



# (12) 发明专利

(10) 授权公告号 CN 101730832 B

(45) 授权公告日 2014. 05. 28

(21) 申请号 200880022393. 9

G10L 19/008 (2013. 01)

(22) 申请日 2008. 04. 16

(56) 对比文件

(30) 优先权数据

0703349 2007. 05. 10 FR

W0 2002/023529 A, 2002. 03. 21,

CN 100458788 C, 2009. 02. 04,

CN 1497426 A, 2004. 05. 19,

CN 1836457 A, 2006. 09. 20,

FR 2820573 A1, 2002. 08. 09,

(85) PCT国际申请进入国家阶段日

2009. 12. 28

(86) PCT国际申请的申请数据

PCT/FR2008/050671 2008. 04. 16

审查员 杨世兴

(87) PCT国际申请的公布数据

W02008/145893 FR 2008. 12. 04

(73) 专利权人 法国电信

地址 法国巴黎

(72) 发明人 艾迪·莫辛 拉蒂夫·杰隆·图伊米

(74) 专利代理机构 上海天协和诚知识产权代理

事务所 31216

代理人 张恒康

(51) Int. Cl.

G10L 19/24 (2013. 01)

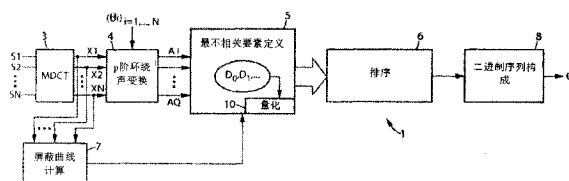
权利要求书2页 说明书25页 附图8页

(54) 发明名称

音频编解码方法及相编解码器和计算机程序

(57) 摘要

本发明涉及一种用于对有待编码的要素 ( $A_1, \dots, A_Q$ ) 进行排序的方法, 这些要素产生于一个包括  $N$  个信号 ( $S_{i=1 \text{ 到 } N}$ ) 的声音场景, 其中  $N > 1$ , 一个有待编码的要素包括与各个频带相关的频谱成份, 其特征在于, 该方法包括以下步骤: 计算至少一些频谱成份的各自影响, 这些成份可以通过产生于  $N$  个信号中的至少一些的频谱参数的函数进行计算, 基于频带上的遮噪比确定, 根据一个对于所述频谱成份编码的函数; 分配至少一个频谱成份的优先级次序, 通过一个函数比较计算出的所述频谱成份的影响与其它成份的影响。



1. 一种用于对有待编码的要素  $A_1, \dots, A_Q$  的频谱成份进行排序的方法, 这些成份产生于一个包括  $N$  个信号  $S_{i=1 \text{ 到 } N}$  的声音场景, 其中  $N > 1$ , 一个有待编码的要素包括与各个不同频带相关的频谱成份, 其特征在于, 该方法包括以下步骤:

- 计算至少一些频谱成份的各自影响, 可以根据产生于  $N$  个信号中的至少一些的频谱参数进行计算, 根据所述频谱成份的编码基于频带上的遮噪比确定;

- 根据针对所述频谱成份所计算出的影响, 与其他频谱成份所计算出的影响相比较, 给至少一个频谱成份分配优先级次序,

通过判断下述两者之间的差异, 计算有待编码的要素中的频谱成份的影响:

- 根据第一速率, 由所述频谱成份编码确定的第一遮噪比; 和

- 根据低于第一速率的第二速率, 由从中删除所述有待编码的要素中的频谱成份的频谱成份的编码确定的第二遮噪比。

2. 根据权利要求 1 所述的方法, 其特征在于, 计算频谱成份的影响通过以下步骤进行:

- a- 对根据第一速率编码的有待编码要素的频谱成份的第一集合进行编码;

- b- 确定每个频带的第一遮噪比;

- c- 确定比所述第一速率低的第二速率;

- d- 删除有待编码的要素中的所述频谱成份, 并对根据第二速率的有待编码的要素的剩余频谱成份进行编码;

- e- 确定每个频带的第二遮噪比;

- f- 根据对于每个频带的第一和第二速率的第一和第二遮噪比之间已确定的差别, 计算遮噪比的变化;

- g- 针对用于排序的有待编码的要素的频谱成份集合中的每一个频谱成份迭代步骤 d 到 f, 并确定最小遮噪比的变化; 分配给符合最小变化的频谱成份的优先级次序作为最小优先级次序。

3. 根据权利要求 2 所述的方法, 其特征在于, 针对用于排序的有待编码的要素的频谱成份集合再重复步骤 a 到 g, 所述集合通过删除已经分配优先级顺序的频谱成份而受到限制。

4. 根据权利要求 2 所述的方法, 其特征在于, 针对用于排序的有待编码的要素的频谱成份集合再重复步骤 a 到 g, 其中已经分配优先级顺序的频谱成份在使用重叠量化器期间被分配了更低的量化值。

5. 根据前述权利要求中任一项所述的方法, 其中所述有待编码的要素包括对于  $N$  个信号计算的频谱参数。

6. 根据权利要求 1 所述的方法, 其中所述有待编码的要素包括通过对于  $N$  个信号计算的频谱参数进行空间变换所得到的要素  $A_1, \dots, A_Q$ 。

7. 根据权利要求 6 所述的方法, 其特征在于, 所述空间变换是环绕声变换。

8. 根据权利要求 6 或 7 所述的方法, 其特征在于, 根据因为编码以及与有待编码要素相关的误差, 空间变换矩阵以及依据所述空间变换矩阵的变换, 确定所述遮噪比。

9. 根据权利要求 6 所述的方法, 一些频谱成份是环绕声成份的频谱参数, 所述方法还包括以下步骤:

- a. 基于角度向量计算所述频谱成份中的至少一些的影响, 所述角度向量根据与葛容准

则相关的能量和速度向量定义,并根据基于量化的环绕声成份的逆环绕声变换来计算;

b. 根据针对至少一个频谱成份所计算的影响与其他频谱成分计算所得的影响相比较,为所述至少一个频谱成份分配优先级顺序。

10. 一种排序模块 (6),用于对有待编码的要素  $A_1, \dots, A_Q$  的频谱成份进行排序,这些成份产生于一个包括  $N$  个信号  $S_{i=1 \text{ 到 } N}$  的声音场景,其中  $N > 1$ ,一个有待编码的要素包括与各个不同频带相关的频谱成份,其特征在于,该模块包括以下装置:

- 计算装置,其计算至少一些频谱成份的各自影响,可以根据产生于  $N$  个信号中的至少一些的频谱参数进行计算,根据所述频谱成份的编码基于频带上的遮噪比确定;

- 分配装置,其根据针对所述频谱成份所计算出的影响,与针对其他频谱成份计算出的影响相比较,给至少一个频谱成份分配优先级次序,

通过判断下述两者之间的差异,计算有待编码的要素中的频谱成份的影响:

- 根据第一速率,由所述频谱成份编码确定的第一遮噪比;和

- 根据低于第一速率的第二速率,由从中删除所述有待编码的要素中的频谱成份的频谱成份的编码确定的第二遮噪比。

11. 一种用于对在输出码流中包括  $N$  个不同信号的 3D 声音场景进行编码的编码器 (1),其中  $N > 1$ ,该编码器包括:

- 变换模块 (3,4),用于根据  $N$  个信号确定与各个频带相关的频谱成份;

- 根据权利要求 10 所述的排序模块 (6),用于对与各个频带相关的频谱成份中的至少一些进行排序;

- 用于构造二进制序列的模块 (8),根据所述排序模块执行的排序,用于构造二进制序列,所述序列包括指示与各个频带相关的频谱成份的数据。

## 音频编解码方法及相编解码器和计算机程序

[0001] 本发明涉及音频信号的编码装置,特别适用于数字和压缩音频信号的存储或传输场合下。

[0002] 更准确地,本发明涉及音频分层编码系统,具有提供可变速率的容量,通过分配关于一个要在分层排列序列中编码的声音信号的信息,以使该信息可以被用于关于声音质量的重要性的次序。用于确定次序所考虑的标准是一个最优化(或者最小劣化)编码声音信号质量的标准。分层编码特别适合于在异构网络间或者那些具有随时间可变速率的网络间的传输,或者也用于到具有不同或可变特性的终端的传输。

[0003] 更具体地,本发明涉及 3D 声音场景的分层编码。3D 声音场景包括多个与单通道声音信号相关的声音信道,其也被认为是空间化的声音。

[0004] 一个已编码的声音场景将被一个声音再现系统再现,该声音再现系统可以包括一个简单的听筒,电脑的两个喇叭或者一个具有五个喇叭的家庭影院 5.1 型系统(其中一个喇叭在屏幕的水平线上,在理论听众的前方一个喇叭在左侧而另一个在右侧;在理论听众的后方一个喇叭在左侧而另一个在右侧),等等。

[0005] 例如,考虑到一个原始声音场景包括三个不同的声源,位于空间的不同位置上。描述该声音场景的信号被编码。产生于该编码的数据被传送到解码器,然后被解码。解码数据被用以产生五个信号,用于声音再现系统中的五个喇叭。五个喇叭中的每个播送其中的一个信号,被喇叭播送的信号集合合成了 3D 声音场景并且因此而在空间中定位了三个虚拟的声源。

[0006] 对声音场景的编码存在不同的技术。

[0007] 例如,所使用的一种技术包括确定描述声音场景的要素,接着进行每一个单声道信号的压缩。产生于这些压缩以及要素描述的数据接着被提供给解码器。

[0008] 依据上述第一种技术的数率适应性(亦称可缩放性)可以通过改变在压缩操作中的速率来获取,但是它是通过单独地考虑针对每一个的信号质量的最优化标准来获得的。

[0009] 另一种编码技术包括从各种通道上的所有单声道音频信号提取并编码空间参数,所述编码技术用于“MPEG 音频环绕”编码器(参见“Text of ISO/IEC FDIS 23003-1, MPEG Surround”, ISO/IEC JTC1/SC29/WG11 N8324, July 2006, Klagenfurt, Austria, 即“ISO/IEC FDIS 23003-1 的文本, MPEG 环绕”, ISO/IEC JTC1/SC29/WG11 N8324, 2006 年 7 月, 克拉根福, 奥地利)中。接着将这些信号混合以获得单声道或立体声信号,所述信号接着由常规单声道或立体声编码器(例如, MPEG-4AAC、HE-AAC 型, 等等)压缩。在解码器层级处,由空间参数和经解码单声道或立体声信号来进行 3D 声音场景的合成。

[0010] 使用此另一技术,可因此通过使用分级单声道或立体声编码器来实现速率适应性,但它是根据用于最佳化单声道或立体声信号的质量的准则来进行的。

[0011] 另外, PSMAC (Progressive Syntax-Rich Multichannel Audio Codec, 渐进富语法多通道音频编解码器)方法使得能够通过使用 KLT 变换(Karhunen Loeve Transform, 卡洛变换)对来自各种通道的信号进行编码, KLT 变换主要可用于信号的解相关,且对应于表示信号统计数据的空间中的主要分量分解。其使得有可能区别能量最多的分量与能量最少的

分量。

[0012] 速率适应性是基于能量较少的分量的消除。尽管如此,这些成份有时对于声音总体质量来说具有重要意义。

[0013] 因此,尽管已知技术在速率适应性方面得到良好结果,但没有一个能基于最优化声音总体质量的标准,提供完全满意的速率适应性的方法,且在还原解码 3D 声音场景过程中,针对定义的压缩数据最优化所感知的声音总体质量。

[0014] 而且,已知的 3D 声音场景编码技术也没有在还原 3D 声音场景的过程中允许基于最优化空间分辨率标准的速率适应性。此适应性将使得有可能保证速率的每一次减少都会尽可能少地损害空间中的声音源的定位准确度及再现区域的尺寸,该尺寸必须尽可能地宽以环绕听众的头部。

[0015] 此外,对于任何用于恢复 3D 声音场景的声音呈现系统,已知 3D 声音场景编码技术均没有使得能直接保证最佳质量的速率适应性成为可能。目前的编码算法经定义以最佳化相对于声音再现系统的特定配置的质量。事实上,在上述“MPEG 音频环绕”编码器的情况下,举例来说,用一耳机或两个扬声器直接收听、或单声道收听是可能的。如果希望用 5.1 或 7.1 型的声音再现系统来处理经压缩位流,那么必须在解码器层级处实施额外处理,例如借助于 OTT (One-To-Two, 一到二) 盒,以便由两个经解码信号产生五个或七个信号。这些盒使得能够在 5.1 或 7.1 型的声音再现系统的情况下获得所需要数目的信号,但不可能再现真实空间方面。此外,这些盒不保证对除了 5.1 或 7.1 型的声音再现系统外的声音再现系统的适应性。

[0016] 本发明旨在改善所述状况。

[0017] 为此,根据第一方面,本发明提出了一种用于对有待编码的要素的频谱成份进行排序的方法,这些成份产生于一个包括 N 个信号的声音场景,其中  $N > 1$ ,一个有待编码的要素包括与各个不同频带相关的频谱成份。

[0018] 该方法包括以下步骤:

[0019] - 计算至少一些频谱成份的各自影响,可以根据产生于 N 个信号中的至少一些的频谱参数进行计算,根据所述频谱成份的编码基于频带上的遮噪比确定;

[0020] - 根据针对所述频谱成份所计算出的影响,与其他计算出的影响相比较,给至少一个频谱成份分配优先级次序。

[0021] 因此,根据本发明的一种方法允许根据有待编码的要素的成份的整体声音质量来安排重要性次序。

[0022] 在整个场景的有待编码的不同要素的不同频谱成份进行相互比较之后,构造一个二进制序列,关于它们对感知的整体声音质量的贡献进行相互比较。因此,为了将多个信号一起进行压缩,信号间的交互作用也因此得到考虑。

[0023] 码流可以这样被排序,以使每一个速率的减少尽量小地降低 3D 声音场景的感知整体声音质量,因为根据它们对于整体声音质量水平的贡献,重要性最低的要素被检测,以便可以不把它们插入(当分配给该传输的速率不足以发送所有有待编码的要素的成份时)或者把它们置于二进制序列的末端(使得由于切断序列而产生的缺陷最小化成为可能)。

[0024] 在一个实施例中,计算一个频谱成份的影响通过以下步骤进行:

[0025] a- 对根据第一速率编码的有待编码要素的频谱成份的第一集合进行编码;

[0026] b- 确定每个频带的第一遮噪比；

[0027] c- 确定比所述第一速率低的第二速率；

[0028] d- 删除有待编码的要素中的所述通常频谱成份，并对根据第二速率的有待编码的要素的剩余频谱成份进行编码；

[0029] e- 确定每个频带的第二遮噪比；

[0030] f- 根据对于每个频带的第一和第二速率的第一和第二遮噪比之间已确定的差别，计算遮噪比的变化；

[0031] g- 针对用于排序的有待编码的要素的频谱成份集合中的每一个频谱成份迭代步骤 d 到 f，并确定最小遮噪比的变化；分配给符合最小变化的频谱成份的优先级次序作为最小优先级次序。

[0032] 这样一个过程使得确定有待编码的要素的至少一个成份成为可能，根据对整体声音质量的贡献，与有待编码用于排序的要素的其它成份的集合相比，该成份的重要性最低。

[0033] 在一个实施例中，针对用于排序的有待编码的要素的频谱成份集合再重复步骤 a 到 g，所述集合通过删除已经分配优先级顺序的频谱成份而受到限制。

[0034] 在另一个实施例中，针对用于排序的有待编码的要素的频谱成份集合再重复步骤 a 到 g，其中已经分配优先级顺序的频谱成份在使用重叠量化器期间被分配了更低的量化值。

[0035] 在一个实施例中，其中所述有待编码的要素包括对于 N 个信号计算的频谱参数。这些要素是例如被直接编码的信号的频谱成份。

[0036] 在另一个实施例中，其中所述有待编码的要素包括通过对于 N 个信号计算的频谱参数进行空间变换所得到的要素，例如环绕声变换。一方面，这一安排使得减少要被发送的数据的数量成为可能，一般地，N 个信号可以通过数量被减少的环绕声成份而得到很好地描述（例如，数量等于 3 到 5），小于 N。该安排还允许适配至任何类型的声音再现系统，因为它在解码侧采用大小为  $Q' \times (2p' + 1)$  的逆环绕声变换是足够的，（其中  $Q'$  等于在解码器输出端所使用的声音再现系统的喇叭数量， $2p' + 1$  等于所接收到的逆环绕声成份的数量），以用于确定提供给声音再现系统的信号，同时保证整体信号质量。

[0037] 在一个实施例中，代替空间变换，使用其它线性变换例如 KLT 变换等。

[0038] 在一个实施例中，根据因为编码以及与有待编码要素相关的误差，空间变换矩阵以及依据所述空间变换矩阵的变换，确定所述遮噪比。

[0039] 在一个实施例中，有待编码的要素是环绕声成份，频谱成份中的一些是环绕声成份的频谱参数。该方法包括以下步骤：

[0040] a. 基于角度向量 (angle vector) 计算所述频谱成份中的至少一些的影响，所述角度向量根据与葛容准则 (Gerzon criteria) 相关的能量和速度向量 (energy and velocity vectors) 定义，并根据基于所述量化的环绕声成份的逆环绕声变换来计算；

[0041] b. 根据针对至少一个频谱成份所计算的影响与其它计算所得的影响，为所述至少一个频谱成份分配优先级顺序。

[0042] 从而，根据本发明的方法使得，根据它们对于空间精度的贡献的重要性，对有待排序的集合的环绕声成份的至少一些频谱参数进行排序成为可能。

[0043] 空间分辨率或空间准确度度量了声音源在空间中的位置的精细程度。增加的空间

分辨率使声音对象在房间中的更精细定位成为可能,且使收听者头部周围更宽阔的恢复区域成为可能。

[0044] 考虑信号之间的相互作用及其与空间准确度有关的后果,以将它们一起进行压缩。

[0045] 码流可以这样被排序,以使每一个速率的减少尽量小地降低 3D 声音场景的感知整体声音质量,因为根据它们对于整体声音质量水平的贡献,重要性最低的要素被检测,以便把它们置于二进制序列的末端(使得由于切断序列而产生的缺陷最小化成为可能)。

[0046] 根据该方法的一个实施例,使用与葛容准则的速度向量 $\vec{V}$ 和能量向量 $\vec{E}$ 有关的角度 $\xi_v$ 和 $\xi_e$ ,正如下文指出的,以便识别出有待编码的要素中哪些与所考虑的对于 3D 声音场景的空间精度的贡献最低程度相关。因此,与惯常手段相反,速度向量 $\vec{V}$ 和能量向量 $\vec{E}$ 并不用于优化所述声音再现系统。

[0047] 在一个实施例中,计算频谱参数的影响是按照以下步骤实施的:

[0048] a- 对根据第一速率编码的有待编码的环绕声成份的频谱参数的第一集合进行编码;

[0049] b- 确定每个频带的第一角向量;

[0050] c- 确定比所述第一速率低的第二速率;

[0051] d- 删除有待编码的成份中的所述通常频谱参数,并对根据第二速率的有待编码的成份的剩余频谱参数进行编码;

[0052] e- 确定每个频带的第二角向量;

[0053] f- 根据对于每个频带的第一和第二速率的第一和第二角向量之间已确定的差别,计算角向量的变化;

[0054] g- 针对用于排序的有待编码的成份的频谱参数集合中的每一个频谱参数迭代步骤 d 到 f,并确定最小角向量的变化;分配给符合最小变化的频谱参数的优先级次序作为最小优先级次序。

[0055] 这一安排使得,通过有限次数的计算,确定所要确定的成份的频谱参数成为可能,其对于空间精度的贡献最小。

[0056] 在一个实施例中,针对用于排序的有待编码的成份的频谱参数集合再重复步骤 a 到 g,所述集合通过删除已经分配优先级顺序的频谱参数而受到限制。

[0057] 在另一个实施例中,针对用于排序的有待编码的成份的频谱参数集合再重复步骤 a 到 g,其中已经分配优先级顺序的频谱参数在使用重叠量化器期间被分配了更低的量化值。

[0058] 这样一种迭代方法使得在还没有分配优先级顺序的环绕声成份的频谱参数中连续识别出那些对于空间精度的贡献最小的频谱参数成为可能。

[0059] 在一个实施例中,能量向量的第一坐标符合函数公式 $\frac{\sum_{1 \leq i \leq Q} T_i^2 \cos \xi_i}{\sum_{1 \leq i \leq Q} T_i^2}$ ,能量

向量的第二坐标符合函数公式 $\frac{\sum_{1 \leq i \leq Q} T_i^2 \sin \xi_i}{\sum_{1 \leq i \leq Q} T_i^2}$ ,速度向量的第一坐标符合函数公式

$\frac{\sum_{1 \leq i \leq Q} T_i \cos \xi_i}{\sum_{1 \leq i \leq Q} T_i}$ , 速度向量的第二坐标符合函数公式  $\frac{\sum_{1 \leq i \leq Q} T_i \sin \xi_i}{\sum_{1 \leq i \leq Q} T_i}$ , 其中  $T_i, i = 1$  到  $Q$ , 代表

基于根据所考虑的速率的所述量化频谱参数的逆环绕声变换所确定的信号,  $\xi_i, i = 1$  到  $Q$  是确定的角度。

[0060] 在一个实施例中, 角度向量的第一坐标表示该角度是速度向量的第二坐标的符号以及速度向量的第一坐标的反余弦的函数, 并且, 角度向量的第二坐标表示该角度是能量向量的第二坐标的符号以及能量向量的第一坐标的反余弦的函数。

[0061] 根据本发明的第二方面, 本发明提出了一种用于实施本发明第一方面的方法的排序模块。

[0062] 根据本发明的第三方面, 本发明提出了一种用于对在输出码流中包括  $N$  个不同信号的 3D 声音场景进行编码的音频编码器, 其中  $N > 1$ , 该编码器包括:

[0063] - 变换模块, 用于根据  $N$  个信号确定与各个频带相关的频谱成份;

[0064] - 根据本发明第二方面所述的排序模块, 用于对与各个频带相关的频谱成份中的至少一些进行排序;

[0065] - 用于构造二进制序列的模块, 根据所述排序模块执行的排序, 用于构造二进制序列, 所述序列包括指示与各个频带相关的频谱成份的数据。

[0066] 根据本发明的第四方面, 本发明提出了一种计算机程序, 用于安装在排序模块中, 在通过所述模块的处理装置执行该程序过程中, 所述程序包括用于执行根据本发明第一方面所述方法的指令。

[0067] 根据本发明的第五方面, 本发明提出了一种用于解码码流的方法, 该码流根据本发明的第一方面所述的方法进行编码, 以确定用于通过  $Q'$  个喇叭还原 3D 声音场景的  $Q'$  个声音信号, 该方法包括:

[0068] - 接收二进制序列;

[0069] - 抽取编码数据, 并且, 根据所抽取的数据, 确定针对  $Q'$  个通道的与各个频带相关的参数集合;

[0070] - 根据每个参数集合, 确定至少一个信号帧。

[0071] 根据本发明的第六方面, 本发明提出了一种解码器, 用于解码根据本发明的第一方面的方法进行编码的码流, 以确定用于通过  $Q'$  个喇叭还原 3D 声音场景的  $Q'$  个声音信号, 该解码器包括用于执行根据本发明的第四方面所述的方法的步骤的装置。

[0072] 根据本发明的第七方面, 本发明提出了一种安装在解码器中的计算机程序, 用于解码根据本发明的第一方面所述的方法进行编码的码流, 以确定用于通过  $Q'$  个喇叭还原 3D 声音场景的  $Q'$  个声音信号, 在通过所述解码器的处理装置执行该程序过程中, 所述程序包括用于执行根据本发明第四方面所述方法的指令。

[0073] 根据本发明的第八方面, 本发明提出了一种二进制序列, 包括与产生于包括  $N$  个信号的声音场景中的待编码的要素的各个频带相关的频谱成份, 其中  $N > 1$ , 其特征在于, 这些频谱成份中的至少一些是根据本发明的第一方面所述的排序方法进行排序的。

[0074] 本发明的其它特点和优点将通过下面的描述变得更加清晰。这是示例性的, 并且必须借助于参照相关附图, 其中:

[0075] - 图 1 表示本发明的一个实施例中的编码器;

- [0076] - 图 2 表示本发明的一个实施例中的解码器；
- [0077] - 图 3 表示空间中的一个平面波的传播；
- [0078] - 图 4 的流程图表示本发明的一个实施例中的第一处理 Proc1 的各个步骤；
- [0079] - 图 5a 表示本发明的一个实施例中构造的二进制序列；
- [0080] - 图 5b 表示本发明的另一个实施例中构造的二进制序列 Seq；
- [0081] - 图 6 的流程图表示本发明的一个实施例中的第二处理 Proc2 的各个步骤；
- [0082] - 图 7 表示一个包括 8 个喇叭  $h_1, h_2, \dots, h_8$  的声音再现系统的构造示例；
- [0083] - 图 8 表示一个处理流程；
- [0084] - 图 9 表示一个第二处理流程；
- [0085] - 图 10 表示一个第三处理流程；
- [0086] - 图 11 的流程图表示本发明的一个实施例中的方法 Proc 的各个步骤。
- [0087] 图 1 表示本发明的一个实施例中的音频编码器。
- [0088] 编码器 1 包括时间 / 频率变换模块 3, 屏蔽曲线计算模块 7, 空间变换模块 4, 用于定义有待编码的最低程度相关要素的模块 5, 其与量化模块 10 相连, 用于对要素进行排序的模块 6, 用于构造二进制序列的模块 8, 用以传输码流  $\Phi$ 。
- [0089] 一个包括 N 个通道的 3D 场景, 在每个通道上分别传输信号  $S_1, \dots, S_N$ 。
- [0090] 图 2 表示本发明的一个实施例中的解码器 100。
- [0091] 解码器 100 包括二进制序列读取模块 104, 逆量化模块 105, 逆环绕声变换模块 101, 以及频率 / 时间变换模块 102。
- [0092] 解码器 100 用于在输入端接收由编码器 1 发送的码流  $\Phi$ , 并且在输出端输出  $Q'$  个信号  $S'_1, S'_2, \dots, S'_{Q'}$ , 用于输入到声音再现系统 103 的  $Q'$  个喇叭  $H_1, H_2, \dots, H_{Q'}$ 。
- [0093] 每个喇叭  $H_i, i = 1$  到  $Q'$  与一个表示来自喇叭的传播的角度  $\beta_i$  相关。
- [0094] 在编码器侧执行的处理
- [0095] 编码器 1 的时间 / 频率变换模块 3 在其输入端接收有待编码的 3D 声音场景的 N 个信号  $S_1, \dots, S_N$ 。
- [0096] 每个信号  $S_i, i = 1$  到 N, 由 3D 场景的声波的声音全向压力 (omnidirectional pressure)  $P_i$  以及传播角度  $\theta_i$  表示。
- [0097] 在每个信号的每个时间帧上, 这些时间帧表示由声音压力  $P_i$  的随时间变化所取的不同值, 时间 / 频率变换模块 3 进行时间 / 频率变换, 在当前情况下, 进行改进的离散余弦变换 (modified discrete cosine transform, MDCT)。
- [0098] 从而对于每个信号  $S_i, i = 1$  到 N, 可以确定其频谱表示为  $X_i$ , 用 M 个 MDCT 变换系数  $X(i, j)$  表征, 其中  $j = 0$  到  $M-1$ 。因此, MDCT 系数表示频带  $F_j$  的信号  $S_i$  的频谱。
- [0099] 信号  $S_i, i = 1$  到 N, 的频谱表示  $X_i$ , 被提供到空间变换模块 4 的输入端, 该模块还在其输入端接收表征输入信号  $S_i$  的声音传播角度  $\theta_i$ 。
- [0100] 信号  $S_i, i = 1$  到 N, 的频谱表示  $X_i$ , 还被提供到屏蔽曲线计算模块 7 的输入端。
- [0101] 屏蔽曲线计算模块 7 用于确定每一个单独的信号  $S_i$  的屏蔽曲线, 通过使用其频谱表示  $X_i$  以及心理声学模型 (psychoacoustic model), 其为每一个频谱表示  $X_i$  的每一个频带  $F_j, j = 0$  到  $M-1$ , 提供了一个屏蔽水平 (masking level)。这些屏蔽曲线的定义要素被

传送到用于定义有待编码的最小程度相关要素的模块 5。

[0102] 空间变换模块 4 用于对所提供的输入信号进行空间变换,也就是,确定来自空间参考系统发射的信号的空间成份,依赖于变换的次序。空间变换的顺序与“搜索”声域的角频率相关。

[0103] 在一个实施例中,空间变换模块 4 执行环绕声变换,给出了 3D 声音场景的紧凑空间表示(compact spatial representation),通过在相关球形或者圆柱形谱函数上制造声域的声音发射。

[0104] 为了获取更多关于环绕声变换的信息,可以参考以下文献:“Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia”(声场的表示方法,应用多媒体环境下的复杂声音场景的传播与重现), Doctoral Thesis of the University of Paris 6, Jérôme DANIEL, 31 July 2001, "A highly scalable spherical microphone array based on an orthonormal decomposition of the sound field"(基于声场的正交分解的高级球形扩音器排列), Jens Meyer-Gary Elko, Vol. II-pp. 1781-1784 in Proc. ICASSP 2002。

[0105] 参考图 3, 以下公式给出了应用于在声场的信号  $S_i$  的无限阶圆柱形谐波的分解公式:

[0106]

$$S_i(r, \varphi) = P_i \cdot [J_0(kr) + \sum_{1 \leq m \leq \infty} 2 \cdot j^m J_m(kr) \cdot (\cos m \cdot \theta_i \cdot \cos m \cdot \varphi + \sin m \cdot \theta_i \cdot \sin m \cdot \varphi)]$$

[0107] 其中 ( $J_m$ ) 表示贝塞尔函数 (the Bessel functions),  $r$  表示从帧中央到位于点  $M$  的听者的位置的距离,  $P_i$  是信号  $S_i$  的声音压力,  $\theta_i$  是与信号  $S_i$  相应的声波的传播角度,  $\varphi$  是听者位置与帧的轴线之间的夹角。

[0108] 如果环绕声变换是  $p$  阶 ( $p$  表示任何一个正整数), 对于 2D 环绕声变换 (在水平面上), 则对于在时域中表示的信号  $S_i$  的环绕声变换包括以下  $2p+1$  个成份:

[0109] ( $P_i, P_i \cdot \cos \theta_i, P_i \cdot \sin \theta_i, P_i \cdot \cos 2 \theta_i, P_i \cdot \sin 2 \theta_i, P_i \cdot \cos 3 \theta_i, P_i \cdot \sin 3 \theta_i, \dots, P_i \cdot \cos p \theta_i, P_i \cdot \sin p \theta_i$ ).

[0110] 在下文中, 已经考虑了 2D 环绕声变换。然而, 本发明可以在 3D 环绕声变换的情况下实施 (在这种情况下, 考虑听众按照球面排列)。

[0111] 在频域中考虑环绕声成份  $A_k, k = 1$  到  $Q = 2p+1$ , 每一个都包括  $M$  个频谱参数  $A(k, j), j = 0$  到  $M-1$ , 分别与频带  $F_j$  相关, 以使得:

[0112] 如果矩阵  $\underline{A}$  包括产生于信号  $S_i, i = 1$  到  $N$  的  $p$  阶的环绕声变换的成份  $A_k, k = 1$  到  $Q, \underline{Amb}(p)$  是环绕声对于空间声场的  $p$  阶的环绕声变换矩阵,  $\underline{X}$  是信号  $S_i, i = 1$  到  $N$  的频率成份的矩阵, 那么:

$$[0113] \quad \underline{A} = \begin{bmatrix} A(1,0) & A(1,1) & . & . & . & A(1,M-1) \\ A(2,0) & & & & & A(2,M-1) \\ . & & & & & . \\ . & & & & & . \\ A(Q,0) & A(Q,1) & . & . & . & A(Q,M-1) \end{bmatrix}$$

[0114]  $\underline{Amb}(p) = [Amb(p)(i, j)]$ , 其中  $i = 1$  到  $Q$ , 并且  $j = 1$  到  $N$ , 其中  $Amb(p)(1, j) = 1$ ,

[0115] 如果  $i$  是偶数,  $Amb(p)(i, j) = \sqrt{2} \cos\left[\left(\frac{i}{2}\right)\theta_j\right]$ ,

[0116] 如果  $i$  是奇数,  $Amb(p)(i, j) = \sqrt{2} \sin\left[\left(\frac{i-1}{2}\right)\theta_j\right]$ , 即:

$$[0117] \quad \underline{Amb}(p) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \sqrt{2} \cos \theta_1 & \sqrt{2} \cos \theta_2 & \dots & \sqrt{2} \cos \theta_N \\ \sqrt{2} \sin \theta_1 & \sqrt{2} \sin \theta_2 & \dots & \sqrt{2} \sin \theta_N \\ \sqrt{2} \cos 2\theta_1 & \sqrt{2} \cos 2\theta_2 & \dots & \sqrt{2} \cos 2\theta_N \\ \sqrt{2} \sin 2\theta_1 & \sqrt{2} \sin 2\theta_2 & \dots & \sqrt{2} \sin 2\theta_N \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{2} \cos p\theta_1 & \sqrt{2} \cos p\theta_2 & \dots & \sqrt{2} \cos p\theta_N \\ \sqrt{2} \sin p\theta_1 & \sqrt{2} \sin p\theta_2 & \dots & \sqrt{2} \sin p\theta_N \end{bmatrix}$$

$$[0118] \quad \text{且 } \underline{X} = \begin{bmatrix} X(1,0) & X(1,1) & \dots & X(1,M-1) \\ X(2,0) & X(2,1) & \dots & X(2,M-1) \\ \vdots & \vdots & \ddots & \vdots \\ X(N,0) & \dots & \dots & X(N,M-1) \end{bmatrix}$$

[0119] 这样我们可以得到  $\underline{A} = \underline{Amb}(p) \times \underline{X}$  等式 (1)。

[0120] 空间变换模块 4 用于确定矩阵  $\underline{A}$ , 根据其输入端收到的数据  $X(i, j)$  及  $\theta_i (i = 1$  到  $N, j = 0$  到  $M-1)$  使用等式 (1)。

[0121] 考虑到一种特殊情况, 环绕声成份  $A_k, k = 1$  到  $Q$ , 也就是矩阵  $\underline{A}$  的参数  $A(k, j), k = 1$  到  $Q$  且  $j = 0$  到  $M-1$ , 是有待被编码器 1 进行编码的二进制序列中的要素。

[0122] 环绕声成份  $A_k, k = 1$  到  $Q$  被传送到模块 5, 模块 5 用于定义量化的最低相关要素以及确定环绕声成份的排序。

[0123] 用于定义量化的最低相关要素的模块 5 通过模块 5 的处理装置, 用于执行第一算法和 / 或第二算法的操作, 用以定义有待编码的最低相关要素并且对有待编码的要素进行彼此排序。

[0124] 该有待编码的要素排序随后在构造将被传输的二进制序列的过程中使用。

[0125] 第一算法包括用于执行如下文参考图 4 所述的过程 Proc1 的指令, 当这些指令在模块 5 的处理装置上被执行时。

[0126] 过程 Proc1

[0127] 过程 Proc1 的原理是: 计算至少一些频谱成份的相应影响, 其可以根据产生于  $N$  个信号中的至少一些频谱参数进行计算, 基于根据对所述频谱成份的编码的在频带上确定的遮噪比 (mask-to-noise ratios)。然后, 根据对至少一个频谱成份计算的影响与其他成

份计算的影响的比值,对所述至少一个频谱成份分配优先级顺序。

[0128] 在一个实施例中,过程 Proc1 具体为:

[0129] 初始化

[0130] • 步骤 1a:

[0131] 在该步骤中,定义第一速率  $D_0 = D_{\max}$  以及对于该速率  $D_0$  在有待编码的要素  $A(k, j)$ ,  $(k, j) \in E_0 = \{(k, j) \text{ 使得 } k = 1 \text{ 到 } Q \text{ 且 } j = 0 \text{ 到 } M-1\}$  之间分配的速率的一部分。在该分配期间分配到有待编码的要素  $A(k, j)$ ,  $(k, j) \in E_0$  的速率 (这些速率的总和  $d_{k, j | k=1 \text{ 到 } Q, j=0 \text{ 到 } M-1}$  等于  $D_0$ ) 被称为  $d_{k, j}$  且对于  $(k, j) \in E_0$ ,  $\delta_0 = \min d_{k, j}$ 。

[0132] 接着,有待编码的要素  $A(k, j)$ ,  $(k, j) \in E_0$ , 根据针对速率  $D_0$  已定义的分配,通过量化模块 10 进行量化。

[0133] • 步骤 1b:

[0134] 接着,计算每个信号  $S_i$  以及每个子带  $F_j$  的屏蔽量化误差 (噪音) 比 (the ratio of the mask to the quantification error) (“遮噪比” “Mask to noise Ratio” 或者 MNR), 其中  $i = 1$  到  $N$  且  $j = 0$  到  $M-1$ , 其等于频带  $F_j$  中的信号  $S_i$  的屏蔽功率除以与频带  $F_j$  中的信号  $S_i$  相关的量化噪音 ( $E(i, j)$ ) 的功率。

[0135] 为了实现这个步骤,有待编码的要素  $A(k, j)$   $(k, j) \in E_0$  的每个频带  $F_j$  中的量化误差  $b(k, j)$  首先由如下步骤确定的:

[0136]  $b(k, j) = A(k, j) - \bar{A}(k, j)$ , 其中  $\bar{A}(k, j)$  是量化结果, 然后对要素  $A(k, j)$  的量化求倒数 (量化通常提供一个量化指数表示在字典中量化的要素值, 量化倒数提供了根据量化指数进行量化的要素的值)。

[0137] 然后, 确定针对每个信号  $S_i$  在每个频带  $F_j$  中的量化误差  $E(i, j)$ , 其中  $i = 1$  到  $N$  且  $j = 0$  到  $M-1$ , 根据速率  $D_0$  下的有待编码的要素的量化值, 通过计算包括要素  $E(i, j)$  的矩阵  $\underline{E}$ :

$$[0138] \quad \underline{E} = \frac{1}{Q^2} (\text{Amb}(p) \cdot \text{Amb}(p)^t)^{-1} \cdot \underline{\text{Amb}(p)}^t \cdot \underline{B}, \text{等式 (2)}$$

[0139] 其中  $Q = 2p+1$ ,  $\text{Amb}(p)$  是  $p$  阶的环绕声变换矩阵, 并且

$$[0140] \quad \underline{E} = \begin{bmatrix} E(1,0) & E(1,1) & \cdot & \cdot & \cdot & E(1,M-1) \\ E(2,0) & E(2,1) & \cdot & \cdot & \cdot & E(2,M-1) \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ E(N,0) & \cdot & \cdot & \cdot & \cdot & E(N,M-1) \end{bmatrix} = [E(i, j)]_{i=1 \text{ 到 } N, j=0 \text{ 到 } M-1} \text{ 并且}$$

[0141]

$$\underline{B} = \begin{bmatrix} b(1,0) & b(1,1) & \cdot & \cdot & \cdot & b(1,M-1) \\ b(2,0) & & & & & b(2,M-1) \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ b(Q,0) & b(Q,1) & \cdot & \cdot & \cdot & b(Q,M-1) \end{bmatrix} = [b(k, j)]_{k=1 \text{ 到 } Q, j=0 \text{ 到 } M-1}.$$

[0142] 然后, 根据因此计算所得的相对于频带  $F_j$  中的信号  $S_i$  的量化噪音  $E(i, j)$ , 以及

屏蔽计算模块 7 提供的频带  $F_j$  中的信号  $S_i$  的屏蔽, 确定针对每个信号  $S_i$  以及每个子带  $F_j$  的屏蔽量化误差比, 其中  $i = 1$  到  $N$  且  $j = 0$  到  $M-1$ 。

[0143]  $MNR(0, D_0)$  指的是矩阵, 矩阵  $MNR(0, D_0)$  的元素  $(i, j)$ ,  $i = 1$  到  $N$  且  $j = 0$  到  $M-1$ , 表示之前执行的量化步骤中的频带  $F_j$  以及信号  $S_i$  的屏蔽量化误差比。

[0144] 在描述过程 Proc1 的第一次迭代之前, 下面将给出等式 (2) 是如何确定的。

[0145] 图 8 表示处理流程 200, 包括  $p$  阶的环绕声变换模块 201 (类似于图 1 中的  $p$  阶的环绕声变换模块 4), 接着是  $p$  阶的逆环绕声变换模块 202。  $p$  阶的环绕声变换模块 201 在其输入端接收信号  $S_1, \dots, S_N$  的频谱表示  $X_1, \dots, X_N$ , 在这些信号上执行  $p$  阶的环绕声变换, 将所获得的环绕声信号  $A_1$  到  $A_Q$  传输至  $p$  阶的逆环绕声变换模块 202, 该模块输出  $N$  个不同的听觉压力信号  $\Pi_i$ ,  $i = 1$  到  $N$ 。

[0146] 接着, 我们有 
$$\begin{pmatrix} \Pi_1 \\ \Pi_2 \\ \vdots \\ \Pi_N \end{pmatrix} = AmbInv(p) \times Amb(p) \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix},$$
 其中  $Amb(p)$  是  $p$  阶环绕声变换

矩阵, 而  $AmbInv(p)$  是  $p$  阶逆环绕声变换矩阵 (也叫作环绕声解码矩阵)。

[0147] 图 9 表示处理流程 210, 包括  $p$  阶环绕声变换模块 201, 接着是量化模块 203, 然后是逆量化模块 204, 以及  $p$  阶逆环绕声变换模块 202。在处理流程 210 的输入端  $p$  阶环绕声变换模块 201 在输入端接收信号  $S_1, \dots, S_N$  的频谱表示  $X_1, \dots, X_N$ , 并输出获得的环绕声信号  $A_1$  到  $A_Q$ , 这些信号被提供到量化模块 203 的输入端。信号  $\bar{A}_1, \dots, \bar{A}_Q$  是通过逆量化模块 204 输出到逆环绕声变换模块 202 的信号, 它们来自对量化模块 203 输出的信号所进行的逆量化。  $Q$  阶逆环绕声变换模块 202 输出  $N$  个不同的听觉压力信号  $\Pi'_i$ ,  $i = 1$  到  $N$ 。

[0148] 图 9 中的处理流程 210 提供了与图 10 中的处理流程 211 相同的输出听觉压力  $\Pi'_i$ , 其中  $p$  阶环绕声变换模块 201 被置于逆量化模块 204 以及  $p$  阶逆环绕声变换模块 202 之间。在处理流程 211 中, 在处理流程 211 的输出端, 量化模块 203 在其输入端接收频谱表示  $X_1, \dots, X_N$ , 对其进行量化, 并将该量化结果输出到逆量化模块 204, 该模块输出  $N$  个信号  $\bar{X}_1, \dots, \bar{X}_N$ 。这些信号  $\bar{X}_1, \dots, \bar{X}_N$  接着被提供到被级联设置的环绕声变换模块以及环绕声逆变换模块 201 和 202。  $p$  阶逆环绕声变换模块 202 输出  $N$  个不同的听觉压力信号  $\Pi'_i$ ,  $i = 1$  到  $N$ 。

[0149] 然后我们可以写成: 
$$\begin{pmatrix} \Pi'_1 \\ \Pi'_2 \\ \vdots \\ \Pi'_N \end{pmatrix} = AmbInv(p) \times Amb(p) \times \begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_N \end{pmatrix}$$

[0150]

$$\begin{pmatrix} \Pi'_1 \\ \Pi'_2 \\ \vdots \\ \Pi'_N \end{pmatrix} - \begin{pmatrix} \Pi_1 \\ \Pi_2 \\ \vdots \\ \Pi_N \end{pmatrix} = AmbInv(p) \times Amb(p) \times \left( \begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_N \end{pmatrix} - \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix} \right) = AmbInv(p) \times Amb(p) \times \underline{E}.$$

$$[0151] \quad \text{设 } \underline{E} = (\text{AmbInv}(p) \times \text{Amb}(p))^{-1} \left( \begin{pmatrix} \Pi'1 \\ \Pi'2 \\ \Pi'N \end{pmatrix} - \begin{pmatrix} \Pi1 \\ \Pi2 \\ \Pi N \end{pmatrix} \right).$$

$$[0152] \quad \text{并且, } \begin{pmatrix} \Pi'1 \\ \Pi'2 \\ \Pi'N \end{pmatrix} - \begin{pmatrix} \Pi1 \\ \Pi2 \\ \Pi N \end{pmatrix} = \text{AmbInv}(p) \times \left( \begin{pmatrix} \bar{A}1 \\ \bar{A}2 \\ \bar{A}Q \end{pmatrix} - \begin{pmatrix} A1 \\ A2 \\ AQ \end{pmatrix} \right) = \text{AmbInv}(p) \times \underline{B}.$$

[0153] 这样我们可以据此推出： $\underline{E} = (\text{AmbInv}(p) \times \text{Amb}(p))^{-1} \text{AmbInv}(p) \times \underline{B}$ 。

[0154] 如果环绕声解码矩阵与具有普通喇叭的系统相应，我们有 **AmbInv(p)**（事实上，N 个量化误差  $\underline{E}$  或  $\underline{B}$  仅仅依赖于所执行的编码而不依赖于解码。根据所使用的解码矩阵，在解码侧将发生的对应于所使用的喇叭的系统的变化，在于在喇叭间分配误差的方式。这取决于所使用的心理声学没有考虑信号间的干扰。因此，如果计算是针对明确定义的解码矩阵进行的，并且量化模块优化了针对该矩阵的误差，那么对于其它解码矩阵，该误差是次优的）。

[0155] 因此从中推出等式 (2)。

[0156] 回到对于图 4 的描述。

[0157] 第 1 次迭代：

[0158] • 步骤 1c：

[0159] 现在定义第二编码速率  $D_1$ ，其中  $D_1 = D_0 - \delta_0$ ，以及该编码速率  $D_1$  在有待编码的要素  $A(k, j)$ ， $k = 1$  到  $Q$  且  $j = 0$  到  $M-1$  间的分配。

[0160] • 步骤 1d：

[0161] 然后，对于每个配对  $(k, j) \in E_0$ ，根据  $E_0$  的配对的词典编纂的 (lexicographical) 读取顺序，连续地考虑从配对 (1, 0) 直到配对 (Q, M-1)，重复以下操作 a1 到 a7：

[0162] a1- 对于操作 a2 到 a5 认为子带  $(k, j)$  被删除；

[0163] a2- 有待编码的要素  $A(i, n)$  ( $(i, n)$  等于除了配对  $(k, j)$  之外的每个  $E_0$  配对) 被量化模块 10 量化，其中  $(i, n) \in E_0 \setminus (k, j)$ ，根据在所述有待编码的要素  $A(i, n)$  之间的对速率  $D_i$  的已定义的分配，其中  $(i, n) \in E_0 \setminus (k, j)$ ；

[0164] a3- 与步骤 1b 中所示的方法相同，基于产生于步骤 a2 中执行的量化操作的要素  $\bar{A}(i, n)$ ， $(i, n) \in E_0 \setminus (k, j)$ ，计算矩阵  $\text{MNR}_{k,j}(1, D_1) = [\text{MNR}_{k,j}(1, D_1)(i, t)]_{i=1 \text{ 到 } N \text{ 且 } t=0 \text{ 到 } M-1}$ ，这样每个矩阵的元素  $\text{MNR}_{k,j}(1, D_1)(i, t)$  表示对于每个信号和每个子带的屏蔽量化误差（或噪音）比，其中  $i = 1$  到  $N$  且  $t = 0$  到  $M-1$ ，遵循在步骤 a2 中执行的量化（所考虑的子带  $(k, j)$  被删除，量化噪音  $b(k, j)$  在计算时被认为是 0）。矩阵  $\text{MNR}_{k,j}(1, D_1)$  的元素的取值被保存；

[0165] a4- 接着，计算并存储屏蔽量化误差比的变化矩阵  $\Delta \text{MNR}_{k,j}(1)$ ， $\Delta \text{MNR}_{k,j}(1) = |\text{MNR}_{k,j}(1, D_1) - \overline{\text{MNR}}_{k,j}(0, D_0)|$ ，其中  $\overline{\text{MNR}}_{k,j}(0, D_0)$  是指示其要素  $(k, j)$  已被删除的矩阵  $\text{MNR}(0, D_0)$ 。

[0166] a5- 计算矩阵  $\Delta \text{MNR}_{k,j}(1)$  的范数 (norm)  $|| \Delta \text{MNR}_{k,j}(1) ||$ 。该范数的值评价了在信

号  $S_i$  的信噪比的集合上,在有待编码的要素  $A(i, n)$  中删除成份  $A(k, j)$  的影响,其中  $(i, n) \in E_0$ 。

[0167] 计算的范数使得测量  $MNR_{k,j}(1, D_1)$  和  $\overline{MNR}_{k,j}(0, D_0)$  之间的差别成为可能,例如等于矩阵  $\Delta MNR_{k,j}(1)$  的每个要素的平方的和的均方根。

[0168] a6- 考虑子带  $(k, j)$  不再被删除;

[0169] a7- 如果  $(k, j) \neq \max E_0 = (Q, M-1)$ , 在  $E_0$  中增大配对  $(k, j)$ , 并且步骤 a1 到 a7 被再重复直到达到  $\max E_0$ 。

[0170] • 步骤 1e:

[0171] 根据当  $(k, j) \in E_0$  时获得的  $\|\Delta MNR_{k,j}(1)\|$  值中的最小值, 确定  $(i_1, j_1)$ , 也就是  $(i_1, j_1) = \arg \min_{(k,j) \in E_0} \|\Delta MNR_{k,j}(1)\|$ 。

[0172] 有待编码的要素  $A(i_1, j_1)$  因此被识别为在有待编码的要素  $A(i, j)$  集合中的针对总体声音质量的最小相关要素, 其中  $(i, j) \in E_0$ 。

[0173] • 步骤 1f:

[0174] 配对  $(i_1, j_1)$  的标识符被输出到排序模块 6, 作为过程 Proc1 的第一迭代的结果。

[0175] • 步骤 1g:

[0176] 在过程 Proc1 的剩余过程中, 在有待编码的要素集合中删除频带  $(i_1, j_1)$ 。得到集合  $E_1 = E_0 \setminus \{(i_1, j_1)\}$ 。

[0177] 第 2 次迭代及后续:

[0178] 针对每个  $n$ , 执行与步骤 1c 到 1g 相似的步骤, 其中  $n \geq 2$ , 在下面进行描述。

[0179] 步骤 1c: 定义第  $(n+1)$  个编码速率  $D_n$ , 其中  $D_n = D_{n-1} - \delta_{n-1}$ , 以使  $\delta_{n-1} = \min(d_{ij})$ , 其中  $(i, j) \in E_{n-1}$ 。

[0180] 步骤 1d: 接着, 对于每个配对  $(k, j) \in E_{n-1}$ , 根据词典编纂的顺序, 连续地考虑, 重复以下操作 a1 到 a7:

[0181] a1- 对于操作 a2 到 a5 认为子带  $(k, j)$  被删除;

[0182] a2- 有待编码的要素  $A(i, n)$  被量化模块 10 量化, 其中  $(i, n) \in E_{n-1} \setminus \{(k, j)\}$ , 根据在所述有待编码的要素  $A(i, n)$  之间的对速率  $D_n$  的分配, 其中  $(i, n) \in E_{n-1} \setminus \{(k, j)\}$ ;

[0183] a3- 基于根据步骤 a2 中的量化所确定的要素  $\bar{A}(i, n)$ ,  $(i, n) \in E_{n-1} \setminus \{(k, j)\}$ , 计算矩阵  $MNR_{k,j}(n, D_n)$ , 其表示对于每个信号和每个子带的屏蔽量化误差 (或噪音) 比, 其中  $i = 1$  到  $N$  且  $j = 0$  到  $M-1$ , 遵循在步骤 a2 中执行的量化;

[0184] a4- 接着, 计算并存储屏蔽量化误差比的变化矩阵  $\Delta MNR_{k,j}(n) = |MNR_{k,j}(n, D_n) - \overline{MNR}_{k,j}(n-1, D_{n-1})|$ , 其中  $\overline{MNR}_{k,j}(n-1, D_{n-1})$  对应于指示要素  $(k, j)$  已被删除的矩阵  $\overline{MNR}(n-1, D_{n-1})$ , 以及矩阵  $\Delta MNR_{k,j}(n)$  的范数  $\|\Delta MNR_{k,j}(n)\|$ 。该范数的值评价了在信号  $S_i$  的信噪比的集合上, 在有待编码的要素  $A(i, n)$  中删除成份  $A(k, j)$  的影响, 其中  $(i, n) \in E_{n-1} \setminus \{(k, j)\}$ 。

[0185] a5- 认为子带  $(k, j)$  不再被删除;

[0186] a6- 如果  $(k, j) \neq \max E_{n-1}$ , 在  $E_{n-1}$  中增大配对  $(k, j)$ , 并且步骤 a1 到 a6 被再重复直到达到  $\max E_{n-1}$ 。

[0187] • 步骤 1e:

[0188] 根据当  $(k, j) \in E_{n-1}$  时获得的  $\|\Delta MNR_{k,j}(n)\|$  值中的最小值, 确定  $(i_n, j_n)$ , 也就

是  $(i_n, j_n) = \arg \min_{(k,j) \in E_{n-1}} \|\Delta \text{MNR}_{k,j}(n)\|$ 。矩阵  $\overline{\text{MNR}}(n, D_n) = \text{MNR}_{i_n, j_n}(n, D_n)$  也被存储。

[0189] 有待编码的要素  $A(i_n, j_n)$  因此被识别为在有待编码的要素  $A(i, j)$  集合中的针对总体声音质量的最小相关要素, 其中  $(i, j) \in E_{n-1}$ 。

[0190] • 步骤 1f :

[0191] 配对  $(i_n, j_n)$  的标识符被输出到排序模块 6, 作为过程 Proc1 的第一迭代的结果。

[0192] • 步骤 1g :

[0193] 在过程 Proc1 的剩余过程中, 在有待编码的要素集合中删除频带  $(i_n, j_n)$ 。得到集合  $E_n = E_{n-1} \setminus \{(i_n, j_n)\}$ 。

[0194] 过程 Proc1 被再迭代  $r$  次, 最大为  $Q * M - 1$  次。

[0195] 然后, 由排序模块 6 将优先级次序分配给不同的频带, 以便将编码数据插入到二进制序列中。

[0196] 对有待编码的要素进行排序并构造二进制序列

[0197] 基于由过程 Proc1 的连续迭代所提供的连续性的结果 :

[0198] 在一个实施例中, 有待编码的要素的排序是通过排序模块 6 进行的, 仅仅基于模块 5 执行的为了定义有待编码的最小相关要素而由处理步骤 Proc1 的连续重复所连续提供的结果, 排除了处理步骤 Proc2 所提供的结果, 后者确定了所述有待编码要素的次序, 反映了涉及总体声音质量的有待编码的重要性。

[0199] 参考图 5a, 对应于配对  $(i_1, j_1)$  的有待编码的要素  $A(i_1, j_1)$  在 Proc1 的第一次重复过程中, 被认为是与总体声音质量最小相关。它因此被模块 5 分配了一个最小优先级指数  $\text{Priol}$ 。

[0200] 对应于配对  $A(i_2, j_2)$  的有待编码的要素在 Proc1 的第二次重复过程中, 被认为是与总体声音质量最小相关, 在分配了优先级  $\text{Priol}$  之后。它因此被分配了一个最小优先级指数  $\text{Prio2}$ , 其中  $\text{Prio2} > \text{Priol}$ 。当步骤的迭代次数  $r$  严格小于  $Q * M - 1$  时, 排序模块 6 因此连续地确定  $r$  个有待编码的要素, 每个要素被分配的优先级指数从  $\text{Priol}$ ,  $\text{Prio2}$  增加到  $\text{Prio } r$ 。在处理步骤 Proc1 中还没有被分配优先级次序的有待编码的要素对于总体声音质量比已经分配了优先级次序的有待编码的要素的重要程度更高。

[0201] 当  $r$  等于  $Q * M - 1$  次时, 所有有待编码的要素被一个接一个地排序。

[0202] 在此情况下, 过程 Proc1 所进行的迭代次数被认为是等于  $Q * M - 1$  次。

[0203] 分配给有待编码的要素  $A(k, j)$  的优先级次序同样被分配给产生于该有待编码的要素的量化的已被编码的要素  $\bar{A}(k, j)$ 。

[0204] 用于构造二进制序列的模块 8 构造了一个二进制序列, 与每个信号  $S_i$  的帧相应, 其中  $i = 1$  到  $N$ , 通过连续地对已被编码的要素  $\bar{A}(k, j)$  求积分, 减小分配的优先级指数的次序, 码流  $\Phi$  中的二进制序列将被发送。

[0205] 这样, 构造的二进制序列按照模块 6 执行的排序步骤被排序。

[0206] 二进制序列被与各个频带相关的频谱成份构造, 这些成份是产生于一个包括  $N$  个信号的声音场景的有待编码的要素中的, 其中  $N > 1$ , 并且根据在频带上确定的对于遮噪比的影响进行排序。

[0207] 二进制序列的频谱成份例如是根据本发明中的方法进行排序的。

[0208] 在一个实施例中, 在包括在构造的二进制序列中的仅仅一些频谱成份使用依据本

发明的方法进行排序。

[0209] 在上述实施例中,对于算法 Proc1 的每一次迭代,都要从有待编码的要素 A(i, j) 中删除一个频谱成份。

[0210] 在另一个实施例中,使用了重叠量化器用于执行量化操作。在这种情况下,一个已被识别的有待编码的要素 A(i<sub>0</sub>, j<sub>0</sub>) 中的频谱成份没有被删除,但是该成份的编码被分配了一个减小了的速率,至于有待编码的要素中的其它频谱成份仍然要被排序。

[0211] 编码器 1 是一个考虑不同单频信号间的干扰的因而允许速率适配性的编码器。它允许定义压缩数据对感知总体声音质量进行优化。

[0212] 对二进制序列的要素的排序与二进制序列的构造使用了上述的本发明实施例中的 Proc1,其中有待编码的要素包括信号的环境声成份。

[0213] 在另一个实施例中,根据本发明的一个编码器没有对这些环境声成份进行编码,而是对信号 S<sub>i</sub> 的频谱系数 X(i, j) 进行编码,其中 j = 0 到 M。

[0214] 在这种情况下,在过程 1 的第一次重复时,例如最小优先级指数(在剩余的有待编码的要素中的最小值)被分配给有待编码的要素,这样删除频谱成份就引起了遮噪比的最小变化。接着,过程 Proc1 被重复。

[0215] 处理步骤 Proc2

[0216] 葛容准则一般被用于表征虚拟声源的位置,通过合成从一个已知声音表现系统的喇叭中还原的信号。

[0217] 这些准则基于研究声压的速度和能量向量,该声压通过使用一个声音表现系统而产生。

[0218] 当一个声音表现系统包括 L 个喇叭时,通过这些喇叭产生的信号 S<sub>i</sub>, i = 1 到 L, 通过一个声压 T<sub>i</sub> 以及一个声音传播角度 ξ<sub>i</sub> 定义。

[0219] 然后速度向量  $\vec{V}$  被这样定义:

$$[0220] \quad \vec{V} = \begin{cases} x_v = \frac{\sum_{1 \leq i \leq L} T_i \cos \xi_i}{\sum_{1 \leq i \leq L} T_i} \\ y_v = \frac{\sum_{1 \leq i \leq L} T_i \sin \xi_i}{\sum_{1 \leq i \leq L} T_i} \end{cases}$$

[0221] 存在一对极坐标 (r<sub>v</sub>, ξ<sub>v</sub>) 以使:

$$[0222] \quad \vec{V} = \begin{cases} x_v = \frac{\sum_{1 \leq i \leq L} T_i \cos \xi_i}{\sum_{1 \leq i \leq L} T_i} = r_v \cos \xi_v \\ y_v = \frac{\sum_{1 \leq i \leq L} T_i \sin \xi_i}{\sum_{1 \leq i \leq L} T_i} = r_v \sin \xi_v \end{cases} \quad \text{等式 (3)}$$

[0223] 能量向量  $\vec{E}$  被定义为:

$$[0224] \quad \vec{E} = \begin{cases} x_E = \frac{\sum_{1 \leq i \leq L} T_i^2 \cos \xi_i}{\sum_{1 \leq i \leq L} T_i^2} \\ y_E = \frac{\sum_{1 \leq i \leq L} T_i^2 \sin \xi_i}{\sum_{1 \leq i \leq L} T_i^2} \end{cases}$$

[0225] 存在一对极坐标  $(r_F, \xi_F)$  以使：

$$[0226] \quad \vec{E} = \begin{cases} x_E = \frac{\sum_{1 \leq i \leq L} Ti^2 \cos \xi_i}{\sum_{1 \leq i \leq L} Ti^2} = r_E \cos \xi_E \\ y_E = \frac{\sum_{1 \leq i \leq L} Ti^2 \sin \xi_i}{\sum_{1 \leq i \leq L} Ti^2} = r_E \sin \xi_E \end{cases} \quad \text{等式 (4)}$$

[0227] 为定位适宜的虚拟声源所需要的条件通过寻找角度  $\xi_i$  而确定, 表征所述声音表现系统的喇叭的位置, 符合以下所述葛容准则, 其中：

[0228] - 准则 1, 与低频下的声源 S 的声音图像的精度有关： $\xi_V = \xi$ ；其中  $\xi$  是寻求获得的真正声源 S 的传播角度。

[0229] - 准则 2, 与低频下的声源 S 的声音图像的稳定性有关： $r_V = 1$ ；

[0230] - 准则 3, 与高频下的声源 S 的声音图像的精度有关： $\xi_E = \xi$ ；

[0231] - 准则 4, 与高频下的声源 S 的声音图像的稳定性有关： $r_E = 1$ ；

[0232] 在下面描述的本发明的一个实施例的操作中, 使用了葛容向量 (Gerzonvector), 除了包括寻找表示所述声音表现系统的喇叭位置的最佳角度  $\xi_i$ 。

[0233] 葛容准则基于研究声压的速度和能量向量, 该声压通过使用一个声音表现系统而产生。

[0234] 等式 3 和 4 中所示的关于能量和速度向量的与葛容准则相关的每个系数  $x_V, y_V, x_E, y_E$  都是  $[-1, 1]$  的元素。因此存在一个简单的配对  $(\xi_V, \xi_E)$  验证以下等式, 符合理想的情况  $(r_V, r_E) = (1, 1)$ ：

$$[0235] \quad \frac{\sum_{1 \leq i \leq L} Ti \cos \xi_i}{\sum_{1 \leq i \leq L} Ti} = \cos \xi_V, \quad \frac{\sum_{1 \leq i \leq L} Ti \sin \xi_i}{\sum_{1 \leq i \leq L} Ti} = \sin \xi_V,$$

$$[0236] \quad \frac{\sum_{1 \leq i \leq L} Ti^2 \cos \xi_i}{\sum_{1 \leq i \leq L} Ti^2} = \cos \xi_E \text{ 且 } \frac{\sum_{1 \leq i \leq L} Ti^2 \sin \xi_i}{\sum_{1 \leq i \leq L} Ti^2} = \sin \xi_E.$$

[0237] 该简单配对中的角度  $\xi_V$  和  $\xi_E$  因此被以下等式定义 (等式 (5))：

$$[0238] \quad \xi_V = \text{sign}\left(\frac{\sum_{1 \leq i \leq L} Ti \sin \xi_i}{\sum_{1 \leq i \leq L} Ti}\right) \cdot \arccos\left(\frac{\sum_{1 \leq i \leq L} Ti \cos \xi_i}{\sum_{1 \leq i \leq L} Ti}\right)$$

$$[0239] \quad \xi_E = \text{sign}\left(\frac{\sum_{1 \leq i \leq L} Ti^2 \sin \xi_i}{\sum_{1 \leq i \leq L} Ti^2}\right) \cdot \arccos\left(\frac{\sum_{1 \leq i \leq L} Ti^2 \cos \xi_i}{\sum_{1 \leq i \leq L} Ti^2}\right)$$

[0240] 此后, 一般的葛容角度向量术语将用于指代向量  $\vec{\xi}$ , 以使  $\vec{\xi} = \begin{pmatrix} \xi_V \\ \xi_E \end{pmatrix}$ 。

[0241] 第二算法包括用于执行的指令, 当它们在模块 5 的处理装置上执行时, 参考图 6 下面描述的过程 Proc2 的步骤的指令。

[0242] 过程 Proc2 的原理如下：对一个要被排序的频谱参数集合中的每一个频谱参数, 计算其对于一个角度向量的影响, 该角度向量根据与葛容准则相关的能量与速度向量而被定义, 并根据一个对于所述量化环绕声成份的逆环绕声变换而被计算。此外, 对至少一个频谱参数分配优先级次序, 根据一个对所述频谱参数所计算的影响与其它计算的影响相比较。

[0243] 在一个实施例中,过程 Proc2 具体如下所述:

[0244] 初始化 ( $n = 0$ )

[0245] • 步骤 2a:

[0246] 一个速率  $D_0 = D_{\max}$  以及在有待编码的要素  $A(k, j)$  间分配的该速率被定义,其中  $(k, j) \in E_0 = \{(k, j) \text{ 以使 } k = 1 \text{ 到 } Q \text{ 且 } j = 0 \text{ 到 } M-1\}$ 。

[0247] 在该初始分配中分配给有待编码的要素的速率被定义为  $d_{k,j}$  (这些速率的和  $d_{k,j} | i = 1 \text{ 到 } Q, j = 0 \text{ 到 } M-1$  等于  $D_0$ ), 并且  $\delta_0 = \min d_{k,j}$ , 其中  $(k, j) \in E_0$ 。

[0248] • 步骤 2b:

[0249] 然后,每个有待编码的要素  $A(k, j)$  被量化模块 10 根据已经在步骤 2a 中被分配的速率  $d_{k,j}$  进行量化,其中  $(k, j) \in E_0$ 。

[0250]  $\bar{A}$  是要素  $\bar{A}(k, j)$  的矩阵,其中  $k = 1 \text{ 到 } Q$  且  $j = 0 \text{ 到 } M-1$ 。每个要素都是量化的结果,环绕声成份  $A(k)$  的参数  $A(k, j)$  的速率  $d_{k,j}$  与频带  $F_j$  有关。因此要素  $\bar{A}(k, j)$  定义对于所述环绕声成份  $A_k$  的频带  $F_j$  的频谱表示的量化值。

$$[0251] \quad \bar{A} = \begin{bmatrix} \bar{A}(1,0) & \bar{A}(1,1) & . & . & . & \bar{A}(1,M-1) \\ \bar{A}(2,0) & & & & & \bar{A}(2,M-1) \\ . & & & & & . \\ . & & & & & . \\ \bar{A}(Q,0) & \bar{A}(Q,1) & . & . & . & \bar{A}(Q,M-1) \end{bmatrix},$$

[0252] • 步骤 2c:

[0253] 然后,这些被量化的环绕声成份  $\bar{A}(k, j)$  ( $k = 1 \text{ 到 } Q$  且  $j = 0 \text{ 到 } M-1$ ) 被进行  $p$  阶的环绕声解码,以使  $2p+1 = Q$ , 并且与一个具有  $N$  个喇叭的规则系统相符合,目的是确定作为该环绕声解码的结果而获取的声音信号的声压  $T1i$ ,  $i = 1 \text{ 到 } N$ 。

[0254] 在所属情况下,  $AmbInv(p)$  是  $p$  阶逆环绕声变换矩阵 (或者  $p$  阶环绕声解码), 输出与  $N$  个不同的喇叭  $H'_1, \dots, H'_N$  相应的  $N$  个信号  $T11, \dots, T1N$ , 这些喇叭围绕一个点规则地安装。结果,矩阵  $AmbInv(p)$  通过矩阵  $Amb(p, N)$  的变换推导而来,矩阵  $Amb(p, N)$  是来自于声音场景的编码的环绕声编码矩阵,该场景由与  $N$  个喇叭  $H'_1, \dots, H'_N$  相应的  $N$  个声源定义,这些喇叭分别置放在位置  $\xi_1, \dots, \xi_N$ 。因此我们可以得到:

$$AmbInv(p) = \frac{1}{N} Amb(p, N)'。$$

[0255]  $T1$  是与频带  $F_j$  ( $j = 0 \text{ 到 } M-1$ ) 相关的信号  $T1i$  ( $i = 1 \text{ 到 } N$ ) 的频谱成份的矩阵。这些频谱成份来自  $p$  阶环绕声逆变换,应用于量化环绕声成份  $\bar{A}(k, j)$ ,  $k = 1 \text{ 到 } Q$  且  $j = 0 \text{ 到 } M-1$ 。

$$[0256] \quad T1 = \begin{bmatrix} T1(1,0) & T1(1,1) & . & . & . & T1(1,M-1) \\ T1(2,0) & T1(2,1) & . & . & . & T1(2,M-1) \\ . & & & & & . \\ . & & & & & . \\ . & & & & & . \\ T1(N,0) & . & . & . & . & T1(N,M-1) \end{bmatrix}$$

[0257] 接着我们有

[0258]  $T1 = \text{AmbInv}(\underline{p}) \times \underline{\bar{A}} = \frac{1}{N} \text{Amb}(p, N)' \times \underline{\bar{A}}$  等式 (6)

[0259] 这样成份  $T1(i, j)$  ( $i = 1$  到  $N$ ) 取决于与所述环绕声成份  $A(k, j)$  ( $k = 1$  到  $Q$  且  $j = 0$  到  $M-1$ ) 的量化相关的量化误差 (事实上, 每一量化要素都是有待量化的环绕声成份的频谱参数以及与所述参数相关的量化噪音的和)。

[0260] 对每一频带  $F_j$ ,  $j = 0$  到  $M-1$ , 使用等式 (5), 接着在过程 Proc2 ( $n = 0$ ) 的初始化时计算一般葛容角度向量  $\tilde{\xi}_j(0)$ , 根据一个由下面的环绕声解码  $\tilde{\xi}_j(0) = \begin{pmatrix} \xi_{vj} \\ \xi_{ej} \end{pmatrix}$  (其由  $\xi_i = \frac{2\pi(i-1)}{N}$ ,  $i = 1$  到  $N$ ) 所确定的频谱成份  $T1(i, j)$  ( $i = 1$  到  $N$  且  $j = 0$  到  $M-1$ )。

$$[0261] \quad \xi_{vj} = \text{sign}\left(\frac{\sum_{1 \leq i \leq N} T1(i, j) \sin \xi_i}{\sum_{1 \leq i \leq N} T1(i, j)}\right) \cdot \arccos\left(\frac{\sum_{1 \leq i \leq N} T1(i, j) \cos \xi_i}{\sum_{1 \leq i \leq N} T1(i, j)}\right)$$

$$[0262] \quad \xi_{ej} = \text{sign}\left(\frac{\sum_{1 \leq i \leq Q} T1(i, j)^2 \sin \xi_i}{\sum_{1 \leq i \leq Q} T1(i, j)^2}\right) \cdot \arccos\left(\frac{\sum_{1 \leq i \leq Q} T1(i, j)^2 \cos \xi_i}{\sum_{1 \leq i \leq Q} T1(i, j)^2}\right)$$

[0263] 从而确定  $\tilde{\xi}_j(0) = \tilde{\xi}_j(0)$ 。

[0264] 需要指出, 这里的环绕声解码矩阵已经被考虑用于一个规则声音表现装置, 其中包括与输入信号数目相等的喇叭, 以使对于环绕声解码矩阵的计算简单。然而, 该步骤对于一个与不规则声音表现装置相应的环绕声解码矩阵, 以及与输入信号数目不相等的喇叭也可以实现。

[0265] 第 1 次迭代 ( $n = 1$ )

[0266] • 步骤 2d

[0267] 速率  $D_1 = D_0 - \delta_0$  以及在有待编码的要素  $A(k, j)$  间分配的该速率  $D_1$  被定义, 其中  $(k, j) \in E_0$ 。

[0268] • 步骤 2e:

[0269] 然后, 每个有待编码的要素  $A(k, j)$  被量化模块 10 根据已经在步骤 2d 中分配的速率  $d_{k,j}$  进行量化, 其中  $(k, j) \in E_0$ 。

[0270]  $\underline{\bar{A}}$  是量化要素  $\bar{A}(k, j)$  的更新矩阵, 其中  $(k, j) \in E_0$ , 每个都来自于根据参数  $A(k, j)$  的总体速率  $D_1$  的最后一次量化。

[0271] • 步骤 2f:

[0272] 以与前述步骤 2c 中相似的方式, 在根据对于总体速率为  $D_1$  的要素执行的  $p$  阶新环绕声解码计算之后, 对于过程 Proc2 的第一次迭代, 对于每个频带  $F_j$  计算其中的第一一般葛容角度向量  $\tilde{\xi}_j(1)$ , 根据从上述新环绕声解码确定的频谱成份  $T1(i, j)$  ( $i = 1$  到  $N$ ,  $j = 0$  到  $M-1$ ), 使用等式 (6)。

[0273] 接着对于向量  $\Delta \tilde{\xi}_j(1)$  进行计算, 该向量等于在初始化步骤 2c 中计算的葛容角度向量  $\tilde{\xi}_j(0)$  与在第 1 次迭代的初始化步骤 2f 中计算的一般葛容角度向量  $\tilde{\xi}_j(1)$  之间的差:  $\Delta \tilde{\xi}_j(1) = \tilde{\xi}_j(1) - \tilde{\xi}_j(0)$ ,  $j = 0$  到  $M-1$ 。

[0274] • 步骤 2g:

[0275] 计算每一频带  $F_j$  中的变化  $\Delta \tilde{\xi}_j(1)$ ,  $j = 0$  到  $M-1$  的范数  $\|\Delta \tilde{\xi}_j(1)\|$ 。

[0276] 该范数表示在每一频带  $F_j$  的速率从  $D_0$  到  $D_1$  的下降时一般葛容角度向量的变化。

[0277] 频带  $F_{j_1}$  的指数  $j_1$  这样被确定, 以使在频带  $F_{j_1}$  中计算的葛容角度变化的范数  $\|\Delta \tilde{\xi}_{j_1}(1)\|$  小于或者等于对于每个频带  $F_j$  ( $j = 0$  到  $M-1$ ) 计算的每个范数  $\|\Delta \tilde{\xi}_j(1)\|$ 。我们因此可以得到  $j_1 = \arg \min_{j=0, \dots, M-1} \|\Delta \tilde{\xi}_j(1)\|$ 。

[0278] • 步骤 2h:

[0279] 现在考虑相对于频带  $F_{j_1}$  的环绕声成份的频谱参数, 也就是参数  $A(k, j_1)$ , 其中  $k \in F_0 = [1, Q]$ 。

[0280] 对任何一个所述  $i \in F_0$ , 都从 1 到  $Q$  依次重复下面的步骤 2h1 到 2h5:

[0281] 2h1- 认为对于步骤 2h2 到 2h4 都删除子带  $(i, j_1)$ ; 因此, 可以认为  $A(i, j_1)$  为 0 并且相应的量化要素  $\bar{A}(i, j_1)$  为 0;

[0282] 2h2- 以与前述步骤 2c 相似的方式, 在根据对于总体速率为  $D_1$  ( $\bar{A}(i, j_1)$  为 0) 的要素执行的  $p$  阶环绕声解码计算之后, 在频带  $F_{j_1}$  中确定一般葛容角度向量  $\tilde{\xi}_{j_1}(A(i, j_1) = 0, 1)$ , 根据所述环绕声解码确定的频谱成份  $T1(i, j)$  ( $i = 1$  到  $N$ ,  $j = 0$  到  $M-1$ ), 使用等式 (6)。

[0283] 2h3- 接下来计算向量  $\Delta \tilde{\xi}_{j_1}(1)$ , 该向量代表频带  $F_{j_1}$  中的上述计算的一般葛容角度向量  $\tilde{\xi}_{j_1}(A(i, j_1) = 0, 1)$  以及在上述第 1 次迭代的步骤 2f 中计算的一般葛容角度向量  $\tilde{\xi}_{j_1}(1)$  之间的差别:  $\Delta \tilde{\xi}_{j_1}(1) = \tilde{\xi}_{j_1}(A(i, j_1) = 0, 1) - \tilde{\xi}_{j_1}(1)$ 。然后, 计算向量  $\Delta \tilde{\xi}_{j_1}(1)$  的范数

[0284]  $\|\Delta \tilde{\xi}_{j_1}(1)\|$ :  $\|\Delta \tilde{\xi}_{j_1}(1)\| = \|\tilde{\xi}_{j_1}(A(i, j_1) = 0, 1) - \tilde{\xi}_{j_1}(1)\|$ 。

[0285] 该范数代表对于速率  $D_1$  频率环绕声成份  $A(i, j_1)$  被删除时, 频带  $F_{j_1}$  中的一般葛容角度向量的变化。

[0286] 2h4- 如果  $i \neq \max F_0$ , 则认为子带  $(i, j_1)$  不再被删除, 并且我们转到步骤 2h5。如果  $i = \max F_0$ , 认为子带  $(i, j_1)$  不再被删除, 并且我们转到步骤 2i。

[0287] 2h5- 在集合  $F_0$  中增加  $i$ , 且重复步骤 2h1 到 2h4, 直到  $i = \max F_0$ 。

[0288] 因此对于每一  $i \in F_0 = [1, Q]$ , 可以获得  $Q$  个一般葛容角度变化值  $\|\Delta \tilde{\xi}_{j_1}(1)\|$ 。

[0289] • 步骤 2i:

[0290] 对于每一个  $i \in F_0 = [1, Q]$ , 互相比对  $\|\Delta \tilde{\xi}_{j_1}(1)\|$  的值, 确定这些值中的最小值, 并且确定相应于最小值的指数  $i_n \in F_n$ , 也就是。

[0291] 因此成份  $A(i_1, j_1)$  被标识为有待编码的要素中对于空间精度的重要性最低的成份, 相比于其它有待编码的要素  $A(k, j)$ ,  $(k, j) \in E_0$ 。

[0292] • 步骤 2j:

[0293] 对于每个频带  $F_j$ , 重新定义并计算针对速率  $D_1$  的来自于第 1 次迭代的一般葛容角度向量  $\tilde{\xi}_j(1)$ :

[0294]  $\tilde{\xi}_j(1) = \tilde{\xi}_j(1)$ , 如果  $j \in [0, M-1] \setminus \{j_1\}$ ;

[0295]  $\tilde{\xi}_{j_1}(1) = \tilde{\xi}_{j_1}(A(i_1, j_1) = 0, 1)$ , 如果  $j = j_1$ 。

[0296] 该重新定义的一般葛容角度向量, 针对等于  $D_1$  的量化速率建立, 考虑到有待编码要素  $A(i_1, j_1)$  的删除, 并且将被用于下面的过程 Proc2 的重复。

[0297] • 步骤 2k :

[0298] 配对  $(i_1, j_1)$  的标识符被传送到排序模块 6, 作为过程 Proc2 的第一次迭代的结果。

[0299] • 步骤 2m :

[0300] 然后, 有待编码要素  $A(i_1, j_1)$  被从过程 Proc2 中的剩余部分的有待编码的要素集合中删除。

[0301] 定义序列  $E_1 = E_0 \setminus (i_1, j_1)$ 。

[0302] 定义  $\delta_1 = \min d_{k,j}$ , 其中  $(k, j) \in E_1$ 。

[0303] 在过程 Proc2 的第二次迭代中, 重复与上述步骤 2d 到 2n 相似的步骤。

[0304] 过程 Proc2 被根据需要尽可能多地重复, 以对剩余的要排序的有待编码的要素  $A(k, j)$ ,  $(k, j) \in E_1$  中的一些或者所有进行排序。

[0305] 因此, 上述步骤 2d 到 2h 被进行第  $n$  次重复。

[0306] 第  $n$  次迭代 ( $n > 1$ ) :

[0307]  $E_{n-1} = E_0 \setminus \{(i_1, j_1), \dots, (i_{n-1}, j_{n-1})\}$ 。

[0308] 有待编码的要素  $A(k, j)$  (其中  $(k, j) \in E_0 \setminus E_{n-1}$ ) 已经在前一次迭代的步骤 2m 中被删除。

[0309] • 步骤 2d

[0310] 速率  $D_n = D_{n-1} - \delta_{n-1}$  以及在有待编码的要素  $A(k, j)$  间分配的该速率  $D_n$  被定义, 其中  $(k, j) \in E_{n-1}$ 。

[0311] 在下面进行的计算环绕声解码的过程中, 认为被量化的要素  $\bar{A}(k, j)$  (其中  $(k, j) \in E_0 \setminus E_{n-1}$ ) 为 0。

[0312] • 步骤 2e :

[0313] 然后, 每个有待编码的要素  $A(k, j)$  被量化模块 10 根据已经在上述步骤 2d 中分配的速率进行量化, 其中  $(k, j) \in E_{n-1}$ 。

[0314] 该有待编码的要素  $A(k, j)$  的量化结果为  $\bar{A}(k, j)$ ,  $(k, j) \in E_{n-1}$ 。

[0315] • 步骤 2f :

[0316] 以与前述第 1 次迭代中相似的方式, 在根据对于总体速率为  $D_n$  的要素执行的  $p$  阶环绕声解码计算之后 (因此认为在该环绕声解码期间成份  $\bar{A}(i_1, j_1), \dots, \bar{A}(i_{n-1}, j_{n-1})$  为 0), 对于过程 Proc2 的第  $n$  次重复, 对于每个频带  $F_j$  计算其中的第一一般葛容角度向量  $\bar{\xi}_j(n)$ , 根据从上述环绕声解码确定的频谱成份  $T1i$ ,  $i = 1$  到  $N$ , 使用等式 (6)。

[0317] 接着对于向量  $\Delta \bar{\xi}_j(n)$  进行计算, 该向量等于在第  $n-1$  次迭代的步骤 2j 中计算的葛容角度向量  $\tilde{\xi}_j(n-1)$  与在当前步骤中计算的一般葛容角度向量  $\bar{\xi}_j(n)$  之间的差:  $\Delta \bar{\xi}_j(n) = \bar{\xi}_j(n) - \tilde{\xi}_j(n-1)$   $j = 0$  到  $M-1$ 。

[0318] • 步骤 2g :

[0319] 计算每一频带  $F_j$  中的变化  $\Delta \bar{\xi}_j(n)$ ,  $j = 0$  到  $M-1$  的范数  $\|\Delta \bar{\xi}_j(n)\|$ 。

[0320] 该范数表示在每一频带  $F_j$  的速率从  $D_n$  到  $D_{n-1}$  的下降时一般葛容角度向量的变化 (参数  $A(i_1, j_1), \dots, A(i_{n-1}, j_{n-1})$  以及  $\bar{A}(i_1, j_1), \dots, \bar{A}(i_{n-1}, j_{n-1})$  被删除)。

[0321] 频带 $F_{j_n}$ 的索引 $j_n$ 这样被确定,以使在频带 $F_{j_n}$ 中计算的葛容角度变化的范数 $\|\Delta\tilde{\xi}_{j_n}(n)\|$ 小于或者等于对于每个频带 $F_j$  ( $j = 0$  到  $M-1$ ) 计算的每个范数 $\|\Delta\tilde{\xi}_j(n)\|$ 。我们因此可以得到 $j_n = \arg \min_{j=0 \dots M-1} \|\Delta\tilde{\xi}_j(n)\|$ 。

[0322] • 步骤 2h :

[0323] 现在考虑相对于频带 $F_{j_n}$ 的环绕声成份的频谱参数,也就是参数 $A(k, j_n)$ ,其中 $k \in F_{n-1} = \{i \in [1 \dots, Q]\}$ 以使 $(i, j_n) \in E_{n-1}$ 。

[0324] 对任何一个所述 $i \in F_{n-1}$ ,从序列 $F_{n-1}$ 中的最小元素( $\min F_{n-1}$ )到序列 $F_{n-1}$ 中的最大元素( $\max F_{n-1}$ )依次重复下面的步骤 2h1 到 2h5 :

[0325] 2h1- 认为对于步骤 2h2 到 2h4 都删除子带 $(i, j_n)$ ;因此,可以认为 $A(i, j_n)$ 为 0 并且相应的量化要素 $\bar{A}(i, j_n)$ 为 0 ;

[0326] 2h2- 以与前述步骤 2c 相似的方式,在根据对于总体速率为 $D_n$  ( $\bar{A}(i, j_n)$  为 0) 的要素执行 p 阶环绕声解码计算之后,在频带 $F_{j_n}$ 中确定一般葛容角度向量 $\tilde{\xi}_{j_n}(A(i, j_n)=0, n)$ ,根据所述环绕声解码确定的频谱成份 $T1(i, j)$  ( $i = 1$  到  $N$ ,  $j = 0$  到  $M-1$ ),使用等式 (6)。

[0327] 2h3- 接下来计算向量 $\Delta\tilde{\xi}_{j_n}(n)$ ,该向量代表频带 $F_{j_n}$ 中的上述步骤 2h2 计算的一般葛容角度向量 $\tilde{\xi}_{j_n}(A(i, j_n)=0, n)$ 以及在上述第 n 次迭代的步骤 2f 中计算的一般葛容角度向量 $\tilde{\xi}_{j_n}(n)$ 之间的差别: $\Delta\tilde{\xi}_{j_n}(n) = \tilde{\xi}_{j_n}(A(i, j_n)=0, n) - \tilde{\xi}_{j_n}(n)$ 。然后,计算向量 $\Delta\tilde{\xi}_{j_n}(n)$ 的范数 $\|\Delta\tilde{\xi}_{j_n}(n)\|$ :  $\|\Delta\tilde{\xi}_{j_n}(n)\| = \|\tilde{\xi}_{j_n}(A(i, j_n)=0, n) - \tilde{\xi}_{j_n}(n)\|$ 。

[0328] 该范数代表对于速率 $D1$  频率环绕声成份 $A(i, j_n)$ 在过程 Proc 2 的第 n 次迭代中被删除时,频带 $F_{j_n}$ 中的一般葛容角度向量的变化。

[0329] 2h4- 如果 $i \neq \max F_{n-1}$ ,则认为子带 $(i, j_n)$ 不再被删除,并且我们转到步骤 2h5。如果 $i = \max F_{n-1}$ ,认为子带 $(i, j_n)$ 不再被删除,并且我们转到步骤 2i。

[0330] 2h5- 在集合 $F_{n-1}$ 中增加 $i$ ,并且重复步骤 2h1 到 2h4,直到 $i = \max F_{n-1}$ 。

[0331] 因此对于每一 $i \in F_{n-1}$ ,可获得值 $\|\Delta\tilde{\xi}_{j_n}(n)\|$ ,其代表了因为成份 $A(i, j_n)$ 被删除时,在频带 $F_{j_n}$ 中的一般葛容角度向量的变化。。

[0332] • 步骤 2i :

[0333] 对于每一个 $i \in F_{n-1}$ ,互相比较 $\|\Delta\tilde{\xi}_{j_n}(n)\|$ 的值,确定这些值中的最小值,并且确定相应于最小值的指数 $i_1 \in F_n$ ,也就是 $i_n = \arg \min_{i \in F_n} \|\Delta\tilde{\xi}_{j_n}(n)\|$ 。

[0334] 因此成份 $A(i_n, j_n)$ 被标识为有待编码的要素中对于空间精度的重要性最低的成份,相比于其它有待编码的要素 $A(k, j)$ ,  $(k, j) \in E_{n-1}$ 。

[0335] • 步骤 2j :

[0336] 对于每个频带 $F_j$ ,重新定义来自于第 n 次迭代的一般葛容角度向量 $\tilde{\xi}_j(n)$ :

[0337]  $\tilde{\xi}_j(n) = \tilde{\xi}_j(n)$ ,若 $j \in [0, M-1] \setminus \{j_n\}$  ;

[0338]  $\tilde{\xi}_{j_n}(n) = \tilde{\xi}_{j_n}(A(i_n, j_n)=0, n)$ ,若 $j = j_n$ 。

[0339] 该重新定义的一般葛容角度向量,针对等于 $D_n$ 的量化速率建立,考虑到有待编码

的要素  $A(i_n, j_n)$  的删除, 并且将被用于下面的迭代。

[0340] • 步骤  $2k$  :

[0341] 配对  $(i_n, j_n)$  的标识符被传送到排序模块 6, 作为过程 Proc2 的第  $n$  次迭代的结果。

[0342] • 步骤  $2m$  :

[0343] 然后, 有待编码的要素  $(i_n, j_n)$  被从过程 Proc2 中的剩余部分的有待编码的要素集合中删除, 也就是有待编码的要素  $A(i_n, j_n)$  被删除。

[0344] 定义序列  $E_n = E_{n-1} \setminus (i_n, j_n)$ 。有待编码的要素  $A(i, j)$  仍然要被排序, 其中  $(i, j) \in E_n$ 。有待编码的要素  $A(i, j)$ , 其中  $(i, j) \in \{(i_1, j_1), \dots, (i_n, j_n)\}$ , 已经在第 1 到  $n$  迭代中被排序。

[0345] 过程 Proc2 被重复  $r$  次, 且最多为  $Q * M - 1$  次。

[0346] 然后, 排序模块 6 将优先级指数分配给有待编码的要素, 其目的在于将编码数据插入到二进制序列。

[0347] 有待编码的要素的排序以及二进制序列的构造, 基于过程 Proc2 的连续重复所提供的连续结果 :

[0348] 在一个实施例中, 基于由模块 5 所执行的过程 Proc2 的连续重复所提供的连续结果有待编码的要素由排序模块 6 执行排序, 其中模块 5 用于定义有待编码的要素的最小相关要素, (排除由过程 Proc1 提供的结果), 排序模块 6 定义了所述有待编码要素的顺序, 其反映了与空间精度相关的有待编码要素的重要性。

[0349] 参考图 5b, 与配对相应  $(i_1, j_1)$  的在过程 Proc2 的第 1 次迭代过程中所确定的有待编码的要素  $A(i_1, j_1)$  被认为是与空间精度最小相关的要素。因此, 模块 5 分配了一个最小优先级指数  $Prio1$ 。

[0350] 在分配了优先级指数  $Prio1$  之后, 与配对相应  $(i_2, j_2)$  的在过程 Proc2 的第二次迭代过程中确定的有待编码的要素  $A(i_2, j_2)$  被认为是与空间精度最小相关的要素。因此分配了一个最小优先级指数  $Prio2$ , 其中  $Prio2 > Prio1$ 。因此, 排序模块 6 连续为  $r$  个有待编码的要素中的每个都分配了逐渐增加的优先级指数  $Prio1, Prio2$  到  $Prio r$ 。

[0351] 在过程 Proc2 的重复过程中还没有被分配优先级次序的有待编码的要素比分配了优先级次序的有待编码的要素的有关空间精度的重要性更高。

[0352] 当  $r$  等于  $Q * M - 1$  次时, 有待编码的要素集合被一个接一个地排序。

[0353] 在上述情况下, 可以认为执行的过程 Proc2 的迭代次数  $r$  的数目等于  $Q * M - 1$ 。

[0354] 分配给有待编码的要素  $A(k, j)$  的优先级次序也被分配给已编码的要素, 根据该有待编码要素的量化结果  $\bar{A}(k, j)$ 。与有待编码的要素  $A(k, j)$  相应的已编码要素同样表示为  $\bar{A}(k, j)$ 。

[0355] 用于构造二进制序列的模块 8 构造一个与每个信号  $S_i (i = 1 \text{ 到 } N)$  的帧相应的二进制序列  $Seq$ , 按照所分配的优先级指数逐渐降低, 这些序列连续地整合成为已编码要素  $\bar{A}(k, j)$ , 二进制序列成为待发送的码流  $\Phi$ 。

[0356] 因此, 组成的二进制序列  $Seq$  由模块 6 所执行的排序进行排序。

[0357] 在上述实施例中, 在过程 Proc2 中的每一次迭代时从有待编码的要素  $A(i, j)$  中删除一个频谱成份。

[0358] 在另一个实施例中, 一个重叠量化器被用于量化步骤。在这种情况下, 在过程

Proc2 的迭代过程中被标识的关于空间精度的重要性最低的有待编码要素的频谱成份未被删除,但相对于有待排序的有待编码要素的其它频谱成份,对该成份的编码分配了一个减小的速率。

[0359] 因此,编码器 1 是一个考虑到不同单频信号之间的干扰,而允许速率可适配性的编码器。它使得定义优化感知空间精度的压缩数据成为可能。

[0360] 过程 Proc1 和 Proc2 的联合

[0361] 在一个实施例中,有待编码的要素中重要性最低的要素通过使用了一个联合上述过程 Proc1 和 Proc2 的的方法的过程 Proc 的方法而得到定义,根据一个考虑到总体声音质量和空间相关性的标准。

[0362] 过程 Proc 的初始化包括上述过程 Proc1 和 Proc2 的初始化。

[0363] 现在结合图 11 对这样一个方法 Proc 的第  $n$  次迭代 ( $n > 1$ ) 进行描述,考虑第  $(n+1)$  个编码速率  $D_n$ , 以及一个有待排序的有待编码的要素  $A(k, j)$  集合,其中  $(k, j) \in E_{n-1}$ 。

[0364] 基于使用了 Proc1 和 Proc2 的的方法的方法 Proc 的前一次迭代,该速率以及该有待编码的要素集合在方法 Proc 的前一次迭代过程中被确定。根据所定义的标准,前一次迭代已经允许确定有待编码要素为最不重要的要素。

[0365] 根据所需要的总体声音质量和空间精度,这些所定义的标准已经被建立。

[0366] 平行地,过程 Proc1 的步骤 1d 和 1e 的重复是基于该有待排序的要素集合,确定与总体声音质量最小相关的有待编码的要素  $A(i_{n1}, j_{n1})$ , 并且过程 Proc2 的步骤 2e 和 2i 的重复是基于该有待排序的要素序列,确定与空间精度最小相关的有待编码的要素  $A(i_{n2}, j_{n2})$ 。

[0367] 根据所定义的标准,在步骤 300 中,两个所确定的有待编码要素中的一个或者两个被选择。这个或者每个被选择的有待编码的要素表示为  $A(i_n, j_n)$ 。

[0368] 接着,另一方面,配对  $(i_n, j_n)$  的一个或多个标识符被提供到排序模块 6,作为过程 Proc2 的第  $n$  次迭代的结果,模块 6 根据已经定义的标准对其分配一个优先级  $Prio_n$ 。该分配的优先级  $Prio_n$  比在过程 Proc 的前一次迭代中根据已定义的标准为所选择的有待编码的要素所分配的优先级大。该步骤代替了如前所述的过程 Proc1 的步骤 1f 以及过程 Proc2 的步骤 2k。

[0369] 该被选择的有待编码的一个或多个要素接着被插入到将被发送的二进制序列,在方法 Proc 的前一次迭代过程中被选择的有待编码的要素之前(由于根据已定义的标准,有待编码的要素  $A(i_n, j_n)$  的重要性比被过程 Proc 之前所选择的有待编码的要素的重要性更高)。该被选择的一个或多个有待编码的要素  $A(i_n, j_n)$  被插入到将被发送的二进制序列,在集合  $E_{n-1}$  的其它有待编码的要素(由于根据已定义的标准,该有待编码的要素  $A(i_n, j_n)$  的重要性比这些其它有待编码的要素的重要性更低)。

[0370] 另一方面,在步骤 301 中,被选择用于方法 Proc 的下一迭代(第  $n+1$  次迭代)(包括方法 Proc 1 和 Proc 2 的第  $n+1$  次迭代)的一个或多个有待编码的要素  $A(i_n, j_n)$  被删除,并接着应用于有待编码的要素集合  $E_n = E_{n-1} \setminus A(i_n, j_n)$ , 基于一个如过程 Proc1 的步骤 1g 以及过程 Proc 2 的步骤 2n 中定义的被减小的速率。

[0371] 该步骤 301 代替了如前所述的方法 Proc1 的步骤 1g 以及过程 Proc 2 的步骤 2m。

[0372] 该被定义的标准使得选择分别在方法 Proc 的步骤 300 中所标识出的一个或多个最小相关要素成为可能。

[0373] 例如,在一个实施例中,由过程 Proc1 在每一次迭代  $n$  中所标识出的要素被删除,其中  $n$  为偶数,并且由过程 Proc2 在每一次迭代  $n$  时所标识出的要素被删除,其中  $n$  为奇数,这使得最佳程度地保持总体声音质量以及空间精度成为可能。

[0374] 也可以使用其它标准。执行这样的方法 Proc 的一种编码使得获得一个关于声音质量以及空间精度的可适配速率的码流成为可能。

[0375] 在解码侧所执行的操作

[0376] 解码器 100 包括二进制序列读取模块 104,逆量化模块 105,逆环绕声变换模块 101 以及频率 / 时间变换模块 102。

[0377] 解码器 100 用于在输入端接收编码器 1 发送的码流  $\Phi$ ,并且在输出端输出  $Q'$  个信号  $S'_1, S'_2, \dots, S'_{Q'}$ ,用于提供给声音表现系统 103 的  $Q'$  个对应的喇叭  $H_1, \dots, H_{Q'}$ 。在一个实施例中,喇叭的数目  $Q'$  可以与所发送的环绕声成份的数目  $Q$  不同。

[0378] 作为示例,图 7 示出了一个包括 8 个喇叭  $h_1, h_2, \dots, h_8$  的声音表现系统的配置。

[0379] 二进制序列读取模块 104 从所接收的二进制序列  $\Phi$  中提取数据,该数据指示为要素所确定的  $\bar{A}(k, j)$  ( $k = 1$  到  $Q$  且  $j = 0$  到  $M-1$ ) 的量化指数,并且将它们提供给逆量化模块 105 的输入端。

[0380] 逆量化模块 105 执行逆量化操作。

[0381] 确定要素  $\bar{A}'(k, j)$ ,  $k = 1$  到  $Q$  且  $j = 0$  到  $M-1$  的要素矩阵  $\bar{A}'$ ,以使得:当接收到的序列包括表示要素  $\bar{A}(k, j)$  的量化指数时,  $\bar{A}'(k, j) = \bar{A}(k, j)$ ,其中  $\bar{A}(k, j)$  来自于解码器的环绕声成份的参数编码  $A(k, j)$ ,以及当接收到的序列不包括表示要素  $\bar{A}(k, j)$  的量化指数时(例如,这些数据已经在码流服务器侧被截断,以便适应网络的可适配速率和 / 或终端的特性),  $\bar{A}'(k, j) = 0$ 。

[0382] 逆空间变换模块 101 适于确定矩阵  $X'$  的元素  $X'(i, j)$ ,  $i = 1$  到  $Q'$ ,  $j = 0$  到  $M-1$ ,该矩阵  $X'$  定义了  $Q'$  个信号中的每一个  $S'_i$  的  $M$  个频谱系数,  $i = 1$  到  $Q'$ ,  $j = 0$  到  $M-1$ ,基于环绕声成份  $A'(k, j)$ ,其中  $k = 1$  到  $Q$  且  $j = 0$  到  $M-1$ ,被逆量化模块 105 所确定。

[0383]  $AmbInv(p', Q')$  是对于 3D 场景的  $p'$  阶逆环绕声变换矩阵,用于确定  $Q'$  个信号  $S'_i$ ,  $i = 1$  到  $Q'$ ,用于与解码器 100 相关的声音表现系统的  $Q'$  个喇叭,基于接收到的  $Q$  个环绕声成份。角度  $\beta_i$ ,其中  $i = 1$  到  $Q'$ ,表示来自喇叭  $H_i$  的声音传播角度。在图 7 的示例中,这些角度对应于由喇叭所发出的声音的传播轴线与轴  $XX$  之间的角度。

[0384]  $X'$  是信号  $S'_i$  ( $i = 1$  到  $Q'$ ) 关于频带  $F_j$  ( $j = 0$  到  $M-1$ ) 的频谱成份  $X'(i, j)$  的矩阵。这样:

$$[0385] \quad \bar{A}' = \begin{bmatrix} \bar{A}'(1,0) & \bar{A}'(1,1) & \dots & \bar{A}'(1,M-1) \\ \bar{A}'(2,0) & & & \bar{A}'(2,M-1) \\ \vdots & & & \vdots \\ \bar{A}'(Q,0) & \bar{A}'(Q,1) & \dots & \bar{A}'(Q,M-1) \end{bmatrix},$$

$$[0386] \quad \underline{\text{AmbInv}}(p', Q') = \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} \cdot \cos \beta_1 & \frac{1}{\sqrt{2}} \cdot \sin \beta_1 & \dots & \frac{1}{\sqrt{2}} \cdot \sin p' \beta_1 \\ 1 & \frac{1}{\sqrt{2}} \cdot \cos \beta_2 & \dots & \dots & \frac{1}{\sqrt{2}} \cdot \sin p' \beta_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{1}{\sqrt{2}} \cdot \cos \beta_{Q'} & \dots & \dots & \frac{1}{\sqrt{2}} \cdot \sin p' \beta_{Q'} \end{bmatrix} \text{ 并且}$$

$$[0387] \quad \underline{X'} = \begin{bmatrix} X'(1,0) & X'(1,1) & \dots & \dots & X'(1,M-1) \\ X'(2,0) & \dots & \dots & \dots & X'(2,M-1) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X'(Q',0) & \dots & \dots & \dots & X'(Q',M-1) \end{bmatrix}$$

[0388] 并且我们有

$$[0389] \quad \underline{X'} = \underline{\text{AmbInv}}(p', Q') \times \bar{A'} \quad \text{等式 (7)}。$$

[0390] 逆空间变换模块 101 用于根据等式 (7) 确定矩阵  $\underline{X'}$  的元素, 频谱系数  $X'(i, j)$ ,  $i = 1$  到  $Q'$ ,  $j = 0$  到  $M-1$ 。

[0391] 这些元素  $X'(i, j)$ ,  $i = 1$  到  $Q'$ ,  $j = 0$  到  $M-1$ , 一旦确定, 就被发送到频率 / 时间变换模块 102 的输入端。

[0392] 解码器 100 的频率 / 时间变换模块 102 基于接收到的频谱系数  $X'(i, j)$ ,  $i = 1$  到  $Q'$ ,  $j = 0$  到  $M-1$ , 将频率空间表示 (the space of frequency representation) 变换为时间空间表示 (the space of time representation) (该变换在当前情况下为逆 MDCT 变换), 并且, 它可以因此确定  $Q'$  个信号  $S'_1, \dots, S'_{Q'}$  中的每一个的时间帧。

[0393] 每一个信号  $S'_i$ ,  $i = 1$  到  $Q'$ , 用于声音表现系统 103 的喇叭  $H_i$ 。

[0394] 在一个实施例中, 由解码器执行的操作中的至少一些通过解码器处理装置的计算机程序指令执行。

[0395] 对于来自前述信号  $S_1, \dots, S_N$  的环绕声变换的成份的编码的一个优点是, 如果声音场景的信号数目  $N$  很大, 它们可以通过一个比  $N$  小很多的  $Q$  个环绕声成份来表示, 同时很小程度地降低信号的空间质量。需要发送的数据量因此而被减小, 同时没有明显降低声音场景的声音质量。

[0396] 根据本发明的编码的另一个优点在于, 这样的编码可以适配于不同类型的声音表现系统, 无论该声音表现系统提供的是何种数量、排列以及类型的喇叭。

[0397] 事实上, 接收包括  $Q$  个环绕声成份的二进制序列的解码器, 后续可以执行任何阶次  $p'$  的逆环绕声变换的操作, 并且与将使用被解码信号的声音表现系统的喇叭数目  $Q'$  相应。

[0398] 编码器 1 执行的编码使得, 利用第一过程 Proc1 根据有待编码的要素对声音质量的贡献, 和 / 或利用第二过程 Proc2 根据有待编码的要素对空间精度以及在声音场景中包含的传播方向的正确再现的贡献, 对有待编码的要素进行排序成为可能。

[0399] 为了与所使用的速率限制相适应, 完全可以截断在序列中排列的具有较低优先级的要素的序列。然后, 可以保证提供最佳总体声音质量 (当执行过程 Proc1 时) 和 / 或

最佳空间精度（当执行过程 Proc2 时）。事实上，对于要素的排序已经以这样的方式执行，即那些对于总体声音质量和 / 或空间精度贡献最小的要素被置于序列的末端。

[0400] 为了确定一个二进制序列，过程 Proc1 和 Proc 2 可以根据各实施例彼此单独或者组合实施。

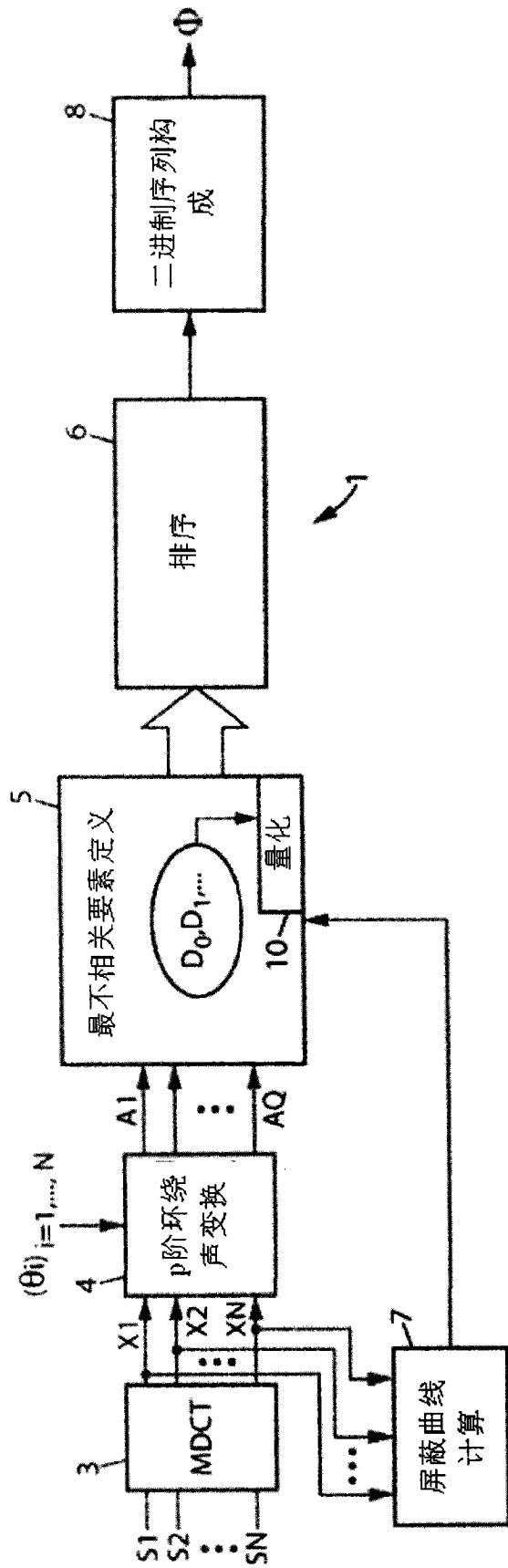


FIG. 1

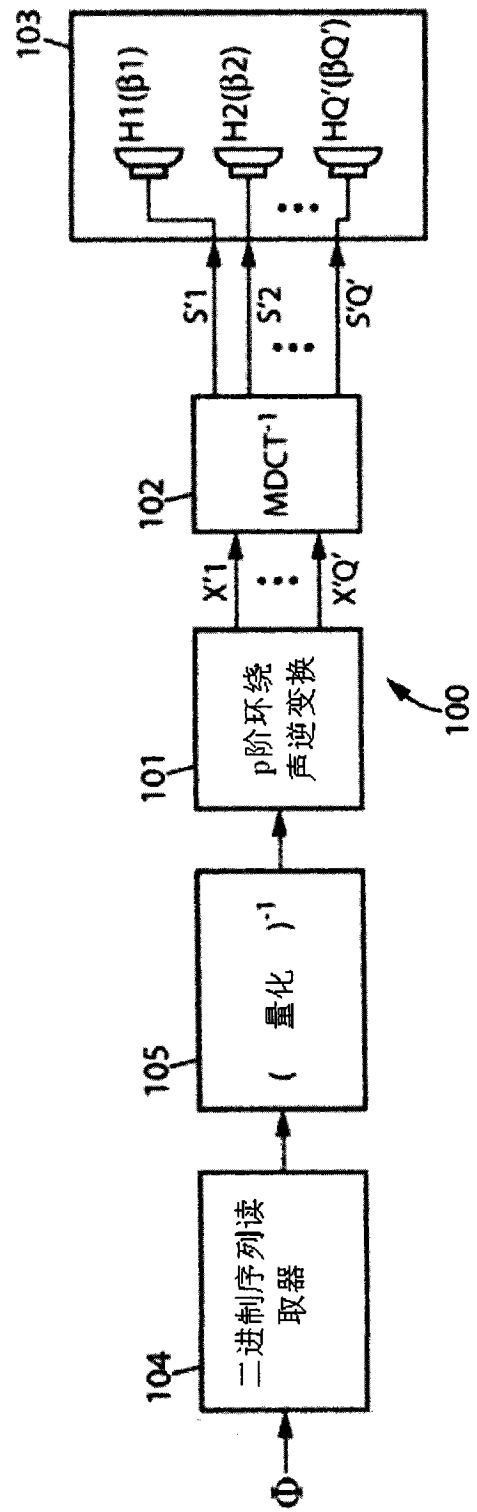


FIG. 2

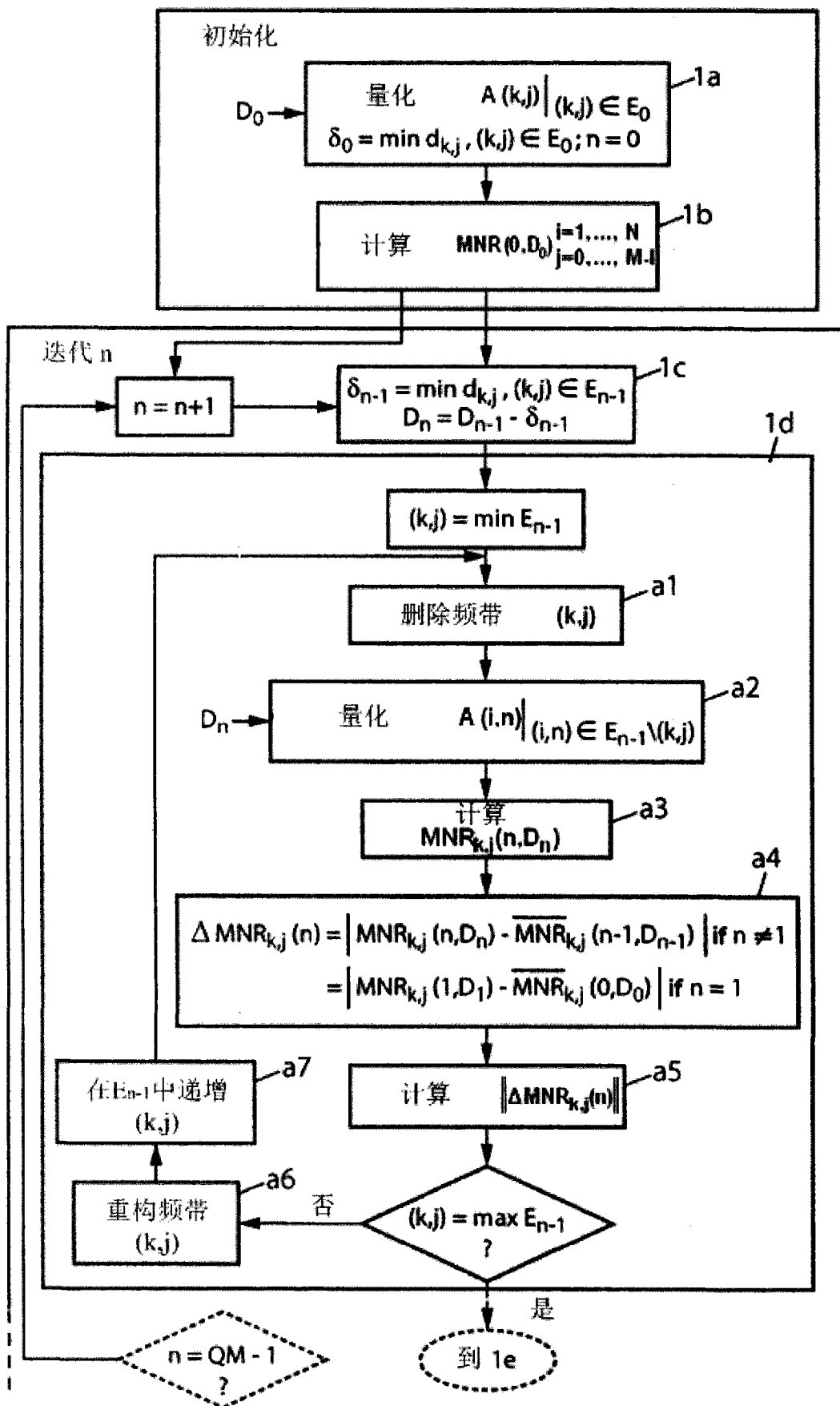


FIG. 4a

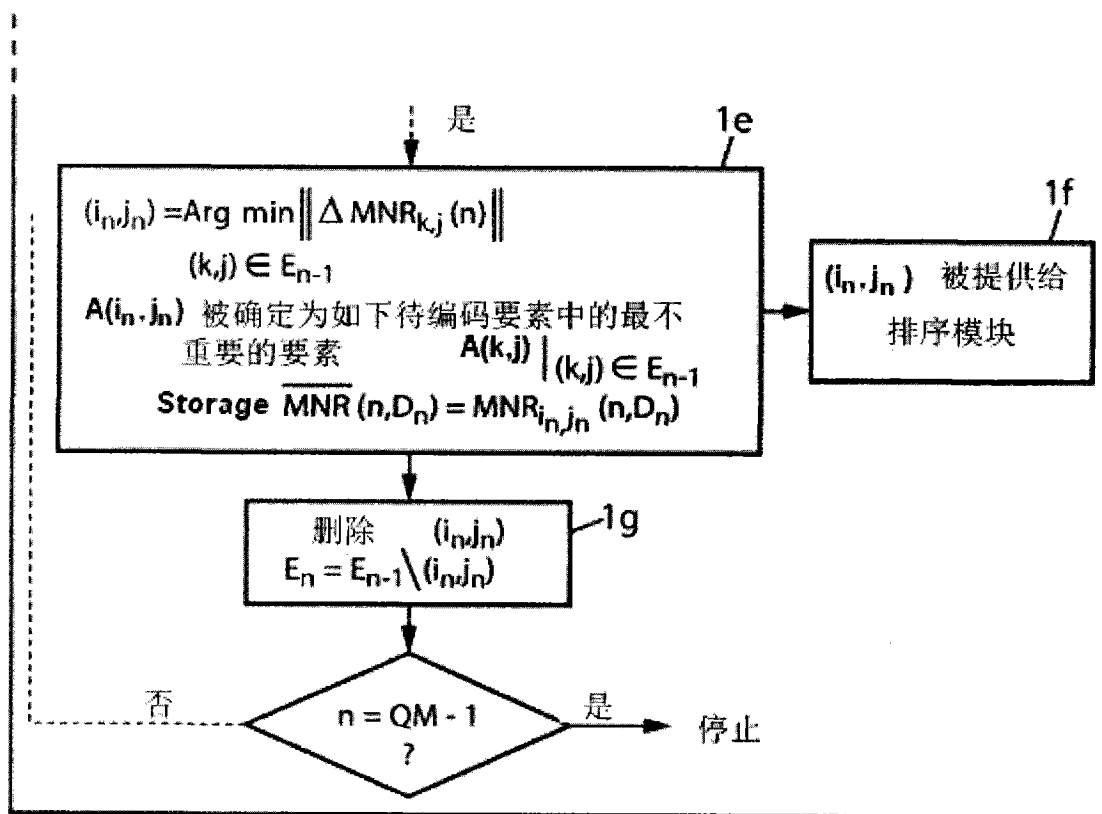


FIG. 4b

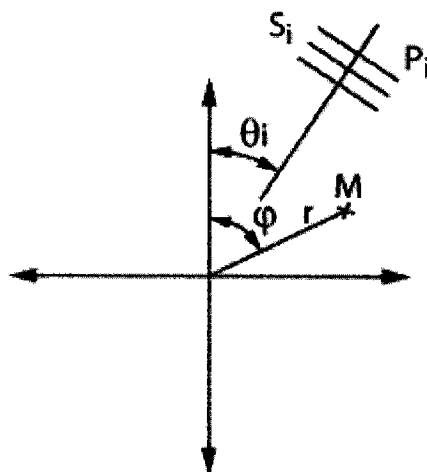


FIG. 3

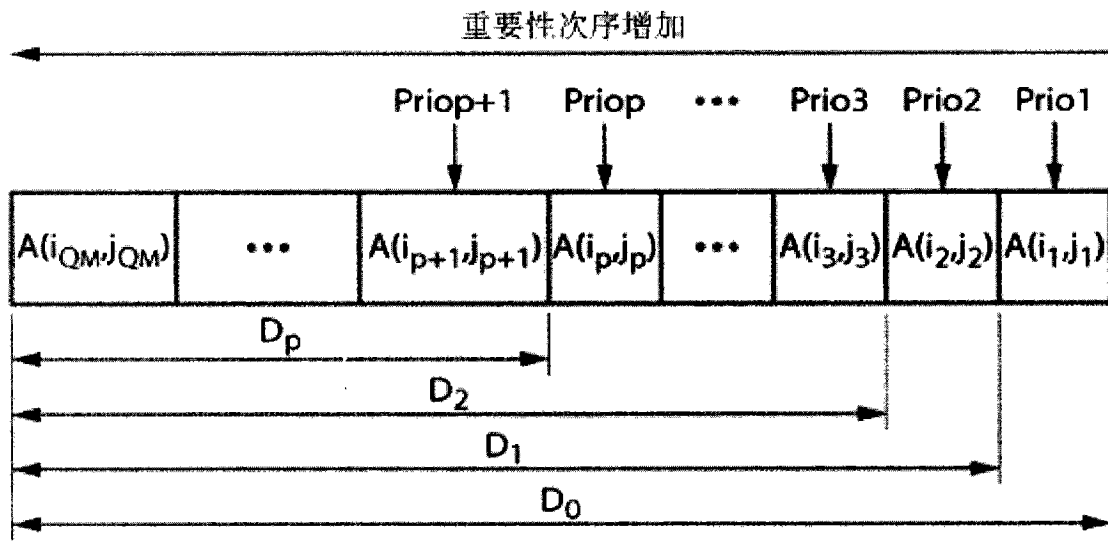


FIG. 5a

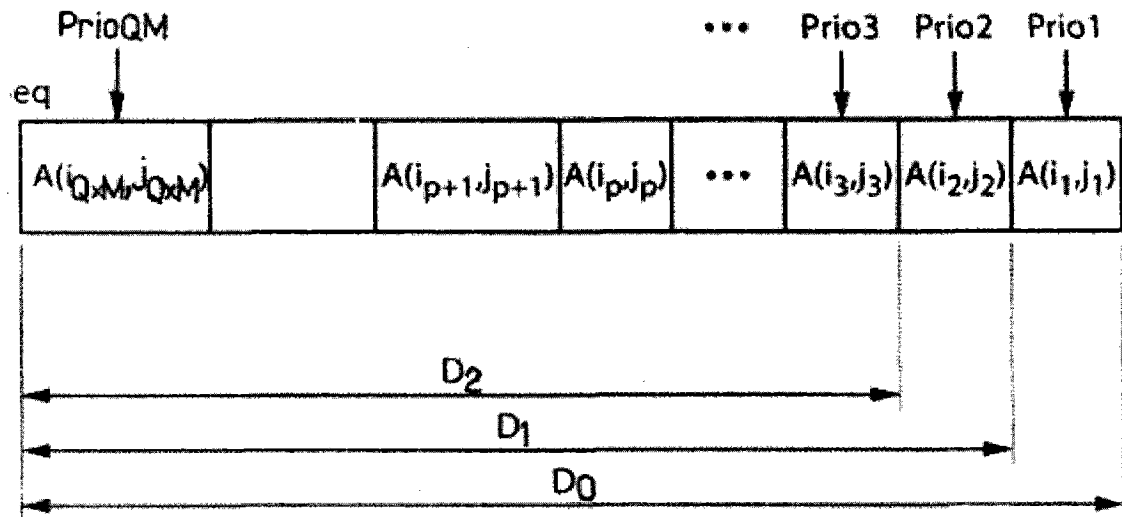


FIG. 5b

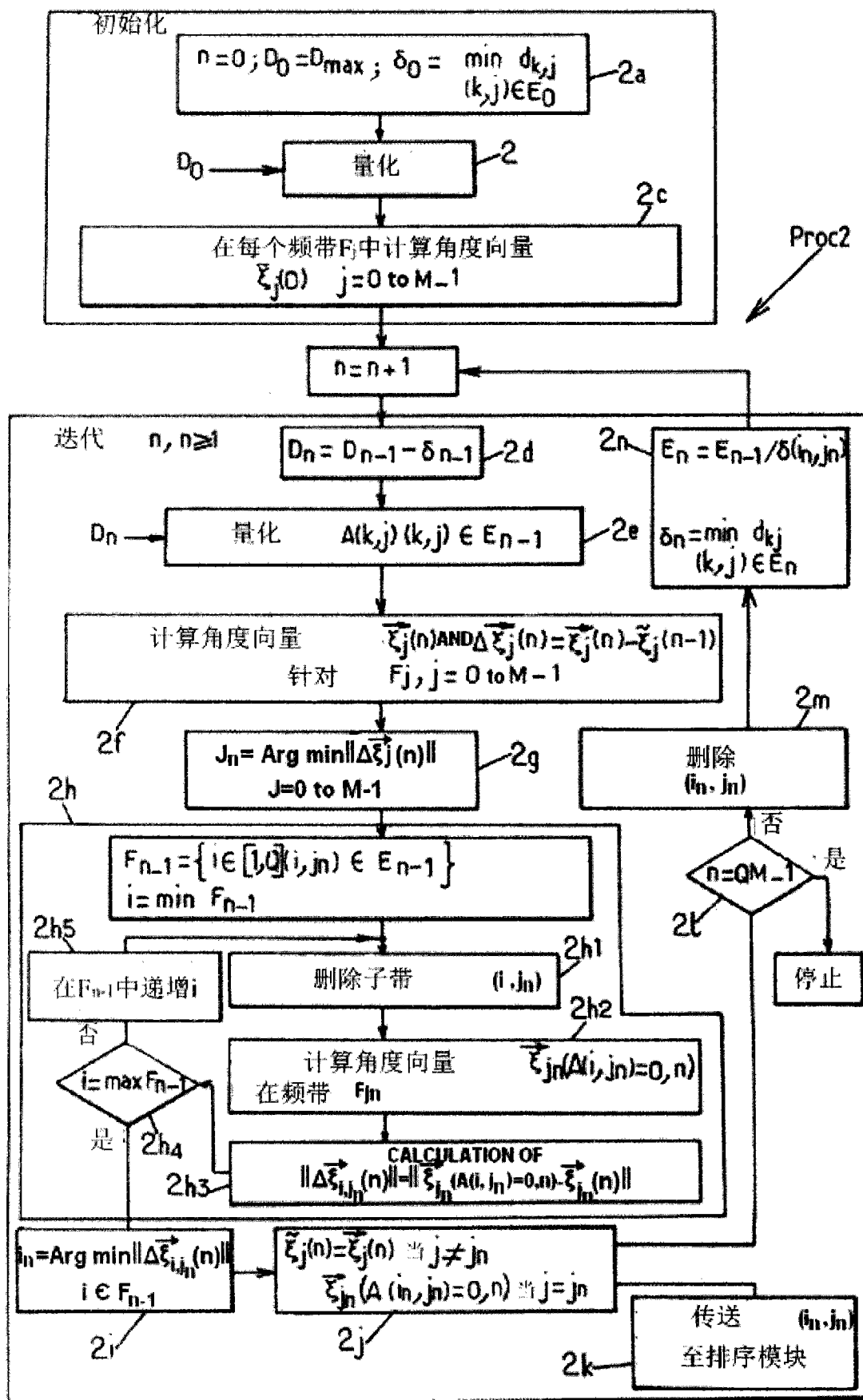


FIG. 6

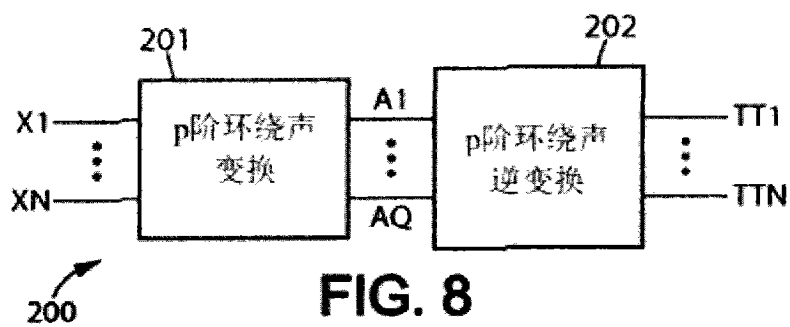


FIG. 8

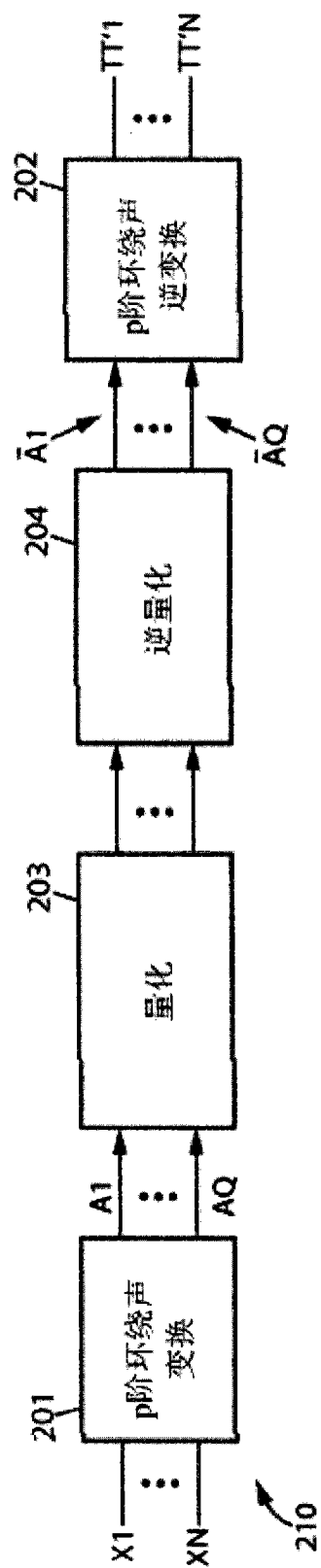


FIG. 9

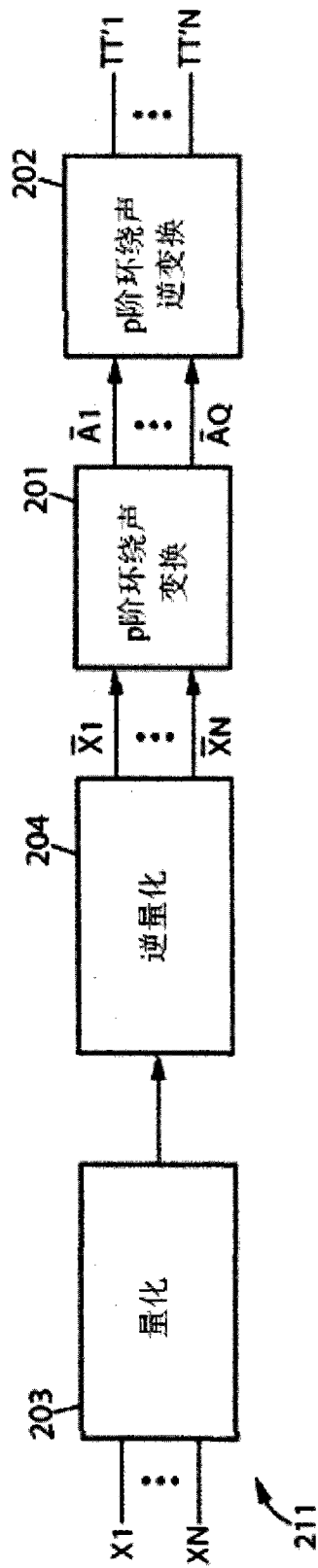


FIG. 10

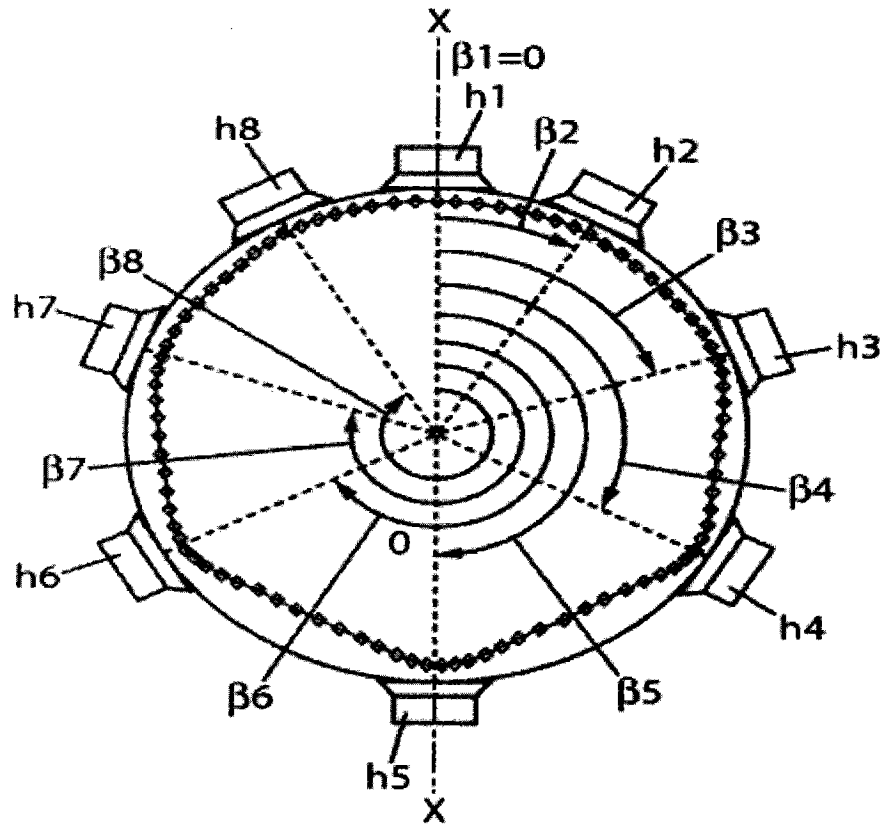


FIG. 7

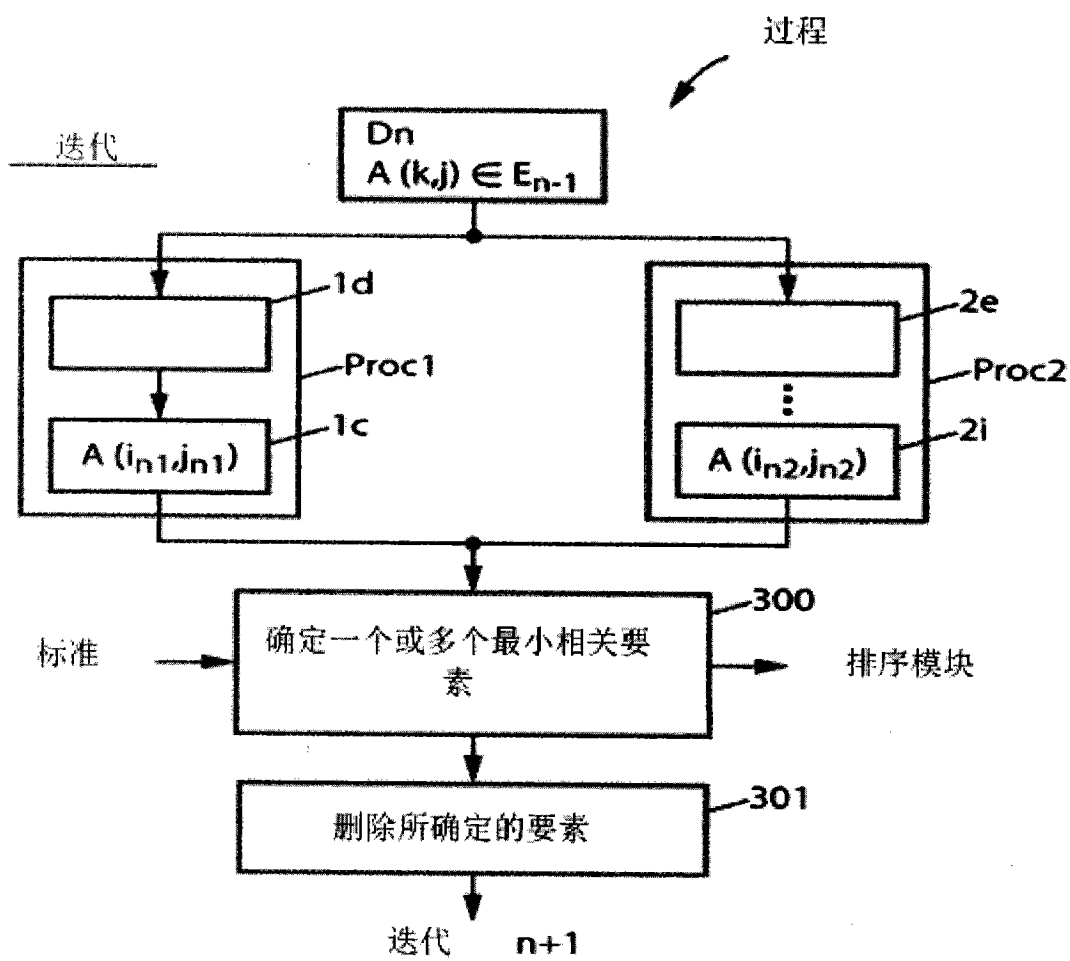


FIG. 11