些公开的参考文献包括:

[0016] ATSC Standard A52/A:Digital Audio Compression Standard(AC-3),Revision

A,Advanced Television Systems Committee,20Aug.2001;以及

[0017] 美国专利5,583,962、5,632,005、5,633,981、5,727,119和6,021,386。

[0018] “Introduction to Dolby Digital Plus,an Enhancement to the Dolby Digital Coding System”,AES Convention Paper 6191,117th AES Convention,October 28,2004中阐明了杜比数字+ (E-AC-3) 编码的细节。

[0019] “Efficient Bit Allocation,Quantization,and Coding in an Audio Distribution System”,AES Preprint 5068,107th AES conference,August1999和 “Professional Audio Coder Optimized for Use with Video”,AES Preprint 5033, 107th AES Conference August 1999中阐明了杜比E编码的细节。

[0020] AC-3编码音频比特流的每个帧包含用于数字音频的1536个样本的音频内容和元数据。对于48kHz的采样速率,这表示32毫秒的数字音频或者音频的每秒31.25个帧的速率。

[0021] 取决于帧包含一个、两个、三个还是六个音频数据块,E-AC-3编码音频比特流的每个帧分别包含用于数字音频的256、512、768或1536个样本的音频内容和元数据。对于48kHz的采样速率,这分别表示5.333、10.667、16或32毫秒的数字音频或者音频的每秒189.9、93.75、62.5或31.25个帧的速率。

[0022] 如图4所示,每个AC-3帧被分成区段(分段),包括:同步信息(SI)区段,其包含(如图5所示)同步字(SW)和两个纠错字中的第一纠错字(CRC1);比特流信息(BSI)区段,其包含元数据的大多数;一个至六个之间的音频块(AB0至AB5),其包含数据压缩的音频内容(也可以包括元数据);浪费比特(W),其包含在音频内容被压缩之后留下的任意的未使用的比特;辅助(AUX)信息区段,其可以包含更多的元数据;以及两个纠错字中的第二纠错字(CRC2)。

[0023] 如图7所示,每个E-AC-3帧被划分成区段(分段),包括:同步信息(SI)区段,其包含(如图5所示)同步字(SW);比特流信息(BSI)区段,其包含元数据的大多数;一个至六个之间的音频块(AB0至AB5),其包含数据压缩的音频内容(也可以包括元数据);浪费比特(W),其包含在音频内容被压缩之后留下的任意的未使用的比特;辅助(AUX)信息区段,其可以包含更多的元数据;以及纠错字(CRC)。

[0024] 在AC-3(或者E-AC-3)比特流中,存在若干音频元数据参数,该音频元数据参数具体地意在用于改变被递送给倾听环境的节目的声音。元数据参数之一是DIALNORM参数,其包括在BSI区段中。

[0025] 如图6所示,AC-3帧的BSI区段包括五比特的参数(“DIALNORM”),其指示用于该节目的DIALNORM值。如果AC-3帧的音频编码模式(“acmod”)为“0”,指示使用双-单或者“1+1”通道配置,则包括用于指示同一AC-3帧中承载的第二音频节目的DIALNORM值的五比特的参数(“DIALNORM2”)。

[0026] BSI分段还包括:用于表示跟随“addbsie”比特的附加比特流信息的存在(或不存在的)标志(“addbsie”)、用于表示跟随“addbsil”值的任意附加比特流信息的长度的参数(“addbsil”)、以及跟随“addbsil”值的最高达64比特的附加比特流信息(“addbsi”)。

[0027] BSI分段包括没有在图6中具体示出的其他元数据值。

发明内容

[0028] 在一类实施例中,本发明是一种包括以下步骤的方法:通过在比特流的至少一个

帧的至少一个分段中包括响度处理状态元数据 (LPSM) 并且在该帧的至少一个其他分段中包括音频数据,对音频数据进行编码以生成包括响度处理状态元数据和音频数据在内的编码音频比特流。在典型的实施例中,该方法包括步骤:在比特流的每个帧中复用音频数据与 LPSM。在典型的解码中,解码器(通过解析和解复用 LPSM 和音频数据)从包括 LPSM 和音频数据的比特流中提取 LPSM,并且对音频数据进行处理以生成解码音频数据流(并且,在某些情况下,还使用 LPSM 执行对音频数据的自适应响度处理或者对 LPSM 和/或音频数据的认证和/或验证中的至少一个)。在某些情况下,解码音频数据和 LPSM 从解码器转发给后处理器,后处理器被配置成使用 LPSM 对解码音频数据执行自适应响度处理。自适应响度处理可以包括动态范围和/或响度控制(如,会话响度调节或其他音量调节),或者自适应响度处理可以由动态范围和/或响度控制构成。响应于 LPSM,音频处理单元可以禁止(如 LPSM 所示)已经对相应的音频内容执行的响度处理。

[0029] 可以对根据本发明的典型的实施例的嵌入在音频比特流中的响度处理状态元数据进行认证和验证,如使能响度调节实体来验证具体的节目的响度是否已经在指定的范围内以及相应的音频数据本身是否还没有被修改过(从而确保符合可应用的规则)。可以读取包括响度处理状态元数据在内的数据块中所包括的响度值,以对其进行验证,代替再次计算响度。响应于 LPSM,管理机构可以判定相应的音频内容是否符合(如 LPSM 所示)响度法规和/或管理要求(如,在商业广告响度降低法案 (Commercial Advertisement Loudness Mitigation Act),也称为“CALM”法案下公布的规则),而不需要计算音频内容的响度。

[0030] 本发明的另外的方面是被配置成执行本发明的方法的任意实施例的音频处理单元 (APU)。在另一类实施例中,本发明是包括缓冲存储器(缓冲器)的 APU,缓冲存储器(如,以非暂时性方式)存储已经由本发明的方法的任意实施例生成的编码音频比特流的至少一个帧。APU 的示例包括但不限于编码器(如,转码器)、解码器、编解码器、预处理系统(预处理器)、后处理系统(后处理器)、音频比特流处理系统以及这样的元件的组合。

[0031] 在另一类实施例中,本发明是被配置成生成包括音频数据分段和元数据分段的编码音频比特流的音频处理单元 (APU),其中,音频数据分段表示音频数据,并且至少部分元数据分段中的每个包括响度处理状态元数据 (LPSM)。通常,比特流的帧中的至少一个这样的元数据分段包括:LPSM 的表示是否已经对帧的音频数据(即在所述帧的至少一个音频数据分段中的音频数据)执行了第一类型的响度处理的至少一个分段;以及 LPSM 的表示帧的音频数据中的至少一些的响度(如,表示会话的帧的音频数据中的至少一些的会话响度)的至少一个其他分段。在这种类型的一个实施例中,APU 是一种被配置成对输入音频进行编码以生成编码音频的编码器,音频数据分段包括编码音频。在这种类型的典型实施例中,每个元数据分段具有本文中要描述的优选格式。

[0032] 在一种优选格式中,编码比特流是 AC-3 比特流或者 E-AC-3 比特流,并且,每个包括 LPSM 的元数据分段作为附加比特流信息被包括在比特流的帧的比特流信息(“BSI”)分段的“addbsi”域中。每个包括 LPSM 的元数据分段具有本文中在以下参考表 1 和表 2 指定的格式(即,其包括表 1 所示的核心元素或者其变型,核心元素或者其变型之后跟随有效载荷 ID(将元数据标识为 LPSM)和有效载荷尺寸值,有效载荷 ID 和有效载荷尺寸值之后跟随有效载荷(如本文中所述的具有如表 2 所示的格式或者表 2 的变型所示的格式的 LPSM 数据))。

[0033] 在另一优选格式中,编码比特流是 AC-3 比特流或者 E-AC-3 比特流,并且,每个包括

LPSM的元数据分段被包括在比特流的帧的比特流信息(“BSI”)分段的“addbsi”域中,或者被包括在比特流的帧的端部处的auxdata(辅助数据)域(如,图4所示的AUX分段)中。帧可以包括一个或两个元数据分段,每个元数据分段包括LPSM,并且,如果帧包括两个元数据分段,则其中一个元数据分段存在于帧的addbsi域中,而另一个元数据分段存在于帧的AUX域中。包括LPSM的每个元数据分段具有本文中在以下参考表1和表2指出的格式(即,其包括表1所示的核心元素或者其变型,之后跟随有效载荷ID(将元数据标识为LPSM))和有效载荷尺寸值,之后跟随有效载荷(如本文中所述的具有如表2所示的格式或者表2的变型所示的格式的LPSM数据)。

[0034] 在另一优选格式中,编码比特流是一种并非AC-3比特流或者E-AC-3比特流的比特流,并且,包括LPSM的元数据分段中的每个被包括在被保留用于存储附加数据的比特流的分段(或域、或时隙)中。包括LPSM的每个元数据分段可以具有与本文中在以下参考表1和表2指出的格式类似或相同的格式(即,其包括与表1所示的核心元素类似或相同的核心元素,之后跟随有效载荷ID(将元数据标识为LPSM)和有效载荷尺寸值,之后跟随有效载荷(具有与如本文中所述的如表2所示的格式或者表2的变型所示的格式类似或相同的格式的LPSM数据))。

[0035] 在某些实施例中,编码比特流包括帧的序列,每个帧包括比特流信息(“BSI”)分段和auxdata域或时隙(如,编码比特流是AC-3比特流或者E-AC-3比特流),其中,比特流信息(“BSI”)分段包括“addbsi”域(有时称为分段或时隙),并且,每个帧包括音频数据分段(如,图4所示的帧的AB0-AB5分段)和元数据分段,其中,元数据分段表示音频数据,至少部分元数据分段的每个包括响度处理状态元数据(LPSM)。LPSM以以下格式存在于比特流中。包括LPSM的元数据分段中的每个被包括在比特流的帧的BSI分段的“addbsi”域中,或者被包括在比特流的帧的auxdata域中。比特流的帧可以包括一个或两个元数据分段,每个元数据分段包括LPSM,并且,如果帧包括两个元数据分段,则其中一个元数据分段存在于帧的addbsi域中,而另一个存在于帧的AUX域中。包括LPSM的每个元数据分段包括具有以下格式的LPSM有效载荷(或容器)分段:

[0036] 首部(通常包括至少一个标识值,如以下表2中所示的LPSM格式版本、长度、周期、计数和子流关联值);以及

[0037] 在首部之后的:

[0038] 至少一个会话指示值(如,表2的参数“会话通道”),其指示相应的音频数据是指示会话还是不指示会话(如,相应的音频数据的哪个通道指示会话)。会话指示值可以指示会话是否存在于相应的音频数据的通道的任意组合或所有通道中;

[0039] 至少一个响度调节相符值(如,表2的参数“响度调节类型”),其指示相应的音频数据是否与所指示的响度调节的集合相符;

[0040] 至少一个响度处理值(如,表2的参数“会话选通的响度校正标志”、“响度校正类型”中的一个或多个),其指示已经对相应的音频数据执行的至少一种类型的响度处理;以及

[0041] 至少一个响度值(如,表2的参数“ITU相对选通的响度”、“ITU语音选通的响度”、“ITU(EBU 3341)短期3s响度”和“真实峰值”中的一个或多个),其指示相应的音频数据的至少一个响度(如,峰值或者平均响度)特性。

[0042] 在本发明的专注、使用或者生成表示相应的音频数据的至少一个响度值的任意实施例中,响度值可以指示用于处理音频数据的响度和/或动态范围的至少一个响度测量特性。

[0043] 在一些实现中,比特流的帧的“addbsi”域或auxdata域中的每个元数据分段具有以下格式:

[0044] 核心首部(通常包括标识元数据分段的开始的同步字,其后跟随标识值,如以下表1中所示的核心元素版本、长度和周期、扩展元素计数以及子流关联值);以及

[0045] 在核心首部之后的至少一个保护值(如,HMAC摘要和音频指纹值,其中,HMAC摘要可以是基于整个帧的音频数据、核心元素和所有的扩展元素计算的256比特的HMAC摘要(使用SHA-2算法),如表1所示,其用于响度处理状态元数据或相应的音频数据中的至少一个的解密、认证或验证中的至少一个);以及

[0046] 如果元数据分段包括LPSM,则也在核心首部之后的LPSM有效载荷标识(ID)和LPSM有效载荷尺寸值,其将跟随的元数据标识为LPSM有效载荷并且指示LPSM有效载荷的尺寸。LPSM有效载荷分段(优选地具有上述格式)跟随LPSM有效载荷ID和LPSM有效载荷尺寸值。

[0047] 在以上段落中描述的类型的一些实施例中,帧的auxdata域(或者“addbsi”域)中的每个元数据分段具有三层结构:

[0048] 高层结构,包括:指示auxdata(或addbsi)域是否包括元数据的标志;指示存在的是什么类型的元数据的至少一个ID值;以及通常还包括指示存在多少比特的元数据(如,每种类型的元数据)的值(如果存在元数据)。可能存在的一种类型的元数据是LPSM,可能存在的另一种类型的元数据是媒体研究元数据(如,尼尔森(Nielsen)媒体研究元数据);

[0049] 中层结构,包括用于每个标识的类型的元数据的核心元素(如,对于每种标识的类型的元数据如上述类型的核心首部、保护值和有效载荷ID以及有效载荷尺寸值);以及

[0050] 低层结构,包括用于一个核心元素的每个有效载荷(如,如果核心元素将其标识为存在,则是LPSM有效载荷,和/或,如果核心元素将其标识为存在,则是另一种类型的元数据有效载荷)。

[0051] 可以对在这样的三层结构中的数据值进行嵌套。例如,可以在由核心元素标识的有效载荷之后(从而在核心元素的核心首部之后),包括用于LPSM有效载荷和/或由核心元素标识的另一元数据有效载荷的保护值。在一种示例中,核心首部可以标识LPSM有效载荷与另一元数据有效载荷,用于第一有效载荷(如,LPSM有效载荷)的有效载荷ID和有效载荷尺寸值可以跟随核心首部,第一有效载荷本身可以跟随该ID和尺寸值,用于第二有效载荷的有效载荷ID和有效载荷尺寸值可以跟随第一有效载荷,第二有效载荷本身可以跟随这些ID和尺寸值,并且,两种有效载荷之一或两者(或者核心元素值和两种有效载荷之一或两者)的保护值可以跟随最后的有效载荷。

[0052] 在一些实施例中,帧的auxdata域(或“addbsi”域)中的元数据分段的核心元素包括核心首部(通常包括标识值,如核心元素版本),并且在核心首部之后包括:指示指纹数据是否被包括用于元数据分段的元数据的值、指示是否存在外部数据(与对应于元数据分段的元数据的音频数据有关)的值、由核心元素标识的每种类型的元数据(如,LPSM和/或除了LPSM之外的类型的元数据)的有效载荷ID和有效载荷尺寸值、以及由核心元素标识的至少一种类型的元数据的保护值。元数据分段的元数据有效载荷跟随核心首部,并且(在一些情

况下)被嵌套在核心元素的值内。在另一优选格式中,编码比特流是杜比E比特流,并且,包括LPSM的元数据分段中的每个被包括在杜比E保护带间隔的前N个样本位置中。

[0053] 在另一类型的实施例中,本发明是APU(如,解码器),APU被耦接和配置来接收包括音频数据分段和元数据分段的编码音频比特流的,其中,音频数据分段表示音频数据,并且至少部分元数据分段中的每个包括响度处理状态元数据(LPSM),并且APU被耦接和配置来从比特流中提取LPSM,响应于音频数据来生成解码音频数据、以及使用LPSM对音频数据执行至少一个自适应响度处理操作。这种类型中的某些实施例还包括耦接至APU的后处理器,其中,后处理器被耦接和配置来使用LPSM对音频数据执行至少一个自适应响度处理操作。

[0054] 在另一类型的实施例中,本发明是包括缓冲存储器(缓冲器)和耦接至缓冲器的处理子系统的音频处理单元(APU)。其中,APU被耦接成接收包括音频数据分段和元数据分段的编码音频比特流,其中,音频数据分段表示音频数据,并且至少部分元数据分段中的每个包括响度处理状态元数据(LPSM)、缓冲器(如,以非暂时性方式)存储编码音频比特流的至少一个帧,并且处理子系统被配置成从比特流中提取LPSM以及使用LPSM对音频数据执行至少一个自适应响度处理操作。在这种类型中的典型的实施例中,APU是编码器、解码器和后处理器中的一种。

[0055] 在本发明的方法的一些实现中,所生成的音频比特流是AC-3编码比特流、E-AC-3比特流或者杜比E比特流中的一种,包括响度处理状态元数据以及其他元数据(如,DIALNORM元数据参数、动态范围控制元数据参数和其他元数据参数)。在方法的一些其他实现中,所生成的音频比特流是另一类型的编码比特流。

[0056] 本发明的各个方面包括被配置(或编程)成执行本发明的方法的任意实施例的系统或装置、以及(如,以非暂时性方式)存储用于实现本发明的方法或其步骤的任意实施例的代码的计算机可读介质(如,磁盘)。例如,本发明的系统可以是或者包括用软件或固件编程的可编程通用处理器、数字信号处理器或微处理器,和/或被配置成执行对数据的各种操作中的任意操作,包括发明的方法或其步骤的实施例。这样的通用处理器可以是或者包括如下计算机系统:其包括输入装置、存储器和处理电路,其被编程为(和/或被配置成)响应于向其传送的数据来执行本发明的方法(或步骤)的实施例。

附图说明

[0057] 图1是可以被配置成执行本发明的方法的实施例的系统的实施例的框图;

[0058] 图2是作为本发明的音频处理单元的实施例的编码器的框图;

[0059] 图3是作为本发明的音频处理单元的实施例的解码器以及作为本发明的音频处理单元的另一实施例的与解码器耦接的后处理器的框图;

[0060] 图4是AC-3帧的图,包括其被划分成的分段;

[0061] 图5是AC-3帧的同步信息(SI)分段的图,包括其被划分成的分段;

[0062] 图6是AC-3帧的比特流信息(BSI)分段的图,包括其被划分成的分段;以及

[0063] 图7是E-AC-3帧的图,包括其被划分成的分段。

[0064] 符号和命名

[0065] 贯穿本公开,包括在权利要求中,在广义上使用“对”信号或数据执行操作(如,对信号或数据进行滤波、缩放、变换或施加增益)的表述来表示直接对信号或数据、或者对信

号或数据的已处理版本(如,对在对其执行该操作之前已经经历了初步的滤波或预处理的信号的版本)执行该操作。

[0066] 贯穿本公开,包括在权利要求中,在广义上使用表述“系统”来表示装置、系统或子系统。例如,实现解码器的子系统可以被称为解码器系统,包括这样的子系统的系统(如,响应于多个输入生成X个输出信号的系统,其中,子系统生成M个输入,其他X-M个输入从外部源来接收)也可以被称为解码器系统。

[0067] 贯穿本公开,包括在权利要求中,在广义上使用术语“处理器”来表示可编程或者否则可(利用软件或固件)配置以对数据(如,音频或视频或其他图像数据)执行操作的系统或装置。处理器的示例包括现场可编程门阵列(或者其他可配置集成电路或芯片组)、被编程和/或否则被配置成对音频或其他声音数据执行流水线处理的数字信号处理器、可编程通用处理器或计算机、以及可编程微处理器芯片或芯片组。

[0068] 贯穿本公开,包括在权利要求中,表述“音频处理器”和“音频处理单元”被可交换地使用,并且在广义上被用来表示被配置成处理音频数据的系统。音频处理单元的示例包括但不限于编码器(如,转码器)、解码器、编解码器、预处理系统、后处理系统和比特流处理系统(有时被称为比特流处理工具)。

[0069] 贯穿本公开,包括在权利要求中,表述“处理状态元数据”(如,在表述“响度处理状态元数据”中)指与相应的音频数据(也包括处理状态元数据在内的音频数据流的音频内容)分离的且不同的数据。处理状态元数据与音频数据相关联,指示相应的音频数据的响度处理状态(如,已经对音频数据执行了什么类型的处理),并且通常还指示音频数据的至少一个特征或特性。处理状态元数据与音频数据的关联是时间同步的。因此,当前(最近接收的或更新的)处理状态元数据指示相应的音频数据同时包括指示类型的音频数据处理的结果。在某些情况下,处理状态元数据可以包括处理历史和/或用在指示类型的处理中和/或根据指示类型的处理得到的参数中的一些或全部。附加地,处理状态元数据可以包括相应的音频数据的至少一个特征或特性,所述至少一个特征或特性已经根据音频数据计算出或从音频数据中提取到。处理状态元数据还可以包括不是与相应的音频数据的任意处理相关的或者不是从相应的音频数据的任意处理中得到的其他元数据。例如,可以通过具体的音频处理单元来添加第三方数据、乐曲(tracking)信息、标识符、专有权或标准信息、用户注解数据、用户偏好数据等,以传递给其他音频处理单元。

[0070] 贯穿本公开,包括在权利要求中,表述“响度处理状态元数据”(或“LPSM”)表示如下处理状态元数据:其表示相应的音频数据的响度处理状态(如,已经对音频数据执行了什么类型的响度处理),并且通常还表示相应的音频数据的至少一个特征或特性(如,响度)。响度处理状态元数据可以包括不是(即,当其被单独考虑时)响度处理状态元数据的数据(如,其他元数据)。

[0071] 贯穿本公开,包括在权利要求中,使用术语“耦接(couples)”或“被耦接(coupled)”来表示直接连接或间接连接。因此,如果第一装置耦接至第二装置,则该连接可以是直接连接或者是通过其他装置和连接实现的间接连接。

具体实施方式

[0072] 根据本发明的典型的实施例,响度处理状态元数据(LPSM)嵌入在音频比特流的元

数据分段的一个或多个保留的域(或时隙)中,该音频比特流在其他分段(音频数据分段)中也包括音频数据。通常,比特流的每个帧的至少一个分段包括LPSM,该帧的至少一个其他分段包括相应的音频数据(即,由LPSM指示其响度处理状态和响度的音频数据)。在一些实施例中,LPSM的数据量可以充分小以在不影响被分配用于承载音频数据的比特速率的情况下被承载。

[0073] 当两个或更多个音频处理单元需要遍及处理链(或内容生命周期)彼此串联工作时,在音频数据处理链中传送响度处理状态元数据特别有用。在音频比特流中不包括响度处理状态元数据的情况下,例如,当在链中使用两个或更多个音频编解码器并且在比特流的至媒体消耗装置(或者比特流的音频内容的渲染点)的行程期间不止一次施加单端音量调节时,可能出现严重的媒体处理问题,如质量、电平和空间的降级。

[0074] 图1是示例性音频处理链(音频数据处理系统)的框图,其中,可以根据本发明的实施例配置系统的元件中的一个或更多个。该系统包括如所示出地那样耦接在一起的以下元件:预处理单元、编码器、信号分析和元数据校正单元、转码器、解码器和预处理单元。在所示出的系统的变型中,省略了其中一个或多个元件,或者包括附加的音频数据处理单元。

[0075] 在一些实现中,图1的预处理单元被配置成:接受包括音频内容在内的PCM(时域)样本作为输入;以及输出经处理的PCM样本。编码器可以被配置成:接受PCM样本作为输入;以及输出表示音频内容的编码比特流(如,压缩)音频比特流。表示音频内容的比特流的数据有时在本文中被称为“音频数据”。如果编码器根据本发明的典型的实施例来配置,则从编码器输出的音频比特流包括响度处理状态元数据(通常还有其他元数据)以及音频数据。

[0076] 图1的信号分析和元数据校正单元可以接受一个或更多个编码音频比特流作为输入,并且通过执行信号分析来判定(如,验证)在每个编码音频比特流中的处理状态元数据是否正确。如果信号分析和元数据校正单元发现所包括的元数据无效,则其通常用根据信号分析获得的正确的值来替代错误的值。因此,从信号分析和元数据校正单元输出的每个编码音频比特流可以包括已校正(或未校正)处理状态元数据以及编码比特流音频数据。

[0077] 图1的转码器可以接受编码音频比特流作为输入,并且相应地输出已修改(如,不同地编码的)音频比特流(如,通过对输入流进行解码并且以不同的编码格式对解码流进行重新编码)。如果转码器根据本发明的典型的实施例来配置,则从转码器输出的音频比特流包括响度处理状态元数据(通常还有其他元数据)以及编码比特流音频数据。元数据已经被包括在比特流中。

[0078] 图1的解码器可以接受编码(如,压缩)音频比特流作为输入,并且(相应地)输出解码PCM音频样本的流。如果解码器根据本发明的典型的实施例来配置,则典型的操作中的解码器的输出是或者包括以下中的任一个:

[0079] 音频样本的流、以及从输入编码比特流中提取的响度处理状态元数据(通常还有其他元数据)的相应的流;或者

[0080] 音频样本的流、以及根据从输入编码比特流中提取的响度处理状态元数据(通常还有其他元数据)确定的相应的控制比特的流;或者

[0081] 在没有处理状态元数据或根据处理状态元数据确定的控制比特的相应的流的情况下的音频样本的流。在该最后的情况下,解码器可以从输入编码比特流中提取响度处理状态元数据(和/或其他元数据),以及对所提取的元数据执行至少一次操作(如,验证),虽

然其没有输出所提取的元数据或者根据其确定的控制比特。

[0082] 通过根据本发明的典型的实施例来配置图1的后处理单元,后处理单元被配置成接受解码PCM音频样本的流,并且使用与样本一起接收的响度处理状态元数据(通常还有其他元数据)或者与样本一起接收的控制比特(其由解码器根据响度处理状态元数据并且通常还根据其他元数据确定)来对其执行后处理(即,音频内容的音量调节)。后处理单元通常还被配置成对经后处理的音频内容进行渲染以由一个或更多个扬声器回放。

[0083] 本发明的典型的实施例提供了增强型音频处理链,其中,根据由通过音频处理单元分别接收的响度处理状态元数据指示的元数据的同时期的状态,音频处理单元(如,编码器、解码器、转码器、预处理单元和后处理单元)适配待施加给音频数据的它们的各个处理。

[0084] 输入到图1的系统的任意音频处理单元(如,图1的编码器或转码器)的音频数据可以包括响度处理状态元数据(还可选地包括其他元数据)以及音频数据(如,编码音频数据)。根据本发明的实施例,该元数据可以已经通过图1的系统的另一元件(或者图1中未示出的另一源)被包括在输入音频中。接收(具有元数据的)输入音频的处理单元可以被配置成对元数据执行至少一个操作(如,验证)或者响应于元数据执行至少一个操作(如,对输入音频的自适应处理),并且通常还被配置成在其输出音频中包括元数据、元数据的已处理版本或者根据元数据确定的控制比特。

[0085] 本发明的音频处理单元(或音频处理器)的典型实施例被配置成基于由与音频数据对应的响度处理状态元数据指示的音频数据的状态来执行对音频数据的自适应处理。在一些实施例中,自适应处理是(或者包括)响度处理(如果元数据指示还没有对音频数据执行响度处理或者与其类似的处理),并且不是(或者不包括)响度处理(如果元数据指示已经对音频数据执行了这样的响度处理或者与其类似的处理)。在一些实施例中,自适应处理是或者包括元数据验证(如,在元数据验证子单元中执行的元数据验证),以确保音频处理单元基于由响度处理状态元数据指示的音频数据的状态来执行对音频数据的其他自适应处理。在一些实施例中,验证判定与该音频数据相关联(如,包括在具有该音频数据的比特流中)的响度处理状态元数据的可靠性。例如,如果元数据被验证为可靠,则可以重复使用来自某种类型的事先执行的音频处理的结果,并且可以避免同一类型的音频处理的新的执行。另一方面,如果发现元数据已经被篡改(或者不可靠),则可以由音频处理单元来重复据称事先被执行的这种类型的媒体处理(如不可靠的元数据指示的那样),和/或可以由音频处理单元对元数据和/或音频数据执行其他处理。音频处理单元还可以被配置成如果音频处理单元判定处理状态元数据有效(如,基于所提取的密码值与参考密码值的匹配),则向增强型媒体处理链中的下游的其他音频处理单元示意响度处理状态元数据(如,存在于媒体比特流中的响度处理状态元数据)是有效的。

[0086] 图2是作为本发明的音频处理单元的实施例的编码器(100)的框图。编码器100的任意部件或元件都可以用硬件、软件或硬件与软件的组合实现为一个或更多个处理和/或一个或更多个电路(如,ASIC、FPGA或其他集成电路)。编码器100包括如所示地连接的帧缓冲器110、解析器111、解码器101、音频状态验证器102、响度处理级103、音频流选择级104、编码器105、填充器/格式器级107、元数据生成级106、会话响度测量子系统108和帧缓冲器109。通常,编码器100还包括其他处理元件(未示出)。

[0087] 编码器100(作为转码器)被配置成通过使用输入比特流中所包括的响度处理状态

元数据执行自适应和自动响度处理,来将输入音频比特流(例如,可以是AC-3比特流、E-AC-3比特流或杜比E比特流中的一种)转换成包括响度处理状态元数据的编码输出音频比特流(例如,可以是AC-3比特流、E-AC-3比特流或杜比E比特流中的另一种)。例如,编码器100可以被配置成将输入的杜比E比特流(通常用在制作和广播设施而非接收已经向其广播的音频节目的消费者装置中的格式)转换成AC-3或E-AC-3格式的编码输出音频比特流(适于广播给用户装置)。

[0088] 图2的系统还包括编码音频递送子系统150(其存储和/或递送从编码器100输出的编码比特流)和解码器152。从编码器100输出的编码音频比特流可以由子系统150来存储(如,以DVD或蓝光光盘的形式)、或者由子系统150来发送(其可以实现发送链路或网络)、或者可以由子系统150来存储和发送。解码器152被配置成通过从比特流的每个帧中提取响度处理状态元数据(LPSM)、以及生成解码音频数据,来对其通过子系统150接收的(由编码器100生成的)包括响度处理状态元数据的编码音频比特流进行解码。通常,解码器152被配置成使用LPSM对解码音频数据执行自适应响度处理,和/或将解码音频数据和LPSM转发给后处理器,后处理器被配置成使用LPSM对解码音频数据执行自适应响度处理。通常,解码器152包括用于(如,以非暂时性方式)存储从子系统150接收的编码音频比特流的缓冲器。

[0089] 编码器100和解码器152的各种实现可以被配置成执行本发明的方法的不同的实施例。

[0090] 帧缓冲器110是被耦接成接收编码输入音频比特流的缓冲存储器。在操作中,缓冲器110(如,以非暂时性方式)存储编码音频比特流的至少一个帧,并且,编码音频比特流的帧的序列被从缓冲器110向解析器111传送(assert)。

[0091] 解析器111被耦接和配置来从编码输入音频的每个帧中提取响度处理状态元数据(LPSM)和其他元数据,以向音频状态验证器102、响度处理级103、级106和子系统108传送至少LPSM,以从编码输入音频中提取音频数据,以及以向解码器101传送音频数据。编码器100的解码器101被配置成对音频数据进行解码以生成解码音频数据,以及向响度处理级103、音频流选择级104、子系统108以及通常也向状态验证器102传送解码音频数据。

[0092] 状态验证器102被配置成对向其传送的LPSM(通常还有其他元数据)进行认证和验证。在一些实施例中,LPSM是已经被包括在(如,根据本发明的实施例的)输入比特流中的数据块(或者被包括在已经被包括在输入比特流中的数据块中)。该块可以包括用于处理LPSM(以及可选地还有其他元数据)和/或潜在的(从解码器101提供给验证器102的)音频数据的加密散列(基于散列的消息认证代码或“HMAC”)。该数据块在这些实施例中可以以数字形式标记,使得下游音频处理单元可以相对容易地认证和验证处理状态元数据。

[0093] 例如,使用HMAC来生成摘要(digest),并且,包括在本发明的比特流中的保护值可以包括摘要。可以如下针对AC-3帧生成摘要:

[0094] 1. 在AC-3数据和LPSM被编码之后,使用帧数据字节(级联的frame_data#1和frame_data#2)和LPSM数据字节作为用于散列函数HMAC的输入。在计算摘要时不考虑可能存在于auxdata域中的其他数据。这样的其他数据可以是既不属于AC-3数据也不属于LPSM数据的字节。可以在计算HMAC摘要时不考虑LPSM中所包括的保护比特。

[0095] 2. 在计算摘要之后,将其写入被保留用于保护比特的域中的比特流中。

[0096] 3. 完整的AC-3帧的生成的最后的步骤是计算CRC校验。其被写在帧的最末端,并

且,将属于该帧的所有数据都考虑在内,包括LPSM比特。

[0097] 可以将包括但不限于一种或更多种非HMAC加密方法中的任一种的其他加密算法用于LPSM的验证(如,在验证器102中),以确保对于LPSM和/或潜在的音频数据的安全的发送和接收。例如,可以在接收本发明的音频比特流的实施例的每个音频处理单元中执行验证(使用这样的加密方法的验证),以判定包括在比特流中的响度处理状态元数据和相应的音频数据是否已经经历了(由元数据指示的)特定的响度处理(和/或是否已经从特定的响度处理得到)以及是否在这样的特定的响度处理执行之后尚未被修改。

[0098] 状态验证器102向音频流选择级104、元数据生成器106和会话响度测量子系统108传送控制数据,以指示验证操作的结果。响应于控制数据,级104可以选择以下中的任一项(并且将其传送给编码器105):

[0099] 响度处理级103的经自适应处理的输出(如,当LPSM指示从解码器101输出的音频数据尚未经过特定类型的响度处理,并且来自验证器102的控制比特指示LPSM有效时);或者

[0100] 从解码器101输出的音频数据(如,当LPSM指示从解码器101输出的音频数据已经经历了可以由级103来执行的特定类型的响度处理,并且来自验证器102的控制比特指示LPSM有效时)。

[0101] 编码器100的级103被配置成基于由解码器101提取的LPSM所指示的一个或多个音频数据特性,对从解码器101输出的解码音频数据执行自适应响度处理。级103可以是自适应变换域实时响度和动态范围控制处理器。级103可以接收用户输入(如,用户目标响度/动态范围值或dialnorm值)、或者其他元数据输入(如,一种或多种类型的第三方数据、乐曲信息、标识符、专有权或标准信息、用户注解数据、用户偏好数据等)和/或其他输入(如,来自指纹处理的其他输入),并且使用这样的输入来对从解码器101输出的解码音频数据进行处理。

[0102] 当来自验证器102的控制比特指示LPSM无效时,会话响度测量子系统108可以操作以使用如由解码器101提取的LPSM(和/或其他元数据)来确定指示会话(或其他语音)的解码音频(来自解码器101)的分段的响度。当来自验证器102的控制比特指示LPSM有效时当LPSM指示之前确定的解码音频(来自解码器101)的会话(或其他语音)分段的响度时,可以禁止会话响度测量子系统108的操作。

[0103] 存在有用的工具(如,杜比LM100响度仪表),用于方便且容易地测量音频内容中的会话的电平。本发明的APU(如,编码器100的级108)的一些实施例被实现为包括这样的工具(或执行这样的工具的功能)以测量音频比特流的音频内容的平均会话响度(如,从编码器100的解码器101向级108传送的解码AC-3比特流)。

[0104] 如果级108被实现为测量音频数据的真实的平均会话响度,则测量可以包括步骤:隔离音频内容的主要包含语音的分段。接着,根据响度测量算法对主要为语音的音频分段进行处理。对于从AC-3比特流解码的音频数据,该算法可以是标准的K加权的响度测量(根据国际标准ITU-R BS.1770)。可替代地,可以使用其他响度测量(如,基于响度的心理声学模型的响度测量)。

[0105] 语音分段的隔离对于测量音频数据的平均会话响度而言并不是至关重要的。然而,从倾听者的角度来说,其提高了测量的精度并且通常提供更令人满意的结果。由于并非

所有的音频内容都包含会话(语音),所以整个音频内容的响度测量可以提供对音频的会话电平的充分的近似,如果语音出现的话。

[0106] 元数据生成器106生成要被级107包括在编码比特流中以从编码器100输出的元数据。元数据生成器106可以将由解码器101提取的LPSM(和/或其他元数据)传递给级107(如,当来自验证器102的控制比特指示LPSM和/或其他元数据有效时),或者生成新的LPSM(和/或其他元数据)并且向级107传送新的元数据(如,当来自验证器102的控制比特指示由解码器101提取的LPSM和/或其他元数据无效时),或者其可以向级107传送由解码器101提取的元数据与新生成的元数据的组合。元数据生成器106可以在其向级107传送的LPSM中包括由子系统108生成的响度数据以及表示由子系统108执行的响度处理的类型的至少一个值,以包括在要从编码器100输出的编码比特流中。

[0107] 元数据生成器106可以生成保护比特(其可以包括基于散列的消息认证代码或“HMAC”或者由基于散列的消息认证代码或“HMAC”构成),保护比特对于要被包括在编码比特流中的LPSM(可选地还有其他元数据)和/或要被包括在编码比特流中的潜在的音频数据的解密、认证或验证中的至少一个而言是很有用的。元数据生成器106可以向级107提供这样的保护比特,以包括在编码比特流中。

[0108] 在典型的操作中,会话响度测量子系统108对从解码器101输出的音频数据进行处理以响应于其生成响度值(如,选通和未选通会话响度值)以及动态范围值。响应于这些值,元数据生成器106可以生成响度处理状态元数据(LPSM),以(由填充器/格式器107)包括在要从编码器100输出的编码比特流中。

[0109] 附加地,可选地,或者可替代地,编码器100的子系统106和/或108可以执行对音频数据的附加分析以生成表示音频数据的至少一个特性的元数据,以包括在要从级107输出的编码比特流中。

[0110] 编码器105对从选择级104输出的音频数据进行编码(如,通过对其执行压缩),并且向级107传送编码音频,以包括在要从级107输出的编码比特流中。

[0111] 级107对来自编码器105的编码音频和来自生成器106的元数据(包括LPSM)进行复用以生成要从级107输出的编码比特流,优选地使得编码比特流具有本发明的优选实施例指定的格式。

[0112] 帧缓冲器109是(如,以非暂时性方式)存储从级107输出的编码音频比特流的至少一个帧的缓冲存储器,接着,编码音频比特流的帧的序列作为来自编码器100的输出从缓冲器109向递送系统150传送。

[0113] 由元数据生成器106生成的并且由级107包括在编码比特流中的LPSM表示相应的音频数据的响度处理状态(如,已经对音频数据执行了什么类型的响度处理)以及相应的音频数据的响度(如,测量的会话响度、选通的和/或未选通的响度、和/或动态范围)。

[0114] 本文中,对音频数据执行的响度和/或电平测量的“选通(gating)”指超过阈值的计算值被包括在最终的测量中的情况下的具体的电平或响度阈值(如,在最终的测量值中忽略低于-60dBFS的短期响度值)。绝对值的选通指固定的电平或响度,而相对值的选通指取决于当前“非选通(ungated)”测量值的值。

[0115] 在编码器100的一些实现中,被缓冲在存储器109中(并且被输出给递送系统150)的编码比特流是AC-3比特流或E-AC-3比特流,并且包括音频数据分段(如,图4所示的帧的

AB0至AB5分段)和元数据分段,其中,音频数据分段表示音频数据,至少部分元数据分段中的每个包括响度处理状态元数据(LPSM)。级107按以下格式将LPSM插入比特流中。包括LPSM的元数据分段中的每个被包括在比特流的帧的比特流信息(“BSI”)分段的“addbsi”域中,或者被包括在比特流的帧的端部处的auxdata域(如,图4所示的AUX分段)中。比特流的帧可以包括一个或两个元数据分段,其中每个包括LPSM,并且,如果帧包括两个元数据分段,则其中一个存在于帧的addbsi域中,而另一个存在于帧的AUX域中。包括LPSM的每个元数据分段包括具有以下格式的LPSM有效载荷(或容器)分段:

[0116] 首部(通常包括标识LPSM有效载荷的开始同步字,其后跟随至少一个标识值,如下面的表2所示的LPSM格式版本、长度、周期、计数和子流关联值);以及

[0117] 在首部之后,

[0118] 至少一个会话指示值(如,表2的参数“会话通道”),其指示相应的音频数据是指示会话还是不指示会话(如,相应的音频数据的哪个通道指示会话);

[0119] 至少一个响度调节相符值(如,表2的参数“响度调节类型”),其指示相应的音频数据是否与所指示的响度调节的集合相符;

[0120] 至少一个响度处理值(如,表2的参数“会话选通的响度校正标志”、“响度校正类型”中的一个或多个),其表示已经对相应的音频数据执行的至少一种类型的响度处理;以及

[0121] 至少一个响度值(如,表2的参数“ITU相对选通的响度”、“ITU语音选通的响度”、“ITU(EBU 3341)短期3s响度”和“真实峰值”中的一个或多个),其表示相应的音频数据的至少一个响度(如,峰值或者平均响度)特性。

[0122] 在一些实施方式中,由级107插入到比特流的帧的“addbsi”域或者auxdata域中的元数据分段中的每个具有以下格式:

[0123] 核心首部(通常包括标识元数据分段的开始同步字,其后跟随标识值,如下面的表1所示的核心元素版本、长度和周期、扩展元素计数以及子流关联值);以及

[0124] 在核心首部之后的至少一个保护值(如,表1的HMAC摘要和音频指纹值,其对于响度处理状态元数据或相应的音频数据中的至少一个的解密、认证或验证中的至少一个而言是很有用的);以及

[0125] 如果元数据分段包括LPSM,则也在核心首部之后的LPSM有效载荷标识(ID)和LPSM有效载荷尺寸值,其将跟随的元数据标识为LPSM有效载荷并且指示LPSM有效载荷的尺寸。

[0126] (优选地具有上述格式的)LPSM有效载荷(或容器)分段跟随LPSM有效载荷ID和LPSM有效载荷尺寸值。

[0127] 在一些实施例中,帧的auxdata域(或“addbsi”域)中的每个元数据分段具有三层结构:

[0128] 高层结构,包括:表示auxdata(或addbsi)域是否包括元数据的标志;指示存在的是什么类型的元数据的至少一个ID值;以及通常还包括指示存在了多少比特的元数据(如,每种类型的元数据)的值(如果存在元数据)。可能存在的一种类型的元数据是LPSM,可能存在的另一种类型的元数据是媒体研究元数据(如,Nielsen媒体研究元数据);

[0129] 中层结构,包括用于每个标识的类型的元数据的核心元素(如,对于每种标识的类型的元数据如上述类型的核心首部、保护值和LPSM有效载荷ID以及有效载荷尺寸值);以及

[0130] 低层结构,包括用于一个核心元素的每个有效载荷(如,如果核心元素将其标识为存在,则是LPSM有效载荷,和/或,如果核心元素将其标识为存在,则是另一种类型的元数据有效载荷)。

[0131] 可以对在这样的三层结构中的数据值进行嵌套。例如,可以在由核心元素标识的有效载荷之后(从而在核心元素的核心首部之后),包括用于LPSM有效载荷和/或由核心元素标识的另一元数据有效载荷的保护值。在一种示例中,核心首部可以标识LPSM有效载荷与另一元数据有效载荷,用于第一有效载荷(如,LPSM有效载荷)的有效载荷ID和有效载荷尺寸值可以跟随核心首部,第一有效载荷本身可以跟随该ID和尺寸值,用于第二有效载荷的有效载荷ID和有效载荷尺寸值可以跟随第一有效载荷,第二有效载荷本身可以跟随这些ID和尺寸值,并且,两种有效载荷(或者核心元素值和两种有效载荷)的保护值可以跟随最后的有效载荷。

[0132] 在一些实施例中,如果解码器101接收具有加密散列的根据本发明的实施例生成的音频比特流,则解码器被配置成从根据比特流确定的数据块中解析并取回加密散列,上述块包括响度处理状态元数据(LPSM)。验证器102可以使用加密散列来验证所接收的比特流和/或相关联的元数据。例如,验证器102基于参考加密散列与从数据块中取回的加密散列之间的匹配发现LPSM是有效的,则其可以禁止处理器103对于相应的音频数据的操作并且使得选择级104通过(未改变的)音频数据。附加地,可选地,或者可替选地,可以使用其他类型的加密技术来替代基于加密散列的方法。

[0133] 图2的编码器100可以判定(响应于由解码器101提取的LPSM)后处理/预处理单元已经对要被编码的音频数据执行了某种类型的响度处理(在元件105、106和107中),因此可以产生(在生成器106中)响度处理状态元数据,该响度处理状态元数据包括用于事先执行的响度处理和/或从事先执行的响度处理中提取的具体的参数。在一些实现中,编码器100可以产生(并且在从其输出的编码比特流中包括)表示对音频内容的处理历史的处理状态元数据,只要编码器知晓已经对音频内容执行的处理的类型。

[0134] 图3是作为本发明的音频处理单元的实施例的解码器(200)的框图、以及耦接至编码器200的后处理器(300)的框图。后处理器(300)也是本发明的音频处理单元的实施例。解码器200和后处理器300的任意部件或元件都可以用硬件、软件或硬件与软件的组合实现为一个或更多个处理和/或一个或更多个电路(如,ASIC、FPGA或其他集成电路)。解码器200包括如所示地连接的帧缓冲器210、解析器205、音频解码器202、音频状态验证级(验证器)203和控制比特生成级204。通常,解码器200还包括其他处理元件(未示出)。

[0135] 帧缓冲器201(缓冲存储器)(以非暂时性方式)存储由解码器200接收的编码音频比特流的至少一个帧。从缓冲器201向解析器205传送编码音频比特流的帧的序列。

[0136] 解析器205被耦接和配置来从编码输入音频的每个帧中提取响度处理状态元数据(LPSM)和其他元数据、向音频状态验证器203和级204传送至少LPSM、将LPSM作为输出(如,给后处理器300的输出)传送、从编码输入音频中提取音频数据、以及向解码器202传送所提取的音频数据。

[0137] 输入给解码器200的编码音频比特流可以是AC-3比特流、E-AC-3比特流或杜比E比特流中的一种。

[0138] 图3的系统还包括后处理器300。后处理器300包括帧缓冲器301和其他处理元件

(未示出),包括耦接至缓冲器301的至少一个处理元件。帧缓冲器301(如,以非暂时性方式)存储由后处理器300从解码器200接收的解码音频比特流的至少一个帧。后处理器300的处理元件被耦接和配置来使用从解码器202输出的元数据(包括LPSM值)和/或从解码器200的级204输出的控制比特,来接收并且自适应地处理从缓冲器301输出的解码音频比特流的帧的序列。通常,后处理器300被配置成(如,基于由LPSM指示的响度处理状态和/或一个或更多个音频数据特性)使用LPSM值来对解码音频数据执行自适应响度处理。

[0139] 解码器200和后处理器300的各种实现都被配置成执行本发明的方法的不同的实施例。

[0140] 解码器200的音频解码器202被配置成对由解析器205提取的音频数据进行解码以生成解码音频数据,并且被配置成将解码音频数据作为输出(如,至后处理器300的输出)传送。

[0141] 状态验证器203被配置成对向其传送的LPSM(通常还有其他元数据)进行认证和验证。在一些实施例中,LPSM是已经被包括在(如,根据本发明的实施例的)输入比特流中的数据块(或者被包括在已经被包括在输入比特流中的数据块中)。该块可以包括用于处理LPSM(以及可选地还处理其他元数据)和/或潜在的音频数据(从解析器205和/或解码器202提供给验证器203)的加密散列(基于散列的消息认证代码或“HMAC”)。该数据块在这些实施例中可以用数字标记,使得下游音频处理单元可以相对容易地认证和验证处理状态元数据。

[0142] 可以将包括但不限于一种或更多种非HMAC加密方法中的任一种的其他加密算法用于LPSM的验证(如,在验证器203中),以确保对于LPSM和/或潜在的音频数据的安全的发送和接收。例如,可以在接收本发明的音频比特流的实施例的每个音频处理单元中执行验证(使用这样的加密方法的验证),以判定包括在比特流中的响度处理状态元数据和相应的音频数据是否已经经历了(由元数据指示的)特定的响度处理(和/或是否已经从特定的响度处理得到)以及是否在这样的特定的响度处理执行之后尚未被修改。

[0143] 状态验证器203向控制比特生成器204传送控制数据,和/或传送控制数据作为输出(如,至后处理器300的输出),以指示验证操作的结果。响应于控制数据(以及可选地也响应于从输入比特流中提取的其他元数据),级204可以生成以下中的任一项(并且将其传送给后处理器300):

[0144] 指示从解码器202输出的解码音频数据已经经历了特定类型的响度处理的控制比特(当LPSM指示从解码器202输出的音频数据已经经历了特定类型的响度处理,并且来自验证器203的控制比特指示LPSM有效时);或者

[0145] 指示从解码器202输出的解码音频数据应当经历特定类型的响度处理的控制比特(如,当LPSM指示从解码器202输出的音频数据尚未经历特定类型的响度处理时,或者当LPSM指示从解码器202输出的音频数据已经经历了特定类型的响度处理而来自验证器203的控制比特指示LPSM无效时)。

[0146] 可替代地,解码器200向后处理器300传送由解码器202从输入比特流中提取的LPSM(以及任意其他元数据),并且,后处理器300使用LPSM对解码音频数据执行响度处理,或者执行LPSM的验证,接着,如果验证指示LPSM有效,则使用LPSM对解码音频数据执行响度处理。

[0147] 在一些实施例中,如果解码器201来接收具有加密散列的根据本发明的实施例生

成的音频比特流,则解码器被配置成从根据比特流确定的数据块来解析和取回加密散列,上述块包括响度处理状态元数据(LPSM)。验证器203可以使用加密散列来验证所接收的比特流和/或相关联的元数据。例如,如果验证器203基于参考加密散列与从数据块中取回的加密散列之间的匹配发现LPSM是有效的,则其可以向下游音频处理单元(如,后处理器300,其可以是或者包括音量调节单元)发信号,以传递(未改变)比特流的音频数据。附加地,可选地,或者可替代地,可以使用其他类型的加密技术来替代基于加密散列的方法。

[0148] 在解码器200的一些实现中,所接收(并且被缓冲在存储器201中)的编码比特流是AC-3比特流或E-AC-3比特流,并且包括音频数据分段(如,图4所示的帧的AB0至AB5分段)和元数据分段,其中,音频数据分段表示音频数据,至少部分元数据分段中的每个包括响度处理状态元数据(LPSM)。解码器级202被配置成从比特流中提取具有以下格式的LPSM。包括LPSM的元数据分段中的每个被包括在比特流的帧的比特流信息(“BSI”)分段的“addbsi”域中,或者被包括在比特流的帧的端部处的auxdata域(如,图4所示的AUX分段)中。比特流的帧可以包括一个或两个元数据分段,其中每个包括LPSM,并且,如果帧包括两个元数据分段,则其中一个存在于帧的addbsi域中,而另一个存在于帧的AUX域中。包括LPSM的每个元数据分段包括具有以下格式的LPSM有效载荷(或容器)分段:

[0149] 首部(通常包括标识LPSM有效载荷的开始同步字,其后跟随标识值,如下面的表2所示的LPSM格式版本、长度、周期、计数和子流关联值);以及

[0150] 在首部之后,

[0151] 至少一个会话指示值(如,表2的参数“会话通道”),其指示相应的音频数据是指示会话还是不指示会话(如,相应的音频数据的哪个通道指示会话);

[0152] 至少一个响度调节相符值(如,表2的参数“响度调节类型”),其指示相应的音频数据是否与指示的响度调节的集合相符;

[0153] 至少一个响度处理值(如,表2的参数“会话选通的响度校正标志”、“响度校正类型”中的一个或多个),其指示已经对相应的音频数据执行的至少一种类型的响度处理;以及

[0154] 至少一个响度值(如,表2的参数“ITU相对选通的响度”、“ITU语音选通的响度”、“ITU(EBU 3341)短期3s响度”和“真实峰值”中的一个或多个),其指示相应的音频数据的至少一个响度(如,峰值或者平均响度)特性。

[0155] 在一些实施中,解码器级202被配置成从比特流的帧的“addbsi”域或auxdata域中提取具有以下格式的每个元数据分段:

[0156] 核心首部(通常包括标识元数据分段的开始的同步字,其后跟随至少一个标识值,如下面的表1中所示的核心元素版本、长度和周期、扩展元素计数以及子流关联值);以及

[0157] 在核心首部之后的至少一个保护值(如,表1的HMAC摘要和音频指纹值),其对于响度处理状态元数据或相应的音频数据中的至少一个的解密、认证或验证中的至少一个而言是很有用的;以及

[0158] 如果元数据分段包括LPSM,则也在核心首部之后的LPSM有效载荷标识(ID)和LPSM有效载荷尺寸值,其将以下元数据标识为LPSM有效载荷并且指示LPSM有效载荷的尺寸。

[0159] (优选地具有上述格式的)LPSM有效载荷(或容器)分段跟随LPSM有效载荷ID和LPSM有效载荷尺寸值。

[0160] 更一般地,由本发明的优选实施例生成的编码音频比特流具有如下结构:其向标签元数据元素和子元素提供机制作为核心(强制)或扩展(可选元素)。这使得比特流(包括其元数据)的数据速率能够跨大量的应用来缩放。优选的比特流句法的核心(强制)元素应当能够发信号告知与音频内容相关联的扩展(可选)元素存在(在带内)和/或在远处(在带外)。

[0161] 需要核心元素存在于比特流的每个帧中。核心元素的一些子元素是可选的并且可以以任意组合存在。扩展元素不需要存在于每个帧中(以防止比特速率过高)。因此,扩展元素可以存在于某些帧中而不存在于其他帧中。扩展元素的某些子元素是可选的并且可以以任意组合存在,而扩展元素的某些子元素可以是强制的(即,如果扩展元素存在于比特流的帧中)。

[0162] 在一类实施例中,(如,由实施本发明的音频处理单元)生成包括音频数据分段和元数据分段的序列的编码音频比特流。音频数据分段表示音频数据,元数据分段中的至少某些中的每个包括响度处理状态元数据(LPSM),音频数据分段与元数据分段时分复用。在这种类型的优选实施例中,每个元数据分段具有要在本文中描述的优选格式。

[0163] 在一种优选格式中,编码比特流是AC-3比特流或E-AC-3比特流,并且,包括LPSM的每个元数据分段作为附加比特流信息被(如,编码器100的优选实现的级107)包括在比特流的帧的比特流信息(“BSI”)分段的“adbsi”域(如图6所示)中,或者被包括在比特流的帧的auxdata域中。

[0164] 在该优选格式中,每个帧在帧的adbsi域中包括具有如以下表1所示的格式的核心元素:

[0165] 表1

参数	描述	强制/可选
SYNC[ID]		强制

[0167]

核心元素版本		强制
核心元素长度		强制
核心元素周期 (XXX)		强制
扩展元素计数	指示与核心元素相关联的扩展元数据元素的数量。该值可以随着比特流从产生通过分发和最终的发射而传递来递增/递减。	强制
子流关联	描述核心元素与哪个或哪些子流相关联。	强制
签名 (HMAC 摘要)	根据整个帧的音频数据、核心元素和所有的扩展元素计算的 256 比特的 HMAC 摘要(使用 SHA-2 算法)。	强制
PGM 边界倒数计数	仅对于在音频节目文件/流的首部或尾部处的一些帧出现的域。因此,可以使用核心元素版本变化来示意该参数的包括。	可选
音频指纹	根据用核心元素周期域表示的一些 PCM 音频样本得到的音频指纹。	可选
视频指纹	根据用核心元素周期域表示的一些压缩视频样本(如果有)得到的视频指纹。	可选
URL/UUID	该域被定义用于承载参考附加节目内容(要素)	可选

[0168]

	的外部位置和/或与比特流相关联的元数据的URL和/或UUID(其对于指纹而言可以是冗余的)。	
--	--	--

[0169] 在优选格式中,包含LPSM的每个addbsi (或auxdata) 域包含核心首部(可选地还有附加核心元素),并且在核心首部(或者核心首部和其他核心元素)之后,还包含以下LPSM值(参数):

[0170] 在核心元素值之后跟随的有效载荷ID(将元数据标识为LPSM) (如表1所示);

[0171] 在有效载荷ID之后跟随的有效载荷尺寸(指示LPSM有效载荷的尺寸);以及

[0172] 具有如下面的表(表2) 所示的格式的LPSM数据(跟随有效载荷ID和有效载荷尺寸值):

[0173] 表2

[0174]

LPSM 参数[智能响应度]	描述	唯一状态的数量	强制/可选	插入速率(参数的更新周期)
LPSM 版本			强制	
LPSM 周期(XXX)	仅适用于XXX域		强制	
LPSM 计数			强制	
LPSM 子流关联			强制	
会话通道	表示在之前	8	强制	~0.5 秒

[0175]

	的 0.5 秒期间 L、C 与 R 音频通道的哪个组合包含语音。当 L、C 或 R 的任意组合中不存在语音时，该参数会表示“无会话”			(通常)
响度调节类型	表示相关联的音频数据流符合规则的特定集合 (如 , ATSC A/85 或 EBU R128)	8	强制	帧
会话选通的响度校正标志	指示是否已经基于会话选通对相关联的音频流进行了校正	2	可选 (仅在 Loudness_Regulation_Type 指示相应的音频“未被校正”的情况下存在)	帧
响度校正类型	指示是否已经用无限的预测未来 (基于文件) 或用实时 (RT) 响度和动态范围控制器对相关联的音	2	可选 (仅在 Loudness_Regulation_Type 指示相应的音频“未被校正”的情况下存在)	帧

[0176]

	频流进行了校正			
ITU 有关的选通的响度(INF)	指示没有应用元数据的相关联的音频流的ITU-R BS.1770-2整体的响度(如, 7比特: -58->+5.5 LKFS, 0.5 LKFS 步长)	128	可选	1 秒
ITU 语音选通的响度(INF)	指示没有应用元数据的相关联的音频流的ITU-R BS.1770-2整体的响度(如, 7比特: -58->+5.5 LKFS, 0.5 LKFS 步长)	128	可选	1 秒
ITU (EBU 3341) 短期 3s 响度	指示 @ ~10Hz 的插入速率的没有应用元数据(滑动窗)	256	可选	0.1 秒

[0177]		的相关联的 音频流的 3 秒的未选通 ITU 响 度 (如, 8 比 特 : 116->+11.5 LKFS, 0.5 LKFS 步 长)			
	真实峰值	指示没有应用元数据的 相关联的音 频 流 的 ITU-R BS.1770 附 件 2 真实峰 值 (dB TP) (即, 在元 素周期域中 示意的帧周 期上的最大 值) 116->+11.5 LKFS, 0.5 LKFS 步长	256	可选	0.5 秒

[0178] 在根据本发明生成的编码比特流的另一优选格式中,比特流是AC-3比特流或E-AC-3比特流,并且,包括LPSM的每个元数据分段(如,通过编码器100的优选实现的级107)被包括在以下任一个中:比特流的帧的比特流信息(“BSI”)分段的“addbsi”域(如图6所示);或者比特流的帧的端部处的auxdata域(如,图4所示的AUX分段)。帧可以包括一个或两个元数据分段,每个元数据分段包括LPSM,并且,如果帧包括两个元数据分段,则其中一个元数据分段存在于帧的addbsi域中,而另一个存在于帧的AUX域中。包括LPSM的每个元数据分段具有以上参考表1和表2指出的格式(即,其包括:表1所示的核心元素,其后跟随有效载荷ID(将元数据标识为LPSM);以及上述有效载荷尺寸值,其后跟随有效载荷(具有表2所示的格式的LPSM数据))。

[0179] 在另一优选格式中,编码比特流是杜比E比特流,并且,每个包括LPSM的元数据分段是杜比E保护带间隔的前N个样本位置。包括这样的元数据分段(其包括LPSM)的杜比E比特流优选地包括表示在SMPTE337M前序的Pd字中示意的LPSM有效载荷长度的值(SMPTE

337M Pa字重复速率优选地保持与相关联的视频帧速率一致)。

[0180] 在其中编码比特流是E-AC-3比特流的优选格式中,每个包括LPSM的元数据分段作为附加比特流信息(如,通过编码器100的优选实现的级107)被包括在比特流的帧的比特流信息(“BSI”)分段的“addbsi”域中。接下来描述具有这种优选格式的LPSM的E-AC-3比特流的编码的另外的方面:

[0181] 1.在E-AC-3比特流的生成期间,当E-AC-3编码器(其将LPSM值插入到比特流中)为“有效”时,对于所生成的每个帧(同步帧),比特流应当包括承载在帧的addbsi域中的元数据块(包括LPSM)。承载元数据块所需的比特不应当增加编码器比特速率(帧长度);

[0182] 2.每个元数据块(包含LPSM)应当包含以下信息:

[0183] 1.loudness_correction_type_flag:其中,“1”指示从编码器的上游对相应的音频数据的响度进行校正,“0”指示由嵌入在编码器中的响度校正器(如,图2的编码器100的响度处理器103)来对响度进行校正;

[0184] 2.speech_channel:指示哪个源通道包含语音(在之前的0.5秒内)。如果没有检测到任何语音,则应这样指示这一点;

[0185] 3.speech_loudness:指示包含语音的每个相应的音频通道的整体的语音响度(在之前的0.5秒内);

[0186] 4.ITU_loudness:指示每个相应的音频信道的整体的ITUBS.1770-2响度;

[0187] 5.gain:用于解码器中的反转的响度合成增益(以示出反转性);

[0188] 3.当E-AC-3编码器(其将LPSM值插入到比特流中)为“有效”并且接收具有“信任”标志的AC-3帧时,编码器中的响度控制器(如,图2的编码器100的响度处理器103)应当被旁路。“信任的”源dialnorm和DRC值应当被传递(如,通过编码器100的生成器106)给E-AC-3编码器部件(如,编码器100的级107)。LPSM块生成继续,并且loudness_correction_type_flag被设置成“1”。响度控制器旁路序列必须与出现“信任”标志的解码AC-3帧的开始同步。响度控制器旁路序列应当如下来实现:在10个音频块周期(即,53.3毫秒)内从值9到值0递减leveler_amount控制,并且将leveler_back_end_meter控制置于旁路模式下(该操作应当导致无缝过渡)。术语校准器的“信任的”旁路表示源比特流的dialnorm值也在编码器的输出端处被重新使用。(即,如果“信任的”源比特流的dialnorm值为-30,则编码器的输出应当将-30用于输出的dialnorm值);

[0189] 4.当E-AC-3编码器(其将LPSM值插入到比特流中)为“有效”并且接收没有“信任”标志的AC-3帧时,嵌入在编码器中的响度控制器(如,图2的编码器100的响度处理器103)应当是有效的。LPSM块生成继续,并且loudness_correction_type_flag被设置为“0”。响度控制器激活序列应当与“信任”标志消失的解码AC-3帧的开始同步。响度控制器激活序列应当如下来实现:在1个音频块周期(即,5.3毫秒)内从值0到值9递增leveler_amount控制,并且将leveler_back_end_meter控制置于“有效”模式下(该操作应当导致无缝过渡并且包括back_end_meter整体重置);以及

[0190] 5.在编码期间,图形用户接口(GUI)应当向用户指明以下参数:“输入音频节目:[信任的/不信任的]”——该参数的状态基于输入信号内的“信任”标志的存在;以及“实时响度校正:[使能/禁止]”——该参数的状态基于嵌入在编码器中的该响度控制器是否有效。

[0191] 当对具有包括在比特流的每个帧的比特流信息 (“BSI”) 分段的 “addbsi” 域中的 LPSM (优选格式的 LPSM) 的 AC-3 或 E-AC-3 比特流进行解码时, 解码器应当解析 LPSM 块数据 (addbsi 域中的 LPSM 块数据) 并且将所有的所提取的 LPSM 值传递给图形用户界面 (GUI)。每帧对所提取的 LPSM 值的组进行更新。

[0192] 在根据本发明生成的编码比特流的另一优选格式中, 编码比特流是 AC-3 比特流或 E-AC-3 比特流, 并且, 包括 LPSM 的每个元数据分段作为附加比特流信息 (如, 通过编码器 100 的优选实现的级 107) 被包括在比特流的帧的比特流信息 (“BSI”) 分段的 “addbsi” 域 (如图 6 所示) 中 (或 AUX 分段中)。在该格式 (其是以上参考表 1 和表 2 所述的格式的变型) 中, 每个包含 LPSM 的 addbsi (或 AUX) 域包含以下 LPSM 值:

[0193] 表 1 所示的核心元素, 其后跟随有效载荷 ID (将元数据标识为 LPSM) 和有效载荷尺寸值, 其后跟随具有以下格式的有效载荷 (LPSM 数据) (类似于上述表 2 所示的强制元素):

[0194] LPSM 有效载荷的版本: 指示 LPSM 有效载荷的版本的 2 比特域;

[0195] dialchan: 指示相应的音频数据的左通道、右通道和/或中心通道是否包含口语会话的 3 比特域。Dialchan 域的比特分配可以如下: 指示左通道中的会话的存在的比特 0 被存储在 dialchan 域的最高有效比特中; 指示中心通道中的会话的存在的比特 2 被存储在 dialchan 域的最低有效比特中。如果相应的通道在节目的在前的 0.5 秒期间包含口语会话, 则 dialchan 域的每个比特被设置为 “1”。

[0196] loudregtyp: 指示节目响度符合哪个响度调节标准的 3 比特域。将 “loudregtyp” 域设置为 “000” 指示 LPSM 不指示响度调节相符性。例如, 该域的一个值 (如, 000) 可以指示与响度调节标准的相符性没有被示出, 该域的另一个值 (如, 001) 可以指示节目的音频数据符合 ATSC A/85 标准, 该域的另一个值 (如, 010) 可以指示节目的音频数据符合 EBU R128 标准。在示例中, 如果域被设置为除了 “000” 以外的任意值, 则 loudcorrdialgat 域和 loudcorrtyp 域应当在有效载荷中跟随。

[0197] loudcorrdialgat: 指示是否已经应用了会话选通的响度校正的 1 比特域。如果已经使用会话选通对节目的响度进行了校正, 则 loudcorrdialgat 域的值被设置为 “1”。否则, 其被设置为 “0”。

[0198] loudcorrtyp: 指示被应用于节目的响度校正的类型的 1 比特域。如果已经用无限的预测未来 (基于文件) 响度校正处理对节目的响度进行了校正, 则 loudcorrtyp 域的值被设置为 “0”。如果已经使用实时响度测量和动态范围控制的组合对节目的响度进行了校正, 则该域的值被设置为 “1”;

[0199] loudrelgate: 指示是否存在有关的选通响度数据 (ITU) 的 1 比特域。如果 loudrelgate 域被设置为 “1”, 则 7 比特的 ituloudrelgat 域应当在有效载荷中跟随;

[0200] loudrelgat: 指示有关的选通节目响度 (ITU) 的 7 比特域。该域指示由于所应用的 dialnorm 和动态范围压缩、在没有任何增益调节的情况下根据 ITU-R BS.1770-2 测量的音频节目的整体的响度。0 至 127 的值被理解为 -58LKFS 至 +5.5LKFS, 步长为 0.5LKFS;

[0201] loudspchgate: 指示语音选通的响度数据 (ITU) 是否存在的 1 比特域。如果 loudspchgate 域被设置为 “1”, 则 7 比特的 loudspchgat 域应当在有效载荷中跟随;

[0202] loudspchgat: 指示语音选通的节目响度的 7 比特域。该域指示由于所应用的 dialnorm 和动态范围压缩、在没有任何增益调节的情况下根据 ITU-R BS.1770-3 的公式 (2)

测量的整个相应的音频节目的整体的响度。0至127的值被理解为-58LKFS至+5.5LKFS,步长为0.5LKFS;

[0203] loudstrm3se:指示是否存在短期(3秒)响度数据的1比特域。如果该域被设置为“1”,则7比特的loudstrm3s域应当在有效载荷中跟随;

[0204] loudstrm3s:指示由于所应用的dialnorm和动态范围压缩、在没有任何增益调节的情况下根据ITU-R BS.1771-1测量的相应的音频节目的在前的3秒的未选通的响度的7比特域。0至256的值被理解为-116LKFS至+11.5LKFS,步长为0.5LKFS;

[0205] truepke:指示是否存在真实的峰值响度数据的1比特域。如果truepke域被设置为“1”,则8比特的truepk域应当在有效载荷中跟随;以及

[0206] truepk:指示由于所应用的dialnorm和动态范围压缩、在没有任何增益调节的情况下根据ITU-R BS.1770-3的附件2测量的的节目的真实的峰值样本值的8比特域。0至256的值被理解为-116LKFS至+11.5LKFS,步长为0.5LKFS。

[0207] 在一些实施例中,AC-3比特流或E-AC-3比特流的帧的auxdata域(或“addbsi”域)中的元数据分段的核心元素包括核心首部(通常包括标识值,如核心元素版本),并且,在核心首部之后,包括:指示指纹数据(或者其他保护值)是否被包括用于元数据分段的元数据的值、指示是否存在(与元数据分段的元数据所对应的音频数据有关的)外部数据的值、用于由核心元素来标识的每种类型的元数据(如,LPSM和/或除了LPSM之外的其他类型的元数据)的有效载荷ID和有效载荷尺寸值、以及用于由核心元素来标识的至少一种类型的元数据的保护值。元数据分段的元数据有效载荷跟随核心首部,并且(在某些情况下)嵌套在核心元素的值内。

[0208] 本发明的实施例可以用硬件、固件或软件或者其组合(如,作为可编程逻辑阵列)来实现。除非特别指出,否则被包括作为本发明的一部分的算法或处理并非固有地与任意具体的计算机或其他装置有关。具体地,各种通用机器可以与根据本文中的教导写出的程序一起使用,或者,构造更专门的设备(如,集成电路)来执行所需要的方法步骤可能更为方便。因此,本发明可以在一个或更多个可编程计算机系统上执行的一个或更多个计算机程序中实现(如,图1的元件、或图2的编码器100(或其元件)、或图3的解码器200(或其元件)或图3的后处理器300(或其元件)中的任一个的实现),每个可编程计算机系统包括至少一个处理器、至少一个数据存储系统(包括易失性和非易失性存储器和/或存储元件)、至少一个输入装置或端口、以及至少一个输出装置或端口。应用节目代码,以输入数据,从而执行本文中所描述的功能并且生成输出信息。输出信息以公知的方式应用于一个或更多个输出装置。

[0209] 可以用任意期望的计算机语言(包括机器、汇编或高层过程、逻辑或面向对象编程语言)来实现每个这样的程序,以与计算机系统通信。在任何情况下,语言可以是编译语言或解释语言。

[0210] 例如,当用计算机软件指令序列来实现时,可以用在合适的数字信号处理硬件中运行的多线程软件指令序列来实现本发明的实施例的各种功能和步骤,在这种情况下,实施例的各种装置、步骤和功能可以对应于软件指令的一部分。

[0211] 每个这样的计算机程序优选地存储在或者下载到由通用或专用可编程计算机可读的存储介质或装置(如固态存储器或媒体、或者磁性或光学媒体),以在计算机系统读取

存储介质和装置以执行本文中所述的过程时,对计算机进行配置和操作。本发明的系统还可以实现为计算机可读的存储介质,其配置有(即,存储)计算机程序,其中,这样配置的存储介质使得计算机系统以具体和预定的方式操作以执行本文中所述的功能。

[0212] 已经描述了本发明的大量实施例。然而,应当理解,可以在不偏离本发明的精神和范围的情况下作出各种修改。鉴于以上教导,对本发明的大量修改和变化是可能的。应当理解,可以在所附权利要求的范围内,以除了如本文中具体地描述的方式之外的方式来实践本发明。

[0213] 附记

[0214] 1. 一种音频处理单元,包括:

[0215] 缓冲存储器;以及

[0216] 至少一个处理子系统,所述至少一个处理子系统耦接至所述缓冲存储器,其中,所述缓冲存储器存储编码音频比特流的至少一个帧,所述帧在所述帧的至少一个元数据分段中包括响度处理状态元数据,并且在所述帧的至少一个其他分段中包括音频数据,其中,所述处理子系统被耦接和配置为执行所述比特流的生成、对所述比特流的解码、或者使用所述比特流的响度处理状态元数据对所述比特流的音频数据进行的自适应响度处理中的至少一个,或者使用所述比特流的响度处理状态元数据来执行对所述比特流的音频数据或响度处理状态元数据中的至少一个的认证或验证中的至少一个,

[0217] 其中,所述编码音频比特流包括帧的序列,所述帧中的每个包括至少一个元数据分段和至少一个音频分段,每个所述音频数据分段表示音频数据,至少一个所述元数据分段包括含有响度处理状态元数据的响度处理状态元数据有效载荷,所述响度处理状态元数据有效载荷包括:

[0218] 首部;以及

[0219] 在所述首部之后的:至少一个会话指示值,所述至少一个会话指示值表示相应的音频数据是指示会话还是不指示会话;至少一个响度调节相符值,所述至少一个响度调节相符值指示所述相应的音频数据是否与所指示的响度调节的集合相符;至少一个响度处理值,所述至少一个响度处理值表示已经对所述相应的音频数据执行的至少一种类型的响度处理;以及至少一个响度值,所述至少一个响度值表示所述相应的音频数据的至少一个响度特性。

[0220] 2. 根据附记1所述的音频处理单元,其中,每个包括一个所述响度处理状态元数据有效载荷的所述元数据分段包括:

[0221] 核心首部;

[0222] 在所述核心首部之后的至少一个保护值,所述至少一个保护值用于所述响度处理状态元数据或者所述相应的音频数据中的至少一个的解密、认证或验证中的至少一个;以及

[0223] 在所述核心首部之后的响度处理状态元数据标识和有效载荷尺寸值,其中,所述响度处理状态元数据有效载荷在所述响度处理状态元数据有效载荷标识和有效载荷尺寸值之后。

[0224] 3. 根据附记2所述的音频处理单元,其中,所述核心首部包括标识所述元数据分段的开始的同步字以及跟随所述同步字的至少一个标识值,并且所述响度处理状态元数据有

效载荷的所述首部包括至少一个标识值。

[0225] 4. 根据附记1所述的音频处理单元,其中,所述编码音频比特流是AC-3比特流或者E-AC-3比特流,所述编码音频比特流的所述帧中的每个包括比特流信息分段和辅助数据auxdata域,所述比特流信息分段包括附加比特流信息adbsi域,并且每个包括响度处理状态元数据有效载荷在内的所述元数据分段被包括在一个所述adbsi域或者一个所述auxdata域中。

[0226] 5. 根据附记1所述的音频处理单元,其中,所述编码音频比特流是AC-3比特流、E-AC-3比特流或者杜比E比特流中的一种。

[0227] 6. 根据附记1所述的音频处理单元,其中,所述缓冲存储器以非暂时性方式存储所述帧。

[0228] 7. 根据附记1所述的音频处理单元,其中,所述音频处理单元是编码器。

[0229] 8. 根据附记7所述的音频处理单元,其中,所述编码器包括:

[0230] 解码子系统,所述解码子系统被配置成接收输入音频比特流以及从所述输入音频比特流中提取输入响度处理状态元数据和输入音频数据;

[0231] 自适应响度处理子系统,所述自适应响度处理子系统被耦接和配置为使用所述输入响度处理状态元数据对所述输入音频数据执行自适应响度处理,从而生成已处理音频数据;以及

[0232] 编码子系统,所述编码子系统被耦接和配置为:通过在所述编码音频比特流中包括所述响度处理状态元数据、或者所述响度处理状态元数据的更新版本,来响应于所述已处理音频数据生成包括所述响度处理状态元数据、或者所述响度处理状态元数据的更新版本在内的所述编码音频比特流;以及将所述编码音频比特流送往所述缓冲存储器。

[0233] 9. 根据附记7所述的音频处理单元,其中,所述解码子系统包括状态验证级,所述状态验证级被耦接和配置为认证和验证所述响度处理状态元数据。

[0234] 10. 根据附记1所述的音频处理单元,其中,所述音频处理单元是解码器。

[0235] 11. 根据附记10所述的音频处理单元,其中,所述处理子系统是解码子系统,所述解码子系统耦接至所述缓冲存储器并且被配置成从所述编码音频比特流中提取所述响度处理状态元数据,并且所述解码器还包括状态验证级,所述状态验证级耦接至所述解码子系统并且被配置成认证和验证所述响度处理状态元数据。

[0236] 12. 根据附记1所述的音频处理单元,包括:

[0237] 解码器,所述解码器耦接至所述缓冲存储器并且被配置成从所述编码音频比特流中提取所述响度处理状态元数据和所述音频数据;以及

[0238] 后处理器,所述后处理器耦接至所述解码器并且被配置成使用从所述编码音频比特流中提取的所述响度处理状态元数据来对从所述编码音频比特流中提取的所述音频数据执行自适应响度处理。

[0239] 13. 根据附记1所述的音频处理单元,其中,所述音频处理单元是数字信号处理器。

[0240] 14. 根据附记1所述的音频处理单元,其中,所述音频处理单元是预处理器,所述预处理器被配置成:从所述编码音频比特流中提取所述响度处理状态元数据和所述音频数据;以及使用从所述编码音频比特流中提取的所述响度处理状态元数据来对从所述编码音频比特流中提取的所述音频数据执行自适应响度处理。

[0241] 15.一种音频处理单元,包括:

[0242] 至少一个输入端,所述至少一个输入端被耦接成接收输入音频比特流;以及

[0243] 处理子系统,所述处理子系统耦接至所述至少一个输入端并且被配置成响应于所述输入音频比特流来生成包括音频数据分段和元数据分段的编码音频比特流,其中,所述编码音频比特流包括帧的序列,所述帧中的每个包括至少一个元数据分段和至少一个音频分段,每个所述音频数据分段表示音频数据,至少一个所述元数据分段包括含响度处理状态元数据的响度处理状态元数据有效载荷,所述响度处理状态元数据有效载荷包括:

[0244] 首部;以及

[0245] 在所述首部之后的:至少一个会话指示值,所述至少一个会话指示值表示相应的音频数据是指示会话还是不指示会话;至少一个响度调节相符值,所述至少一个响度调节相符值指示所述相应的音频数据是否与所指示的响度调节的集合相符;至少一个响度处理值,所述至少一个响度处理值表示已经对所述相应的音频数据执行的至少一种类型的响度处理;以及至少一个响度值,所述至少一个响度值表示所述相应的音频数据的至少一个响度特性。

[0246] 16.根据附记15所述的音频处理单元,其中,每个包括一个所述响度处理状态元数据有效载荷的所述元数据分段包括:

[0247] 核心首部;

[0248] 在所述核心首部之后的至少一个保护值,所述至少一个保护值用于所述响度处理状态元数据或者所述相应的音频数据中的至少一个的解密、认证或验证中的至少一个;以及

[0249] 在所述核心首部之后的响度处理状态元数据标识和有效载荷尺寸值,其中,所述响度处理状态元数据有效载荷在所述响度处理状态元数据有效载荷标识和有效载荷尺寸值之后。

[0250] 17.根据附记16所述的音频处理单元,其中,所述核心首部包括标识所述元数据分段的开始的同步字以及跟随所述同步字的至少一个标识值,并且所述响度处理状态元数据有效载荷的所述首部包括至少一个标识值。

[0251] 18.根据附记15所述的音频处理单元,其中,所述编码音频比特流是AC-3比特流或者E-AC-3比特流,所述编码音频比特流的所述帧中的每个包括比特流信息分段和辅助数据auxdata域,所述比特流信息分段包括附加比特流信息adbsi域,并且每个包括响度处理状态元数据有效载荷在内的所述元数据分段被包括在一个所述adbsi域或者一个所述auxdata域中。

[0252] 19.根据附记15所述的音频处理单元,其中,所述编码音频比特流是AC-3比特流、E-AC-3比特流或者杜比E比特流中的一种。

[0253] 20.根据附记15所述的音频处理单元,其中,所述音频处理单元是数字信号处理器。

[0254] 21.一种音频处理单元,包括:

[0255] 处理子系统,所述处理子系统被耦接和配置为接收编码音频比特流以及从所述编码音频比特流中提取响度处理状态元数据和音频数据,

[0256] 其中,所述编码音频比特流包括帧的序列,所述帧中的每个包括至少一个元数据

分段和至少一个音频分段,每个所述音频数据分段表示至少部分所述音频数据,至少一个所述元数据分段包括响度处理状态元数据有效载荷,所述响度处理状态元数据有效载荷包括至少部分所述响度处理状态元数据,所述响度处理状态元数据有效载荷包括:

[0257] 首部;以及

[0258] 在所述首部之后的:至少一个会话指示值,所述至少一个会话指示值表示相应的音频数据是指示会话还是不指示会话;至少一个响度调节相符值,所述至少一个响度调节相符值指示所述相应的音频数据是否与所指示的响度调节的集合相符;至少一个响度处理值,所述至少一个响度处理值表示已经对所述相应的音频数据执行的至少一种类型的响度处理;以及至少一个响度值,所述至少一个响度值表示所述相应的音频数据的至少一个响度特性。

[0259] 22.根据附记21所述的音频处理单元,其中,每个包括一个所述响度处理状态元数据有效载荷的所述元数据分段包括:

[0260] 核心首部;

[0261] 在所述核心首部之后的至少一个保护值,所述至少一个保护值所述响度处理状态元数据或者所述相应的音频数据中的至少一个的解密、认证或验证中的至少一个;以及

[0262] 在所述核心首部之后的响度处理状态元数据标识和有效载荷尺寸值,其中,所述响度处理状态元数据有效载荷在所述响度处理状态元数据有效载荷标识和有效载荷尺寸值之后。

[0263] 23.根据附记22所述的音频处理单元,其中,所述核心首部包括标识所述元数据分段的开始的同步字以及跟随所述同步字的至少一个标识值,并且所述响度处理状态元数据有效载荷的所述首部包括至少一个标识值。

[0264] 24.根据附记21所述的音频处理单元,其中,所述编码比特流是AC-3比特流或者E-AC-3比特流,所述编码音频比特流的所述帧中的每个包括比特流信息分段和辅助数据auxdata域,所述比特流信息分段包括附加比特流信息addbsi域,并且每个包括响度处理状态元数据有效载荷在内的所述元数据分段被包括在一个所述addbsi域或者一个所述auxdata域中。

[0265] 25.根据附记21所述的音频处理单元,其中,所述编码音频比特流是AC-3比特流、E-AC-3比特流或者杜比E比特流中的一种。

[0266] 26.根据附记21所述的音频处理单元,还包括:

[0267] 状态验证级,所述状态验证级耦接至所述处理子系统并且被配置成认证和验证所述响度处理状态元数据。

[0268] 27.根据附记21所述的音频处理单元,还包括:

[0269] 响度处理级,所述响度处理级耦接至所述处理子系统并且被配置成使用从所述编码音频比特流中提取的所述响度处理状态元数据来对从所述编码音频比特流中提取的所述音频数据执行自适应响度处理。

[0270] 28.根据附记21所述的音频处理单元,还包括:

[0271] 后处理器,所述后处理器耦接至所述音频处理单元并且被配置成使用从所述编码音频比特流中提取的所述响度处理状态元数据来对从所述编码音频比特流中提取的所述音频数据执行自适应响度处理。

[0272] 29. 根据附记21所述的音频处理单元,其中,所述音频处理单元是数字信号处理器。

[0273] 30. 一种用于生成包括音频数据分段和元数据分段的编码音频比特流的方法,所述方法包括下列步骤:

[0274] 提供音频数据;

[0275] 生成表示所述音频数据的响度处理状态元数据;以及

[0276] 复用所述编码音频数据与所述响度处理状态元数据,以生成所述编码音频比特流,使得所述编码音频比特流包括帧的序列,所述帧中的每个包括至少一个元数据分段和至少一个音频分段,每个所述音频数据分段表示至少部分所述音频数据,至少一个所述元数据分段包括响度处理状态元数据有效载荷,所述响度处理状态元数据有效载荷包括至少部分所述响度处理状态元数据,所述响度处理状态元数据有效载荷包括:

[0277] 首部;以及

[0278] 在所述首部之后的:至少一个会话指示值,所述至少一个会话指示值表示相应的音频数据是指示会话还是不指示会话;至少一个响度调节相符值,所述至少一个响度调节相符值指示所述相应的音频数据是否与所指示的响度调节的集合相符;至少一个响度处理值,所述至少一个响度处理值表示已经对所述相应的音频数据执行的至少一种类型的响度处理;以及至少一个响度值,所述至少一个响度值表示所述相应的音频数据的至少一个响度特性。

[0279] 31. 根据附记30所述的方法,其中,每个包括一个所述响度处理状态元数据有效载荷的所述元数据分段包括:

[0280] 核心首部;

[0281] 在所述核心首部之后的至少一个保护值,所述至少一个保护值用于所述响度处理状态元数据或者所述相应的音频数据中的至少一个的解密、认证或验证中的至少一个;以及

[0282] 在所述核心首部之后的响度处理状态元数据标识和有效载荷尺寸值,其中,所述响度处理状态元数据有效载荷在所述响度处理状态元数据有效载荷标识和有效载荷尺寸值之后。

[0283] 32. 根据附记31所述的方法,其中,所述核心首部包括标识所述元数据分段的开始的同步字以及跟随所述同步字的至少一个标识值,并且所述响度处理状态元数据有效载荷的所述首部包括至少一个标识值。

[0284] 33. 根据附记30所述的方法,其中,所述编码音频比特流是AC-3比特流或者E-AC-3比特流,所述编码音频比特流的所述帧中的每个包括比特流信息分段和辅助数据auxdata域,所述比特流信息分段包括附加比特流信息adbsi域,并且每个包括响度处理状态元数据有效载荷在内的所述元数据分段被包括在一个所述adbsi域或者一个所述auxdata域中。

[0285] 34. 根据附记30所述的方法,其中,所述编码音频比特流是AC-3比特流、E-AC-3比特流或者杜比E比特流中的一种。

[0286] 35. 一种用于对包括音频数据分段和元数据分段的编码音频比特流进行解码的方法,所述方法包括步骤:

[0287] 接收编码音频比特流;以及

[0288] 从所述编码音频比特流中提取响度处理状态元数据和音频数据,

[0289] 其中,所述编码音频比特流包括帧的序列,所述帧中的每个包括至少一个元数据分段和至少一个音频分段,每个所述音频数据分段表示至少部分所述音频数据,至少一个所述元数据分段包括响度处理状态元数据有效载荷,所述响度处理状态元数据有效载荷包括至少部分所述响度处理状态元数据,所述响度处理状态元数据有效载荷包括:

[0290] 首部;以及

[0291] 在所述首部之后的:至少一个会话指示值,所述至少一个会话指示值表示相应的音频数据是指示会话还是不指示会话;至少一个响度调节相符值,所述至少一个响度调节相符值指示所述相应的音频数据是否与所指示的响度调节的集合相符;至少一个响度处理值,所述至少一个响度处理值表示已经对所述相应的音频数据执行的至少一种类型的响度处理;以及至少一个响度值,所述至少一个响度值表示所述相应的音频数据的至少一个响度特性。

[0292] 36.根据附记35所述的方法,其中,每个包括一个所述响度处理状态元数据有效载荷的所述元数据分段包括:

[0293] 核心首部;

[0294] 在所述核心首部之后的至少一个保护值,所述至少一个保护值用于所述响度处理状态元数据或者所述相应的音频数据中的至少一个的解密、认证或验证中的至少一个;以及

[0295] 在所述核心首部之后的响度处理状态元数据标识和有效载荷尺寸值,其中,所述响度处理状态元数据有效载荷在所述响度处理状态元数据有效载荷标识和有效载荷尺寸值之后。

[0296] 37.根据附记35所述的方法,其中,所述核心首部包括标识所述元数据分段的开始的同步字以及跟随所述同步字的至少一个标识值,并且所述响度处理状态元数据有效载荷的所述首部包括至少一个标识值。

[0297] 38.根据附记36所述的方法,其中,所述编码比特流是AC-3比特流或者E-AC-3比特流,所述编码音频比特流的所述帧中的每个包括比特流信息分段和辅助数据auxdata域,所述比特流信息分段包括附加比特流信息addbsi域,并且每个包括响度处理状态元数据有效载荷在内的所述元数据分段被包括在一个所述addbsi域或者一个所述auxdata域中。

[0298] 39.根据附记36所述的方法,其中,所述编码音频比特流是AC-3比特流、E-AC-3比特流或者杜比E比特流中的一种。

[0299] 40.根据附记36所述的方法,还包括步骤:

[0300] 使用从所述编码音频比特流中提取的所述响度处理状态元数据来对从所述编码音频比特流中提取的所述音频数据执行自适应响度处理。

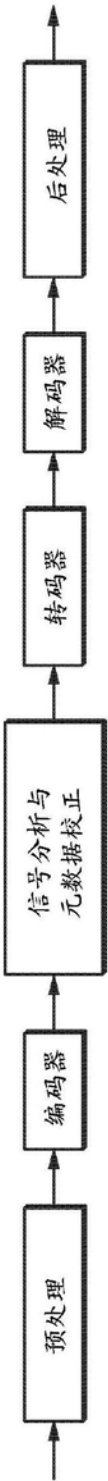


图1

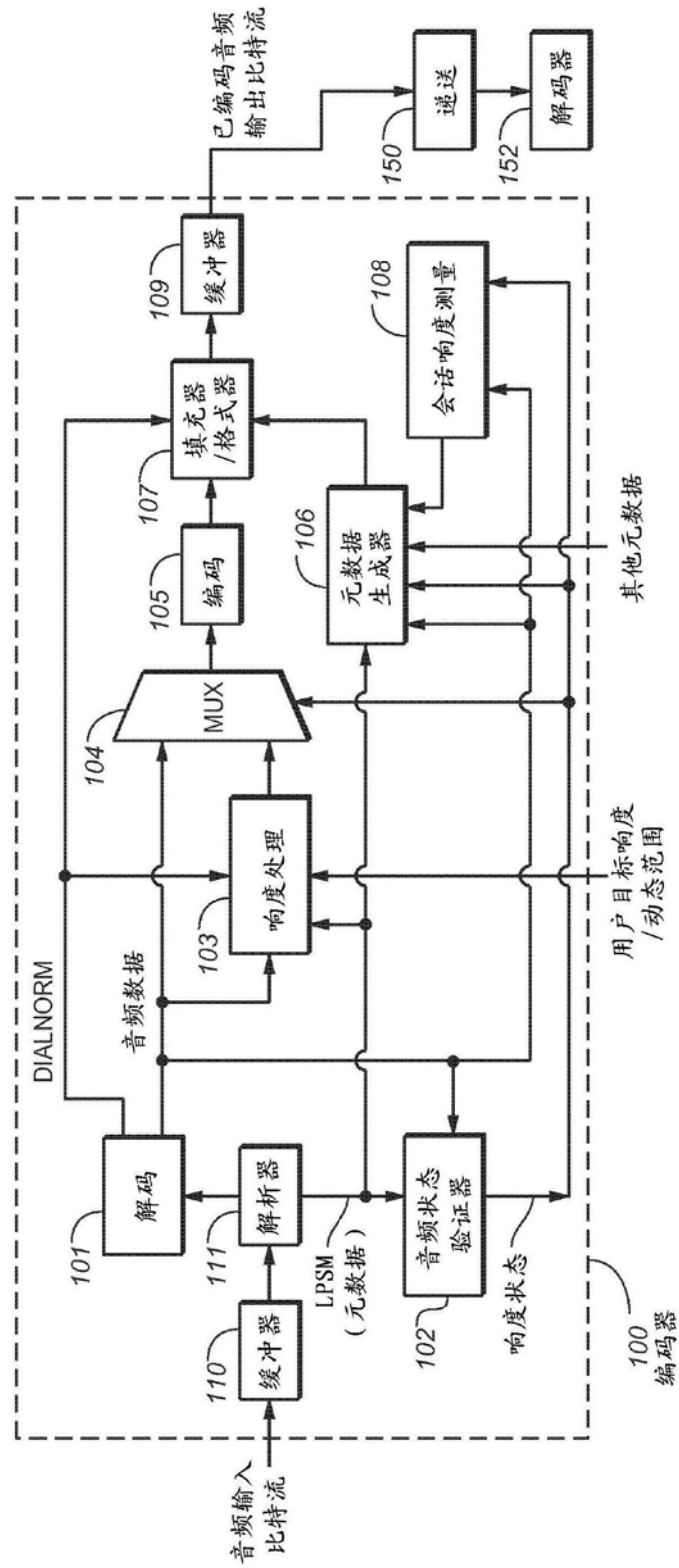


图2

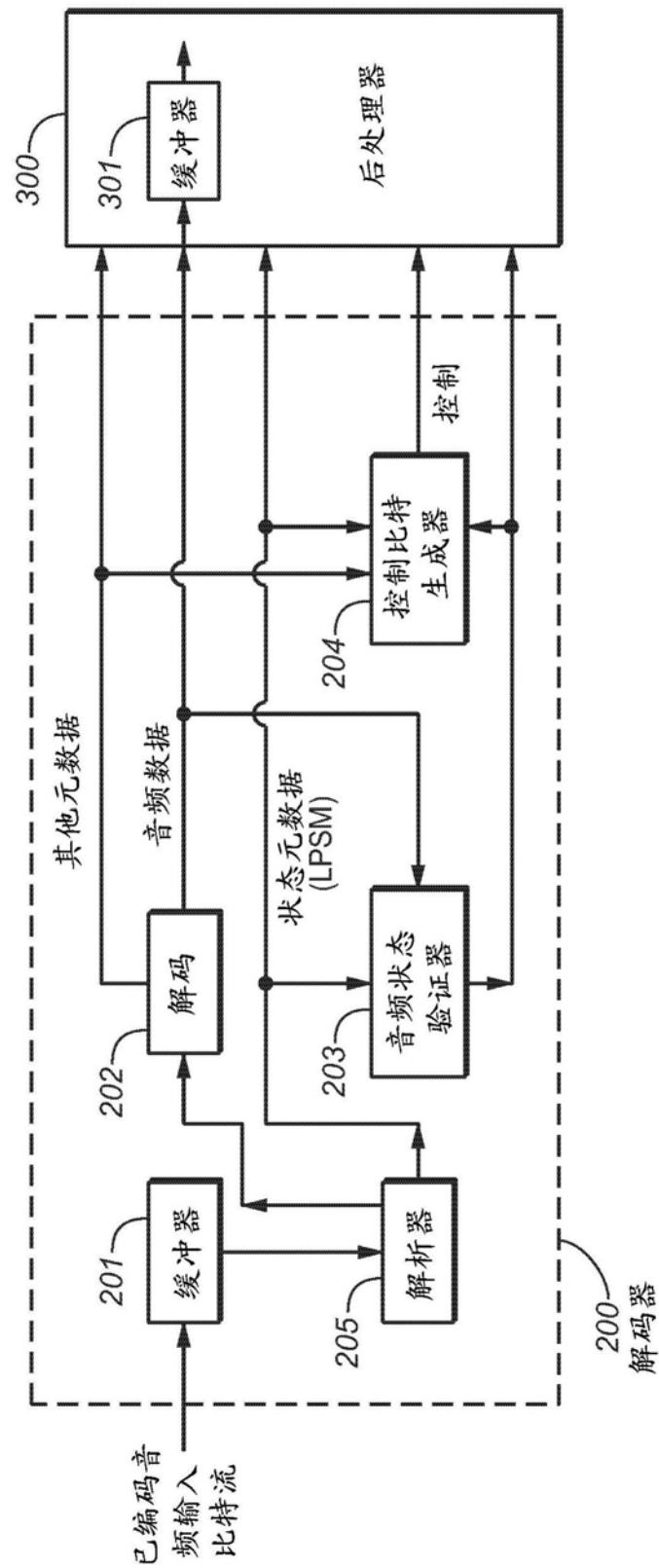


图3

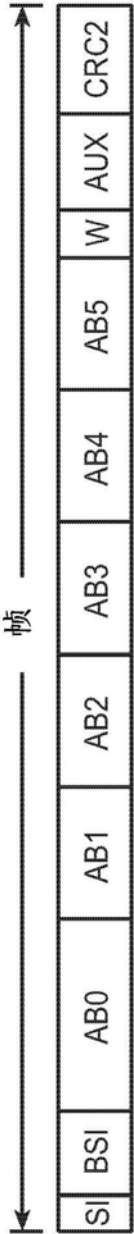


图4

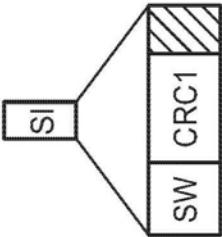


图5

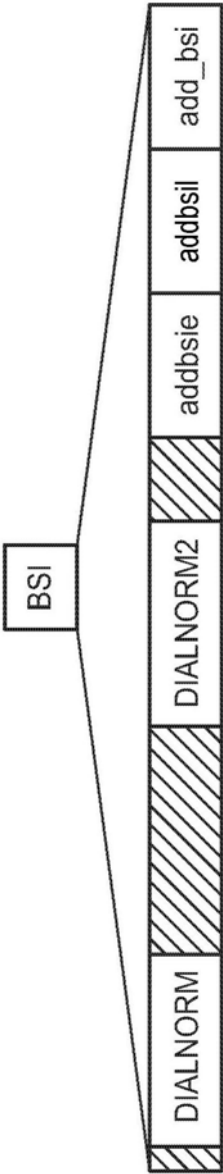


图6

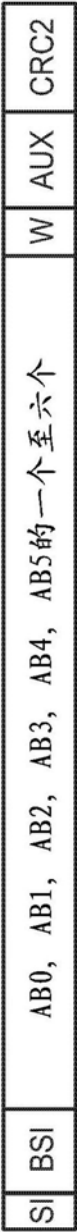


图7