



(12) 发明专利

(10) 授权公告号 CN 112997248 B

(45) 授权公告日 2024. 11. 01

(21) 申请号 201980072488.X

(22) 申请日 2019.10.01

(65) 同一申请的已公布的文献号
申请公布号 CN 112997248 A

(43) 申请公布日 2021.06.18

(30) 优先权数据
1817807.9 2018.10.31 GB
1903850.4 2019.03.21 GB

(85) PCT国际申请进入国家阶段日
2021.04.30

(86) PCT国际申请的申请数据
PCT/FI2019/050704 2019.10.01

(87) PCT国际申请的公布数据
W02020/089510 EN 2020.05.07

(73) 专利权人 诺基亚技术有限公司
地址 芬兰埃斯波

(72) 发明人 A·瓦西拉凯 M·V·莱蒂南

(74) 专利代理机构 北京市中咨律师事务所
11247
专利代理师 杨晓光

(51) Int.Cl.
G10L 19/008 (2013.01)
H04S 3/00 (2006.01)
G10L 19/032 (2013.01)
G10L 19/02 (2013.01)

(56) 对比文件
CN 112219237 A, 2021.01.12
CN 113228168 A, 2021.08.06
GB 201817807 D0, 2018.12.19

审查员 宋玉香

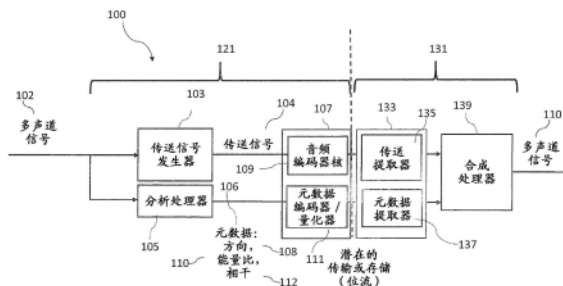
权利要求书5页 说明书23页 附图10页

(54) 发明名称

确定空间音频参数的编码和相关联解码

(57) 摘要

一种装置,其包括用于以下操作的模块:接收用于音频信号的帧的子带的值,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值和至少一个扩展和/或环绕相干值;基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本;对至少一个向量进行离散余弦变换,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码。



1. 一种装置,其包括用于以下操作的模块:

接收用于音频信号的帧的子带的值,所述值包括用于每个子带的至少一个方位角值、至少一个能量比值和至少一个扩展和/或环绕相干值;

基于用于帧的每个子带的所述至少一个能量比值和所述至少一个方位角值的方差,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本;

对至少一个向量进行离散余弦变换,所述至少一个向量包括用于所述帧的子带的所述至少一个扩展和/或环绕相干值;以及

基于所确定的码本对所述离散余弦变换向量的第一数量的分量进行编码。

2. 根据权利要求1所述的装置,其中,用于基于用于帧的每个子带的所述至少一个能量比值和所述至少一个方位角值的方差,确定用于对用于每个子带的至少一个相干值进行编码的码本的模块进一步用于:

获得表示用于所述帧的每个子带的所述至少一个能量比值的加权平均的指数;

确定用于帧的所述子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值;以及

基于所述指数和确定用于帧的所述子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值,来选择所述码本。

3. 根据权利要求2所述的装置,其中,用于基于所述指数和确定用于帧的子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值,来选择所述码本的模块进一步用于基于所述指数选择用于所述码本的码字的数量。

4. 根据权利要求1至3中任一项所述的装置,其中,用于基于所述所确定的码本对所述离散余弦变换向量的第一数量的分量进行编码的模块进一步用于:

确定所述离散余弦变换向量的所述第一数量的分量取决于所述子带;

基于所述码本对所述第一数量的所述离散余弦变换向量分量的第一分量进行编码。

5. 根据权利要求4所述的装置,其中,用于基于所述所确定的码本对所述离散余弦变换向量的第一数量的分量进行编码的模块进一步用于:

基于子带的指数确定用于标量量化的码本,每个码本包括确定数量的码字;

基于所述所确定的码本,为所述第一数量的所述离散余弦变换向量分量的分量剩余部分生成至少一个进一步指数;

基于用于所述第一数量的所述离散余弦变换向量分量的所述分量剩余部分的所述至少一个进一步指数,生成均值移除指数;以及

对所述均值移除指数进行熵编码。

6. 根据权利要求4所述的装置,其中,用于基于所述所确定的码本对所述离散余弦变换向量的第一数量的分量进行编码的模块进一步用于:

基于具有所定义数量的码字的码本,确定用于所述第一数量的所述离散余弦变换向量分量的所述分量剩余部分的至少一个进一步指数,所述码本是基于所述向量的子带指数的;

基于用于所述第一数量的所述离散余弦变换向量分量的所述分量剩余部分的所述至少一个进一步指数,确定均值移除指数;以及

对所述均值移除指数进行熵编码。

7. 根据权利要求5和6中任一项所述的装置,其中,用于对所述均值移除指数进行熵编码的模块进一步用于对所述均值移除指数进行Golomb-Rice编码。

8. 根据权利要求1至7中任一项所述的装置,其中,所述模块进一步用于:存储和/或发送所述离散余弦变换向量的所述编码的第一数量的分量。

9. 根据权利要求1至8中任一项所述的装置,其中所述模块进一步用于对所述至少一个能量比值进行标量量化,以生成适合于确定所述码本的至少一个能量比值指数,所述码本用于对用于每个子带的至少一个相干值进行编码。

10. 根据从属于权利要求5或6时的权利要求9所述的装置,其中所述模块进一步用于:接收用于每个子带的至少一个仰角值;

基于目标位数、在所述编码前基于所述所确定的码本估计的用于对所述离散余弦变换向量的所述第一数量的分量进行编码的位数、表示所述至少一个能量比值指数的位数、和表示所述均值移除指数的所述熵编码的位数,来估计用于对所述至少一个方位角值和所述至少一个仰角值进行编码的剩余位数;

对所述至少一个方位角值和所述至少一个仰角值进行编码,以基于所述剩余位数生成至少一个方位角值指数和至少一个仰角值指数,其中基于所述至少一个方位角值指数确定用于对用于每个子带的至少一个相干值进行编码的所述码本。

11. 一种装置,其包括用于以下操作的模块:

获得用于音频信号的帧的子带的编码值,所述值包括用于每个子带的至少一个方位角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;

解码所述至少一个方位角指数,以给出至少一个方位角值;

基于所述至少一个能量比指数和所述至少一个方位角值的方差,确定用于对用于每个子带的所述至少一个扩展和/或环绕相干指数进行解码的码本;

对所述至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量,所述至少一个向量包括用于所述帧的子带的所述至少一个扩展和/或环绕相干值;以及

解析所述向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

12. 根据权利要求11所述的装置,其中,用于基于所述至少一个能量比指数和所述至少一个方位角值的方差,确定用于对用于每个子带的所述至少一个扩展和/或环绕相干指数进行解码的码本的所述模块进一步用于:

确定用于帧的子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值;以及

基于所述至少一个能量比指数和确定用于帧的所述子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值,来选择所述码本。

13. 根据权利要求12所述的装置,其中,用于基于所述至少一个能量比指数,和确定用于帧的子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值来选择所述码本的所述模块进一步用于基于所述至少一个能量比指数选择用于所述码本的码字的数量。

14. 根据权利要求11至13中任一项所述的装置,其中,所述装置进一步包括用于通过以下来基于所确定的码本对所述离散余弦变换向量的第一数量的分量进行解码的模块:

基于所述码本对所述第一数量的所述离散余弦变换向量分量的第一分量进行解码;

基于所述码本对所述第一数量的所述离散余弦变换向量分量的进一步分量进行解码;

以及

对所述解码的第一分量和进一步分量进行逆余弦变换。

15. 一种方法,其包括:

接收用于音频信号的帧的子带的值,所述值包括用于每个子带的至少一个方位角值、至少一个能量比值和至少一个扩展和/或环绕相干值;

基于用于帧的每个子带的所述至少一个能量比值和所述至少一个方位角值的方差,确定用于对用于每个子带的所述至少一个扩展和/或环绕相干值进行编码的码本;

对至少一个向量进行离散余弦变换,所述至少一个向量包括用于所述帧的子带的所述至少一个扩展和/或环绕相干值;以及

基于所确定的码本对所述离散余弦变换向量的第一数量的分量进行编码。

16. 根据权利要求15所述的方法,其中,基于用于帧的每个子带的所述至少一个能量比值和所述至少一个方位角值的方差,确定用于对用于每个子带的至少一个相干值进行编码的码本进一步包括:

获得表示用于所述帧的每个子带的所述至少一个能量比值的加权平均的指数;

确定用于帧的所述子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值;以及

基于所述指数和确定用于所述帧的所述子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值,来选择所述码本。

17. 根据权利要求16所述的方法,其中,基于所述指数和所述确定来选择所述码本进一步包括基于所述指数选择用于所述码本的码字的数量。

18. 根据权利要求15至17中任一项所述的方法,其中,基于所述所确定的码本对所述离散余弦变换向量的第一数量的分量进行编码进一步包括:

确定所述离散余弦变换向量的所述第一数量的分量取决于所述子带;

基于所述码本对所述第一数量的所述离散余弦变换向量分量的第一分量进行编码。

19. 根据权利要求18所述的方法,其中,基于所述所确定的码本对所述离散余弦变换向量的第一数量的分量进行编码进一步包括:

基于子带的指数确定用于标量量化的码本,每个码本包括确定数量的码字;

基于所述所确定的码本,为所述第一数量的所述离散余弦变换向量分量的分量剩余部分生成至少一个进一步指数;

基于用于所述第一数量的所述离散余弦变换向量分量的所述分量剩余部分的至少一个进一步指数,生成均值移除指数;以及

对所述均值移除指数进行熵编码。

20. 根据权利要求19所述的方法,其中基于所述所确定的码本对所述离散余弦变换向量的第一数量的分量进行编码进一步包括:

基于具有所定义数量的码字的码本,确定用于所述第一数量的所述离散余弦变换向量分量的所述分量剩余部分的至少一个进一步指数,所述码本是基于所述向量的子带指数的;

基于用于所述第一数量的所述离散余弦变换向量分量的所述分量剩余部分的所述至少一个进一步指数,确定均值移除指数;以及

对所述均值移除指数进行熵编码。

21. 根据权利要求19和20中任一项所述的方法,其中,对所述均值移除指数进行熵编码进一步包括对所述均值移除指数进行Golomb-Rice编码。

22. 根据权利要求15至21中任一项所述的方法,进一步包括:存储和/或发送所述离散余弦变换向量的所述编码的第一数量的分量。

23. 根据权利要求15至22中任一项所述的方法,进一步包括:对所述至少一个能量比值进行标量量化,以生成适合于确定所述码本的至少一个能量比值指数,所述码本用于对用于每个子带的至少一个相干值进行编码。

24. 根据从属于权利要求19或20时的权利要求23所述的方法,进一步包括:

接收用于每个子带的至少一个仰角值;

基于目标位数、在所述编码前基于所述所确定的码本估计的用于对所述离散余弦变换向量的所述第一数量的分量进行编码的位数、表示所述至少一个能量比值指数的位数、和表示所述均值移除指数的所述熵编码的位数,来估计用于对所述至少一个方位角值和所述至少一个仰角值进行编码的剩余位数;

对所述至少一个方位角值和所述至少一个仰角值进行编码,以基于所述剩余位数生成至少一个方位角值指数和至少一个仰角值指数,其中基于所述至少一个方位角值指数确定用于对用于每个子带的至少一个相干值进行编码的所述码本。

25. 一种方法,其包括:

获得用于音频信号的帧的子带的编码值,所述值包括用于每个子带的至少一个方位角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;

解码所述至少一个方位角指数,以给出至少一个方位角值;

基于所述至少一个能量比指数和所述至少一个方位角值的方差,确定用于对用于每个子带的所述至少一个扩展和/或环绕相干指数进行解码的码本;

对所述至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量,所述至少一个向量包括用于所述帧的子带的所述至少一个扩展和/或环绕相干值;以及

解析所述向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

26. 根据权利要求25所述的方法,其中,基于所述至少一个能量比指数和所述至少一个方位角值的方差,确定用于对用于每个子带的所述至少一个扩展和/或环绕相干指数进行解码的码本进一步包括:

确定用于帧的子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值;以及

基于所述至少一个能量比指数和确定用于帧的所述子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值来选择所述码本。

27. 根据权利要求26所述的方法,其中,基于所述至少一个能量比指数,和确定用于帧的子带的所述至少一个方位角值的方差是否大于或等于所确定的阈值来选择所述码本进一步包括:基于所述至少一个能量比指数选择用于所述码本的码字的数量。

28. 根据权利要求25至27中任一项所述的方法,其中,基于所确定的码本对所述离散余弦变换向量的第一数量的分量进行解码进一步包括:

基于所述码本对所述第一数量的所述离散余弦变换向量分量的第一分量进行解码;

基于所述码本对所述第一数量的所述离散余弦变换向量分量的进一步分量进行解码；
以及
对所述解码的第一分量和进一步分量进行逆余弦变换。

确定空间音频参数的编码和相关联解码

技术领域

[0001] 本申请涉及用于声场相关参数编码的装置和方法,但不排他地针对用于音频编码器和解码器的时频域方向相关参数编码。

背景技术

[0002] 参数化空间音频处理是其中使用一组参数描述声音的空间方面的音频信号处理的领域。例如,在来自麦克风阵列的参数化空间音频采集中,典型且有效的选择是从麦克风阵列信号估计出一组参数,诸如声音在频带中的方向,以及所采集的声音在频带中的方向性部分和非方向性部分之间的比率。已知这些参数良好地描述所采集的声音在麦克风阵列的位置处的感知空间性质。这些参数可以因此在空间声音的合成中用于双耳耳机、用于扬声器,或用于其他格式,诸如全景声(Ambisonic)。

[0003] 因此,频带中的方向和直接能量与总能量之比是对空间音频采集特别有效的参数化。

[0004] 由频带中的方向参数和频带中的能量比参数(指示声音的方向性)组成的参数集也可以用作作用于音频编解码器的空间元数据(其也可以包括其他参数,诸如相干性、扩展相干性、方向数量、距离等)。例如,这些参数可以从麦克风阵列捕获的音频信号估计,并且例如可以从麦克风阵列信号生成立体声信号,与空间元数据一起传送。立体声信号可以例如用AAC编码器进行编码。解码器可以将音频信号解码成PCM信号,并对声音进行频带处理(使用空间元数据)以获得空间输出,例如双耳输出。

[0005] 上述解决方案特别适合于对来自麦克风阵列(例如,在移动电话、VR相机、独立麦克风阵列中)的所捕获的空间声音进行编码。然而,对于此种编码器,可能期望还具有除麦克风阵列捕获的信号之外的其他输入类型,例如,扬声器信号、音频对象信号或全景声信号。

[0006] 在与定向音频编码(DirAC)和谐波平面波扩展(Harpex)相关的科学文献中,已彻底记录了分析用于空间元数据提取的一阶全景声(FOA)输入。这是因为存在直接提供FOA信号(更准确地说:其变体,B格式信号)的麦克风阵列,并且分析此种输入因此已是该领域的研究重点。

[0007] 用于编码器的另一个输入也是多声道扬声器输入,例如5.1或7.1声道环绕声输入。

[0008] 然而关于元数据分量,压缩是目前的研究课题。

发明内容

[0009] 根据第一方面,提供了一种装置,其包括用于以下操作的模块:接收用于音频信号的帧的子带的值,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值和至少一个扩展和/或环绕相干值;基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编

码的码本;对至少一个向量进行离散余弦变换,该至少一个向量包括用于所述帧的子带的至少一个扩展和/或环绕相干值;以及基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码。

[0010] 用于基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个相干值进行编码的码本的模块可以进一步用于:获得表示用于帧的每个子带的至少一个能量比值的加权平均的指数;确定用于帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值;以及基于指数和确定用于帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值,来选择码本。

[0011] 用于基于指数,和确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值来选择码本的模块可以进一步用于基于指数选择用于码本的码字的数量。

[0012] 分布的度量可以是以下中的一个:连续方位角值之间的平均绝对差;相对于子带中的平均方位角值的平均绝对差;用于帧的子带的至少一个方位角值的标准偏差;以及用于帧的子带的至少一个方位角值的方差。

[0013] 用于基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码的模块可以进一步用于:确定离散余弦变换向量的第一数量取决于子带;基于码本对第一数量的离散余弦变换向量分量中的第一分量进行编码。

[0014] 用于基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码的模块可以进一步用于:基于子带的指数确定用于标量量化的码本,每个码本包括确定数量的码字;基于所确定的码本,为第一数量的离散余弦变换向量分量的分量剩余部分生成至少一个进一步指数;基于用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数,生成均值移除指数;以及对均值移除指数进行熵编码。

[0015] 用于基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码的模块可以进一步用于:基于具有所定义数量的码字的码本,确定用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数,该码本进一步基于向量的子带指数;基于用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数,确定均值移除指数;以及对均值移除指数进行熵编码。

[0016] 用于对均值移除指数进行熵编码的模块可以进一步用于对均值移除指数进行Golomb-Rice编码。

[0017] 该模块可以进一步用于:存储和/或发送离散余弦变换向量的编码的第一数量的分量。

[0018] 该模块可以进一步用于对至少一个能量比值进行标量量化,以生成适合于确定码本的至少一个能量比值指数,该码本用于对用于每个子带的至少一个相干值进行编码。

[0019] 该模块可以进一步用于:基于目标位数、在编码前基于所确定的码本估计的用于对离散余弦变换向量的第一数量的分量进行编码的位数、表示至少一个能量比值指数的位数、和表示均值移除指数的熵编码的位数,来估计用于对至少一个方位角值和至少一个仰角值进行编码的剩余位数;对至少一个方位角值和至少一个仰角值进行编码,以基于剩余位数生成至少一个方位角值指数和至少一个仰角值指数,其中基于至少一个方位角值指数确定用于对用于每个子带的至少一个相干值进行编码的码本。

[0020] 根据第二方面,提供了一种装置,其包括用于以下动作的模块:获得用于音频信号的帧的子带的编码值,该值包括用于每个子带的至少一个方位角指数、至少一个仰角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本;对至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及解析该向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

[0021] 用于基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本的模块可以进一步用于:确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值;以及基于至少一个能量比指数和确定用于帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值,来选择码本。

[0022] 用于基于至少一个能量比指数,和确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值来选择码本的模块可以进一步用于基于至少一个能量比指数选择用于所述码本的码字的数量。

[0023] 分布的度量可以是以下中的一个:连续方位角值之间的平均绝对差;相对于子带中的平均方位角值的平均绝对差;用于帧的子带的至少一个方位角值的方差;以及用于帧的子带的至少一个方位角值的方差。

[0024] 用于基于所确定的码本对离散余弦变换向量的第一数量的分量进行解码的模块可以进一步用于:基于码本对第一数量的离散余弦变换向量分量的第一分量进行解码;基于码本对第一数量的离散余弦变换向量分量的进一步分量进行解码;以及对解码的第一分量和进一步分量进行逆余弦变换。

[0025] 根据第三方面,提供了一种方法,其包括:接收用于音频信号的帧的子带的值,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值和至少一个扩展和/或环绕相干值;基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本;对至少一个向量进行离散余弦变换,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码。

[0026] 基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个相干值进行编码的码本可以进一步包括:获得表示用于帧的每个子带的至少一个能量比值的加权平均值的指数;确定用于帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值;以及基于指数和确定用于上述帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值,来选择码本。

[0027] 基于指数,和确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值来选择码本可以进一步包括基于指数选择用于所述码本的码字的数量。

[0028] 分布的度量可以是以下中的一个:连续方位角值之间的平均绝对差;相对于子带中的平均方位角值的平均绝对差;用于帧的子带的至少一个方位角值的标准偏差;以及用于帧的子带的至少一个方位角值的方差。

[0029] 基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码可以进一步

包括:确定离散余弦变换向量的第一数量取决于子带;基于码本对第一数量的离散余弦变换向量分量的第一分量进行编码。

[0030] 基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码可以进一步包括:基于子带的指数确定用于标量量化的码本,每个码本包括确定数量的码字;基于所确定的码本为第一数量的离散余弦变换向量分量的分量剩余部分生成至少一个进一步指数;基于用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数生成均值移除指数;以及对均值移除指数进行熵编码。

[0031] 基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码可以进一步包括:基于具有所定义数量的码字的码本,确定用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数,该码本进一步基于向量的子带指数;基于用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数,确定均值移除指数;以及对均值移除指数进行熵编码。

[0032] 对均值移除指数进行熵编码可以进一步包括对均值移除指数进行Golomb-Rice编码。

[0033] 该方法可以进一步包括:存储和/或发送离散余弦变换向量的编码的第一数量的分量。

[0034] 该方法可以进一步包括:对至少一个能量比值进行标量量化,以生成适合于确定码本的至少一个能量比值指数,该码本用于对用于每个子带的至少一个相干值进行编码。

[0035] 该方法可以进一步包括:基于目标位数、在所述编码前基于所确定的码本估计的用于对离散余弦变换向量的第一数量的分量进行编码的位数、表示至少一个能量比值指数的位数、和表示均值移除指数的熵编码的位数,来估计用于对至少一个方位角值和至少一个仰角值进行编码的剩余位数;对至少一个方位角值和至少一个仰角值进行编码,以基于剩余位数生成至少一个方位角值指数和至少一个仰角值指数,其中基于至少一个方位角值指数确定用于对用于每个子带的至少一个相干值进行编码的码本。

[0036] 根据第四方面,提供了一种方法,其包括:获得用于音频信号的帧的子带的编码值,该值包括用于每个子带的至少一个方位角指数、至少一个仰角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本;对至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量,该至少一个向量包括用于所述帧的子带的至少一个扩展和/或环绕相干值;以及解析该向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

[0037] 基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本可以进一步包括:确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值;以及基于至少一个能量比指数和确定用于帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值,来选择码本。

[0038] 基于至少一个能量比指数,和确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值来选择码本可以进一步包括:基于至少一个能量比指数选择用于所述码本的码字的数量。

[0039] 分布的度量可以是以下中的一个:连续方位角值之间的平均绝对差;相对于子带中的平均方位角值的平均绝对差;用于帧的子带的至少一个方位角值的方差;以及用于帧的子带的至少一个方位角值的方差。

[0040] 基于所确定的码本对离散余弦变换向量的第一数量的分量进行解码可以进一步包括:基于码本对第一数量的离散余弦变换向量分量的第一分量进行解码;基于码本对第一数量的离散余弦变换向量分量的进一步分量进行解码;以及对解码的第一分量和进一步分量进行逆余弦变换。

[0041] 根据第五方面,提供了一种装置,其包括至少一个处理器和至少一个包括计算机程序代码的存储器,该至少一个存储器和计算机程序代码被配置为与至少一个处理器一起,致使该装置至少:接收用于音频信号的帧的子带的值,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值和至少一个扩展和/或环绕相干值;基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本;对至少一个向量进行离散余弦变换,该至少一个向量包括用于所述帧的子带的至少一个扩展和/或环绕相干值;以及基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码。

[0042] 被致使基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个相干值进行编码的码本的装置可以进一步被致使:获得表示用于帧的每个子带的至少一个能量比值的加权平均的指数;确定用于帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值;以及基于指数和确定用于帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值,来选择码本。

[0043] 被致使基于指数,和确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值来选择码本的装置可以进一步被致使:基于指数选择用于所述码本的码字的数量。

[0044] 所述分布的度量可以是以下中的一个:连续方位角值之间的平均绝对差;相对于子带中的平均方位角值的平均绝对差;用于帧的子带的至少一个方位角值的标准偏差;以及用于帧的子带的至少一个方位角值的方差。

[0045] 被致使基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码的装置可以进一步被致使:确定离散余弦变换向量的第一数量取决于子带;基于码本对第一数量的离散余弦变换向量分量的第一分量进行编码。

[0046] 被致使基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码的装置可以进一步被致使:基于子带的指数确定用于标量量化的码本,每个码本包括确定数量的码字;基于所确定的码本,为第一数量的离散余弦变换向量分量的分量剩余部分生成至少一个进一步指数;基于用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数,生成均值移除指数;以及对均值移除指数进行熵编码。

[0047] 被致使基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码的装置可以进一步被致使:基于具有所定义数量的码字的码本,确定用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数,该码本进一步基于向量的子带指数;基于用于第一数量的离散余弦变换向量分量的分量剩余部分的至少一个进一步指数,确定均值移除指数;以及对均值移除指数进行熵编码。

[0048] 被致使对均值移除指数进行熵编码的装置可以进一步被致使对均值移除指数进行Golomb-Rice编码。

[0049] 该装置可以进一步被致使:存储和/或发送离散余弦变换向量的编码的第一数量的分量。

[0050] 该装置可以进一步被致使:对至少一个能量比值进行标量量化,以生成适合于确定码本的至少一个能量比值指数,该码本用于对用于每个子带的至少一个相干值进行编码。

[0051] 该装置可以进一步被致使:基于目标位数、在编码前基于所确定的码本估计用于对离散余弦变换向量的第一数量的分量进行编码的位数、表示至少一个能量比值指数的位数、和表示均值移除指数的熵编码的位数,来估计用于对至少一个方位角值和至少一个仰角值进行编码的剩余位数;对至少一个方位角值和至少一个仰角值进行编码,以基于剩余位数生成至少一个方位角值指数和至少一个仰角值指数,其中基于至少一个方位角值指数确定用于对用于每个子带的至少一个相干值进行编码的码本。

[0052] 根据第六方面,提供了一种装置,其包括至少一个处理器和至少一个包括计算机程序代码的存储器,该至少一个存储器和计算机程序代码被配置为与至少一个处理器一起,致使该装置至少:获得用于音频信号的帧的子带的编码值,该值包括用于每个子带的至少一个方位角指数、至少一个仰角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本;对至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及解析该向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

[0053] 被致使基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本的装置可以进一步被致使:确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值;以及基于至少一个能量比指数和确定用于帧的子带的至少一个方位角值的分布的度量是否大于或等于所确定的阈值,来选择码本。

[0054] 被致使基于至少一个能量比指数,和确定用于帧的子带的至少一个方位角指数的分布的度量是否大于或等于所确定的阈值来选择码本的装置可以进一步被致使:基于至少一个能量比指数选择用于所述码本的码字的数量。

[0055] 所述分布的度量可以是以下中的一个:连续方位角值之间的平均绝对差;相对于子带中的平均方位角值的平均绝对差;用于帧的子带的至少一个方位角值的方差;以及用于帧的子带的至少一个方位角值的方差。

[0056] 被致使基于所确定的码本对离散余弦变换向量的第一数量的分量进行解码的装置可以进一步被致使:基于码本对第一数量的离散余弦变换向量分量的第一分量进行解码;基于码本对第一数量的离散余弦变换向量分量的进一步分量进行解码;以及对解码的第一分量和进一步分量进行逆余弦变换。

[0057] 根据第七方面,提供了一种装置,其包括:用于接收用于音频信号的帧的子带的值的模块,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值

和至少一个扩展和/或环绕相干值;用于基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本的模块;用于对至少一个向量进行离散余弦变换的模块,该至少一个向量包括用于所述帧的子带的至少一个扩展和/或环绕相干值;以及用于基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码的模块。

[0058] 根据第八方面,提供了一种装置,其包括:用于获得用于音频信号的帧的子带的编码值的模块,该值包括用于每个子带的至少一个方位角指数、至少一个仰角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;用于基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本的模块;用于对至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量的模块,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及用于解析该向量以生成用于每个子带的至少一个扩展和/或环绕相干值的模块。

[0059] 根据第九方面,提供了一种包括指令的计算机程序[或包括程序指令的计算机可读介质],其用于致使装置至少执行以下操作:接收用于音频信号的帧的子带的值,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值和至少一个扩展和/或环绕相干值;基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本;对至少一个向量进行离散余弦变换,该至少一个向量包括用于所述帧的子带的至少一个扩展和/或环绕相干值;以及基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码。

[0060] 根据第十方面,提供了一种包括指令的计算机程序[或包括程序指令的计算机可读介质],其用于致使装置至少执行以下操作:获得用于音频信号的帧的子带的编码值,该值包括用于每个子带的至少一个方位角指数、至少一个仰角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本;对至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及解析该向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

[0061] 根据第十一方面,提供了一种非暂时性计算机可读介质,其包括用于引起装置至少执行以下操作的程序指令:接收用于音频信号的帧的子带的值,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值和至少一个扩展和/或环绕相干值;基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本;对至少一个向量进行离散余弦变换,该至少一个向量包括用于所述帧的子带的至少一个扩展和/或环绕相干值;以及基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码。

[0062] 根据第十二方面,提供了一种非暂时性计算机可读介质,其包括用于致使装置至少执行以下操作的程序指令:获得用于音频信号的帧的子带的编码值,该值包括用于每个子带的至少一个方位角指数、至少一个仰角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本;对至少一个扩展和/或环绕相干

指数进行逆离散余弦变换以生成至少一个向量,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及解析该向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

[0063] 根据第十三方面,提供了一种装置,其包括:接收电路系统,其被配置为接收用于音频信号的帧的子带的值,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值和至少一个扩展和/或环绕相干值;确定电路系统,其被配置为基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本;变换电路系统,其被配置为对至少一个向量进行离散余弦变换,该至少一个向量包括用于所述帧的子带的至少一个扩展和/或环绕相干值;以及编码电路系统,其被配置为基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码。

[0064] 根据第十四方面,提供了一种装置,其包括:获得电路系统,其被配置为获得用于音频信号的帧的子带的编码值,该值包括用于每个子带的至少一个方位角指数、至少一个仰角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;确定电路系统,其被配置为基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本;变换电路系统,其被配置为对至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及解析电路系统,其被配置为解析该向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

[0065] 根据第十五方面,提供了一种计算机可读介质,其包括用于致使装置至少执行以下操作的程序指令:接收用于音频信号的帧的子带的值,该值包括用于每个子带的至少一个方位角值、至少一个仰角值、至少一个能量比值和至少一个扩展和/或环绕相干值;基于用于帧的每个子带的至少一个能量比值和至少一个方位角值,确定用于对用于每个子带的至少一个扩展和/或环绕相干值进行编码的码本;对至少一个向量进行离散余弦变换,该至少一个向量包括用于所述帧的子带的至少一个扩展和/或环绕相干值;以及基于所确定的码本对离散余弦变换向量的第一数量的分量进行编码。

[0066] 根据第十六方面,提供了一种计算机可读介质,其包括用于致使装置至少执行以下操作的程序指令:获得用于音频信号的帧的子带的编码值,该值包括用于每个子带的至少一个方位角指数、至少一个仰角指数、至少一个能量比指数和至少一个扩展和/或环绕相干指数;基于至少一个能量比指数和至少一个方位角指数,确定用于对用于每个子带的至少一个扩展和/或环绕相干指数进行解码的码本;对至少一个扩展和/或环绕相干指数进行逆离散余弦变换以生成至少一个向量,该至少一个向量包括用于帧的子带的至少一个扩展和/或环绕相干值;以及解析该向量以生成用于每个子带的至少一个扩展和/或环绕相干值。

[0067] 一种装置,其包括用于执行如上所述的方法的动作用的模块。

[0068] 一种装置,其被配置为执行如上所述的方法的动作。

[0069] 一种计算机程序,其包括用于致使计算机执行如上所述的方法的程序指令。

[0070] 一种存储在介质上的计算机程序产品,可以致使装置执行如本文中所述的方法。

[0071] 一种电子设备,可以包括如本文中所述的装置。

- [0072] 一种芯片组,可以包括如本文中所述的装置。
- [0073] 本申请的实施例旨在解决与现有技术相关联的问题。

附图说明

- [0074] 为了更好地理解本申请,现在将以举例的方式参考附图,在附图中:
- [0075] 图1示意性地示出了适合于实现一些实施例的装置的系统;
- [0076] 图2示意性地示出了根据一些实施例的元数据编码器;
- [0077] 图3示出了根据一些实施例的如图2所示的元数据编码器的操作的流程图;
- [0078] 图4示意性地示出了根据一些实施例的如图2所示的相干编码器;
- [0079] 图5示出了根据一些实施例的如图4所示的相干编码器的操作的流程图;
- [0080] 图6示出了根据一些实施例的相干编码器对第一相干分量和进一步相干分量进行编码的操作的流程图;
- [0081] 图7示出了根据一些进一步实施例的相干编码器对第一相干分量和进一步相干分量进行编码的进一步操作的流程图;
- [0082] 图8示意性地示出了根据一些实施例的关于相干解码的元数据解码器;
- [0083] 图9示出了根据一些实施例的如图8所示的元数据解码器的操作的流程图;以及
- [0084] 图10示意性地示出了适合于实现所示装置的示例设备。

具体实施方式

[0085] 下面进一步详细描述用于提供有效的空间分析导出的元数据参数的合适设备和可能的机制。在下面的讨论中,就多声道麦克风的实现方式来讨论多声道系统。然而如上面所讨论的,输入格式可以是任何合适的输入格式,诸如多声道扬声器、全景声(FOA/HOA)等。应理解,在一些实施例中,声道位置基于麦克风的位置,或者是虚拟位置或方向。此外,示例系统的输出是多声道扬声器布置。然而应理解,输出可以经由扬声器以外的手段表现给用户。此外,多声道扬声器信号可以概括为两个或更多个回放音频信号。

[0086] 对于每个所考虑的时频块(时间/频率子带),元数据至少由方向(仰角、方位角)、所得方向的能量比,以及所得方向的扩展相干分量组成。此外,独立于方向,可以针对每个时频块确定并包括环绕相干性。所有该数据由编码器编码和发送(或存储),以便能够在解码器处重建空间信号。

[0087] 编解码器的典型总体操作位率留有3.0kbps、4.0kbps、8kbps或10kbps用于发送/存储元数据。前面已研究了方向参数和能量比分量的编码,但对相干数据的编码还没有探讨,并且在较低的位率下,相干数据被去除并且不被发送或存储。

[0088] 如下文中讨论的概念是将相干参数与用于每个时频块的方向参数和能量比参数一起进行编码。在下面的示例中,编码在离散余弦变换域中执行,并且取决于当前的子带指数,以及当前的能量比和方位角值。在下面的实施例中选择了DCT变换,因为其是针对低复杂度实现方式而被优化的,然而可以代替地应用和使用其他时频域变换。

[0089] 在一些实施例中,固定位率编码方法可以与可变位率编码组合,该可变位率编码在不同的段之间分配要压缩的数据的编码位,使得每帧的总体位率是固定的。在时频块内,位可以在频率子带之间转移。

[0090] 关于图1示出了用于实现本申请的实施例的示例装置和系统。系统100被示出具有“分析”部分121和“合成”部分131。“分析”部分121是从接收多声道扬声器信号到对元数据和混缩信号进行编码的部分,并且“合成”部分131是从对编码的元数据和混缩信号进行解码到重新生成的信号的呈现(例如以多声道扬声器形式)的部分。

[0091] 系统100和“分析”部分121的输入是多声道信号102。在下面的示例中描述了麦克风声道信号输入,然而任何合适的输入(或合成的多声道)格式可以在其他实施例中实现。例如在一些实施例中,空间分析器和空间分析可以在编码器外部实现。例如在一些实施例中,与音频信号相关联的空间元数据可以作为单独的位流提供给编码器。在一些实施例中,空间元数据可以作为一组空间(方向)指数值提供。

[0092] 多声道信号被传递给传送信号发生器103和分析处理器105。

[0093] 在一些实施例中,传送信号发生器103被配置为接收多声道信号,并生成包括确定的声道数的合适传送信号,并输出传送信号104。例如,传送信号发生器103可以被配置为生成多声道信号的2音频声道混缩。确定的声道数可以是任何合适的声道数。传送信号发生器在一些实施例中被配置为以其他方式选择或组合,例如通过波束成形技术将输入的音频信号选择或组合到确定的声道数,并将这些信号作为传送信号输出。

[0094] 在一些实施例中,传送信号发生器103是任选的,并且多声道信号以与该示例中传送信号相同的方式未经处理地传递给编码器107。

[0095] 在一些实施例中,分析处理器105还被配置为接收多声道信号,并分析信号以产生与多声道信号相关联并因此与传送信号104相关联的元数据106。分析处理器105可以被配置为生成元数据,对于每个时频分析区间,该元数据可以包括方向参数108和能量比参数110以及相干参数112(并且在一些实施例中是扩散性参数)。方向参数、能量比参数和相干参数在一些实施例中可以被认为是空间音频参数。换句话说,空间音频参数包括旨在表征由多声道信号(或通常两个或更多个回放音频信号)创建的声场的参数。

[0096] 在一些实施例中,所生成的参数可能因频带不同而不同。因此,例如在频带X中,所有的参数都被生成和发送,而在频带Y中,仅参数中的一个被生成和发送,并且此外在频带Z中没有参数被生成或发送。这点的实际示例可能是,对于一些频带,诸如最高频带,出于感知的原因,参数中的一些是不需要的。传送信号104和元数据106可以被传递给编码器107。

[0097] 编码器107可以包括音频编码器核109,其被配置为接收传送(例如混缩)信号104并生成这些音频信号的合适编码。编码器107在一些实施例中可以是计算机(运行存储在存储器上和至少一个处理器上的合适软件),或者可替代地是利用例如FPGA或ASIC的特定设备。编码可以使用任何合适的方案来实现。编码器107此外可以包括元数据编码器/量化器111,其被配置为接收元数据并输出信息的编码或压缩形式。在一些实施例中,编码器107可以在图1中虚线所示的传输或存储之前,进一步地将元数据交织、多路复用到单个数据流,或者将元数据嵌入在编码的混缩信号内。多路复用可以使用任何合适的方案来实现。

[0098] 在解码器侧中,所接收或所检索的数据(流)可由解码器/解多路复用器133接收。解码器/解多路复用器133可将编码流解多路复用,并将音频编码流传递给传送提取器135,其被配置为对音频信号进行解码以获得传送信号。类似地,解码器/解多路复用器133可以包括元数据提取器137,其被配置为接收编码的元数据并生成元数据。解码器/解多路复用器133在一些实施例中可以是计算机(运行存储在存储器上和至少一个处理器上的合适软

件),或者可替代地是利用例如FPGA或ASIC的特定设备。

[0099] 解码的元数据和传送音频信号可以被传递该合成处理器139。

[0100] 系统100“合成”部分131进一步示出了合成处理器139,其被配置为接收传送和元数据,并基于传送信号和元数据以任何合适的格式重新创建采用多声道信号110的形式(这些可以是多声道扬声器格式,或在一些实施例中可以是任何合适的输出格式,诸如双耳或全景声信号,取决于使用情况)的合成的空间音频。

[0101] 因此,总之首先系统(分析部分)被配置为接收多声道音频信号。

[0102] 然后,系统(分析部分)被配置为生成合适的传送音频信号(例如通过选择或混缩音频信号声道中的一些)。

[0103] 然后,系统被配置为对传送信号和元数据进行编码以用于存储/传输。

[0104] 此后,系统可以存储/发送编码的传送和元数据。

[0105] 系统可以检索/接收编码的传送和元数据。

[0106] 然后,系统被配置为从编码的传送和元数据参数提取传送和元数据,例如对编码的传送和元数据参数进行解多路复用和解码。

[0107] 系统(合成部分)被配置为基于提取的传送音频信号和元数据合成输出多声道音频信号。

[0108] 关于图2,进一步详细描述根据一些实施例的示例分析处理器105和元数据编码器/量化器111(如图1所示)。

[0109] 在一些实施例中,分析处理器105包括时频域变换器201。

[0110] 在一些实施例中,时频域变换器201被配置为接收多声道信号102并应用合适的时频域变换,诸如短时傅里叶变换(STFT),以便将输入的时域信号转换为合适的时频信号。这些时频信号可以被传递给空间分析器203和信号分析器205。

[0111] 因此例如时频信号202可以在时频域表示中由下式表示

[0112] $s_i(b, n)$,

[0113] 其中 b 为频率仓(frequency bin)指数,并且 n 为时频块(帧)指数,并且 i 为信道指数。在另一种表述中, n 可以认为是采样率比原时域信号的采样率更低的时间指数。这些频率仓可以被分组到子带中,该子带将仓中的一个或多个分组到带指数 $k=0, \dots, K-1$ 的子带中。每个子带 k 具有最低仓 $b_{k,low}$ 和最高仓 $b_{k,high}$,并且子带包含从 $b_{k,low}$ 到 $b_{k,high}$ 的所有仓。子带的宽度可以近似于任何合适的分布。例如等效矩形带宽(ERB)尺度或Bark尺度。

[0114] 在一些实施例中,分析处理器105包括空间分析器203。空间分析器203可以被配置为接收时频信号202,并基于这些信号估计方向参数108。方向参数可以基于任何基于音频的“方向”而被确定。

[0115] 例如在一些实施例中,空间分析器203被配置为用两个或更多个信号输入来估计方向。这表示估计“方向”的最简单配置,更复杂的处理可以用甚至更多的信号来执行。

[0116] 因此,空间分析器203可以被配置为在音频信号帧内为每个频带和时间的时频块提供至少一个方位角和仰角,表示为方位角 $\varphi(k, n)$ 和仰角 $\theta(k, n)$ 。方向参数108也可以被传递给方向指数发生器205。

[0117] 空间分析器203还可以被配置为确定能量比参数110。能量比可以被认为是对这样的音频信号的能量确定,该这样的音频信号可以被认为从一方向到达的。可以估计直接

能量与总能量之比 $r(k, n)$, 例如, 使用方向估计的稳定性度量, 或使用任何相关性度量, 或任何其他合适的方法来获得比率参数。能量比可以被传递给能量比编码器207。

[0118] 空间分析器203此外可以被配置为确定多个相干参数112, 其可以包括环绕相干($\gamma(k, n)$)和扩展相干($\zeta(k, n)$), 两者都在时频域中分析。扩展相干参数的值可以为从0到1。扩展相干值为0表示为点声源, 换句话说, 当使用多扬声器系统重现音频信号时, 应该用尽可能少的扬声器重现声音(例如当方向为中心时, 仅用中心扬声器)。随着扩展相干值的增加, 更多的能量被扩展到围绕中心扬声器的其他扬声器, 直到值为0.5时, 能量在中心扬声器和邻近的扬声器之间均匀扩展。当扩展相干的值增加到高于0.5时, 中心扬声器中的能量减少, 直到值为1时, 中心扬声器中没有能量, 并且所有的能量都在邻近的扬声器中。环绕相干参数的值为从0到1。值为1意味着所有(或几乎所有)扬声器声道之间都有相干。值为0意味着所有(或甚至几乎所有)扬声器声道之间没有相干。这在GB申请号1718341.9和PCT申请PCT/FI2018/050788中进一步解释。

[0119] 因此总之, 分析处理器被配置为接收时域多声道或其他格式诸如麦克风或全景声音频信号。

[0120] 此后, 分析处理器可以应用时域到频域变换(例如STFT)以生成合适的时频域信号用于分析, 并然后应用方向分析以确定方向和能量比参数。

[0121] 然后, 分析处理器可以被配置为输出所确定的参数。

[0122] 尽管方向、能量比和相干参数在这里针对每个时间指数 n 表达, 但在一些实施例中, 参数可以在若干个时间指数上组合。同样的适用于频率轴线, 正如已表达的那样, 若干个频率仓 b 的方向可以通过由若干个频率仓 b 组成的频带 k 中的一个方向参数表达。同样的适用于本文中讨论的所有空间参数。

[0123] 在一些实施例中, 方向数据可以使用16位来表示, 使得每个方位角参数大约在9位上表示, 并且仰角在7位上表示。在此类实施例中, 能量比参数可以在8位上表示。对于每个帧, 可以有 $N=5$ 个子带和 $M=4$ 个时频(TF)块。因此在该示例中, 需要有 $(16+8) \times M \times N$ 位以存储用于每个帧的未压缩方向和能量比元数据。用于每个TF块的相干数据可以是0和1之间的浮点表示, 并且可以最初在8位上表示。

[0124] 同样如图2所示, 示出了根据一些实施例的示例元数据编码器/量化器111。

[0125] 元数据编码器/量化器111可以包括方向编码器205。方向编码器205被配置为接收方向参数(诸如方位角 $\varphi(k, n)$ 和仰角 $\theta(k, n)$) 108(以及在一些实施例中的预期位分配), 并由此生成合适的编码输出。在一些实施例中, 编码基于形成球面网格的球面布置, 该球面网格以“表面”球面上的环布置, 该环由所确定的量化分辨率定义的查找表定义。换句话说, 球面网格使用这样的想法: 用较小球面覆盖球面, 并将较小球面的中心视为定义几乎等距离方向的网格的点。因此, 较小球面定义关于中心点的圆锥或实心角, 该中心点可根据任何合适的索引算法进行索引。尽管这里描述了球面量化, 但线性或非线性的任何合适的量化都可以使用。

[0126] 此外, 在一些实施例中, 方向编码器205被配置为确定方位角参数值的方差, 并将该方差传递给相干编码器209。

[0127] 然后, 编码的方向参数可以被传递给组合器211。

[0128] 元数据编码器/量化器111可以包括能量比编码器207。能量比编码器207被配置为

接收能量比并确定用于压缩用于子带和时频块的能量比的合适编码。例如在一些实施例中,能量比编码器207被配置为使用3位来对每个能量比参数值进行编码。

[0129] 此外,在一些实施例中,不是发送或存储用于所有TF块的所有能量比值,而是仅发送或存储每个子带的一个加权平均值。该平均值可以通过考虑每个时间块的总能量来确定,因此有利于具有更多能量的子带的值。

[0130] 在此类实施例中,量化的能量比值对于给定子带的所有TF块是相同的。

[0131] 在一些实施例中,能量比编码器207进一步被配置为将量化的(编码的)能量比值传递给组合器211和相干编码器209。

[0132] 元数据编码器/量化器111可以包括相干编码器209。相干编码器209被配置为接收相干值,并确定用于对用于子带和时频块的相干值进行压缩的合适编码。用于相干参数值的3位精度值已示出产生可接受的音频合成结果,但即使如此,这也需要为所有TF块(在示例中每帧8个子带和5个TF块)的相干数据提供总共 3×20 位。

[0133] 如下文中所述,在一些实施例中,编码在DCT域中实现,并且可以取决于当前的子带指数,以及当前的能量比和方位角值。

[0134] 然后,编码的相干参数值可以传递给组合器211。

[0135] 元数据编码器/量化器111可以包括组合器211。该组合器被配置为接收编码的(或量化/压缩的)方向参数、能量比参数和相干参数,并将这些组合以生成合适的输出(例如,元数据位流,其可与传送信号组合,或与传送信号分开发送或存储)。

[0136] 关于图3,示出了根据一些实施例的如图2所示的元数据编码器/量化器的示例操作。

[0137] 初始操作是获得元数据(诸如方位角值、仰角值、能量比、相干等),如图3中通过步骤301所示。

[0138] 然后,方向值(仰角、方位角)可以被压缩或编码(例如通过应用球面量化,或任何合适的压缩),如图3中通过步骤303所示。

[0139] 能量比值被压缩或编码(例如通过生成每子带的加权平均值,并然后将这些量化为3位值),如图3中通过步骤305所示。

[0140] 相干值也被压缩或编码(例如通过在DCT域中编码,如下文中所指示),如图3中通过步骤307所示。

[0141] 然后将编码的方向值、能量比、相干值组合以生成编码的元数据,如图3中通过步骤305所示。

[0142] 关于图4示出了如图2所示的示例相干编码器209。

[0143] 在一些实施例中,相干编码器209包括相干向量发生器401。相干向量发生器401被配置为接收相干值112,其可以是0和1之间的8位浮点表示。

[0144] 相干向量发生器401被配置为针对每个子带生成相干值的向量。因此,在其中有M个时频块的示例中,则相干向量发生器401被配置为生成M维的相干数据向量402。

[0145] 相干数据向量402被输出到离散余弦变换器403。

[0146] 在一些实施例中,相干编码器209包括离散余弦变换器。离散余弦变换器可以被配置为接收M维相干数据向量402并对该向量进行离散余弦变换(DCT)。

[0147] 可以实现用于执行DCT的任何合适的方法。例如在一些实施例中,其中向量包括对

应于子带的4维相干向量。于是向量 $x = (x_1, x_2, x_3, x_4)'$ 与4阶DCT矩阵的矩阵相乘等于：

$$[0148] \quad y = DCT(x) = \begin{bmatrix} 0.5(a+b) \\ 0.6533c + 0.2706d \\ 0.5(a-b) \\ 0.2706c - 0.6533d \end{bmatrix}$$

[0149] 其中

[0150] $a = x_1 + x_2$

[0151] $b = x_2 + x_3$

[0152] $c = x_1 - x_4$

[0153] $d = x_2 - x_3$

[0154] 这将DCT变换的操作次数从28次减少到14次。

[0155] 然后，DCT相干向量404可以被输出到向量编码器405。

[0156] 在一些实施例中，相干编码器209包括向量编码器405。向量编码器405被配置为接收DCT相干向量404并通过使用合适的码本对其进行编码。

[0157] 在一些实施例中，向量编码器405包括码本确定器415。该码本确定器被配置为接收编码/量化的能量比412和量化的方位角414的方差（其可以从能量比编码器和方向编码器确定，如图2所示），并确定应用于DCT相干向量值的合适码本。

[0158] 在一些实施例中，第一DCT参数的编码以不同于进一步DCT参数的编码的方式实现。这是因为第一DCT参数和进一步DCT参数具有显著不同的分布。此外，第一DCT参数的分布还取决于两个因素：用于当前子带的能量比值和当前子带内方位角的方差。

[0159] 在一些实施例中（并且如先前所讨论的），3位用于对每个能量比值进行编码，并且每个子带仅生成和发送（和/或存储）一个加权平均值。这意味着量化的能量比值对于给定子带的所有TF块都是相同的。

[0160] 此外，方位角的方差基于子带内的量化方位角的方差是非常小（在所确定的阈值下）还是大于阈值而影响第一DCT参数的分布。

[0161] 在一些实施例中，此外，子带的数量选择为 I_N 。例如在一些实施例中 $I_N = 3$ 。在此类实施例中，对直到所选择的子带限制的子带使用第一数量的次级DCT参数进行编码，并且对剩余的子带使用第二数量的次级DCT参数进行编码。在一些实施例中，第一数量是1并且第二数量是2。换句话说，在一些实施例中，向量编码器被配置为使得子带 $\leq I_N$ 对DCT变换向量的前2个分量（一个初级和一个次级）进行编码，并且子带 $> I_N$ 对DCT变换向量的前3个分量（一个初级和两个次级）进行编码。这两个额外分量可以用2维向量量化器进行编码，或者它们可以作为额外维度添加到第二DCT参数的N维向量量化器，并立即使用N+2维向量量化器用于对所有的次级参数进行编码。

[0162] 相干参数的编码的概述在流程图，即图6中示出。

[0163] 第一操作是获得相干参数值，如图6中通过步骤501所示。

[0164] 在获得了用于帧的相干参数值后，下一个操作是生成用于每个子带的M维相干向量，如图6中通过步骤503所示。

[0165] 然后对M维相干向量进行变换，例如使用离散余弦变换（DCT），如图6中通过步骤505所示。

[0166] 然后,将DCT表示排序为低于所确定的子带选择值的子带和高于该值的子带,如图6中通过步骤507所示。换句话说,确定正在处理的当前子带是小于或等于I_N还是大于I_N。

[0167] 然后,通过对DCT变换向量的前2个分量进行编码,对针对小于或等于I_N的子带的M维相干向量的DCT表示进行编码,如图6中通过步骤509所示。

[0168] 然后,通过对DCT变换向量的前3个分量进行编码,对针对大于I_N的子带的M维相干向量的DCT表示进行编码,如图6中通过步骤511所示。

[0169] 例如,这可以概括为以下伪码形式。

For each subband i=1:N

[0170] The M dimensional vector of coherence data is DCT transformed

If i <= I_N

Encode the first 2 components of the DCT transformed vector

Else

[0171] Encode the first 3 components of the DCT transformed vector

End if

End for

[0172] 关于图5进一步详细示出了向量编码器405,根据一些实施例向量编码器405被示出接收DCT相干向量404作为输入。

[0173] 在一些实施例中,向量编码器包括DCT 0阶扩展相干位编码估计器(或第一/初级DCT相干参数估计器)451。

[0174] DCT 0阶扩展相干位编码估计器(或第一/初级DCT相干参数估计器)451被配置为接收DCT相干向量404,并由此确定是否所有的相干值都为非空的。当至少一个相干值为非空时,DCT 0阶扩展相干位编码估计器被配置为估计用于扩展相干的0阶DCT参数的编码的位数,用于联合编码: $\lceil \log_2 \Pi_i \text{len_cb_dct0}[\text{indexER}_i] \rceil$,其中 indexER_i 是子带i的量化能量比的指数,并且 $\text{len_cb_dct0}[\] = \{7, 6, 5, 4, 4, 4, 3, 2\}$ 。

[0175] 该估计被传递给码本确定器415。

[0176] 向量编码器此外在一些实施例中可以包括DCT1阶(和2阶往后)扩展相干编码器(或进一步的/次级DCT相干参数编码器)455。DCT1阶(和2阶往后)扩展相干编码器455被配置为接收DCT相干向量404,并由此对用于扩展相干的1阶(以及对于对进一步的次级参数进行编码的子带,2阶往后)的DCT参数进行编码,使用Golomb Rice编码用于量化指数的均值移除指数。在一些实施例中,指数从取决于子带的指数的码本中的标量量化获得。码字的数量对于所有子带是相同的,例如5个码字。

[0177] 输出的编码DCT1阶(和2阶往后)所、编码的扩展相干参数可以被准备作为编码的相干向量404的一部分输出。

[0178] 在一些实施例中,向量编码器此外可以包括环绕相干编码器457。环绕相干编码器457被配置为接收环绕相干参数,并由此对环绕相干参数进行编码并计算用于环绕相干的位数。在一些实施例中,环绕相干编码器457被配置每子带发送一个环绕相干值。以如关于

能量比的编码来描述的方式,在一些实施例中,该值可以作为子带的时频块的加权平均值获得,权重由信号能量确定。

[0179] 在一些实施例中,平均环绕相干值是用码本标量量化的,该码本的长度(码字数量)取决于能量比指数(对于指数:0,1,2,3,4,5,6,7,码字2,3,4,5,6,7,8,8)。在一些实施例中,使用针对均值移除值的Golomb Rice编码器,或者通过考虑到所使用的码字数量的联合编码,来对指数进行编码(换句话说,基于哪一个将值编码为较少的位,选择诸如GR编码的熵编码,或联合编码)。

[0180] 在一些实施例中,确定估计的(用于对初级扩展相干进行编码)和使用的(用于对二次扩展和环绕相干参数进行编码)的总位数,并由此确定可用于对方向参数进行编码的剩余位数。这例如可以在数学上确定为

[0181] $ED = B - (EPSC + SSC + SC + EP)$

[0182] 其中ED是可用的剩余位数,B是原始位目标,EPSC是用于对初级扩展相干参数进行编码的估计位数,SSC是用于对次级扩展相干参数进行编码的位数,SC是用于对环绕相干参数进行编码的位数,并且EP是用于对能量比进行编码的位数。

[0183] 可将可用的剩余位数传递给方向编码器,并确定用于根据任何合适的编码方法(例如如上所述)对方向参数进行编码将使用的位数。

[0184] 此外,在一些实施例中,向量编码器此外可以包括如前所讨论的码本确定器415。在一些实施例中,码本确定器415被配置为接收用于对DCT0阶扩展相干参数进行编码的位数的估计,并且此外接收编码/量化的能量比412和方位角414的编码方差。码本确定器415可以从这些输入确定用于对DCT 0阶扩展相干参数进行编码的合适码本。在一些实施例中,该确定基于能量比和量化方位角值(用于当前子带的量化方位角值的方差)。如果用于子带的方位角的方差低于所确定的阈值(例如阈值为30),则使用第一所确定码本,否则使用另一个所确定码本。在一些实施例中,对于0阶DCT系数,总共有16个码本(基于有用于能量比的8个指数和与给定阈值有关的对于方位角方差的2种可能性)。

[0185] 所选择的码本被传递给DCT 0阶扩展相干编码器453。

[0186] 此外,在一些实施例中,向量编码器此外可以包括DCT 0阶扩展相干编码器453。在接收到所确定的码本和DCT相干向量后,DCT 0阶扩展相干编码器453被配置为使用码本对DCT 0阶扩展相干进行编码,并将其作为编码的相干向量404输出。

[0187] 关于图7,示出了根据一些实施例的用于对能量比参数和方向参数(如虚线左侧所示)和相干参数(虚线右侧)进行编码的方法的流程图。

[0188] 在一些实施例中,使用每个值3位并通过使用优化标量量化(SQ)方法对能量比进行编码,如图7中通过步骤601所示。

[0189] 然后,如果至少一个相干值是非空的,则估计用于扩展相干的0阶DCT参数的编码的位数,如图7中通过步骤603所示。否则,如果输出全部为零,则仅发送一个位来表明该值为零。

[0190] 此外,该方法可以包括使用用于量化指数的均值移除指数的Golomb Rice编码,对用于扩展相干的1阶DCT参数进行编码,如图7中通过步骤605所示。如上面讨论的指数在一些实施例中可以从取决于子带的指数的码本中的标量量化获得。对于所有子带来说,码字的数量是相同的(例如5)。

[0191] 另外在一些实施例中,该方法进一步包括对环绕相干进行编码并计算用于环绕相干的位数,如图7中通过步骤607所示。在一些实施例中,如上面所讨论的,每个子带发送一个环绕相干值。此外在一些实施例中,该值是以与步骤601中用于能量比的方法类似的方式获得,作为子带的时频块的加权平均值,权重是信号能量。然后将平均的环绕相干值用码本进行标量化,该码本的长度(码字数量)取决于能量比指数(针对指数:0,1,2,3,4,5,6,7,码字2,3,4,5,6,7,8,8)。通过针对均值移除值进行编码的Golomb Rice,或通过考虑到所使用的码字数量的联合编码,对指数进行编码。

[0192] 在一些实施例中,该方法包括计算用于对方向参数进行编码的剩余位数,如图7中通过步骤609所示。

[0193] 在确定了用于对方向参数进行编码的剩余位数之后,于是对方向参数进行编码,如图7中通过步骤611所示。

[0194] 此外,该方法包括,使用取决于能量比和量化方位角值(用于当前子带的量化方位角值的方差)的码本,对用于扩展相干的0阶DCT系数进行编码,如图7中通过步骤613所示。该确定可以基于为能量比值范围选择两个可能的码本中的一个或另一个,该选择基于用于子带的方位角的方差低于(或高于)阈值。以此方式,可以有用于0阶DCT系数的总共16个码本(用于能量比的8个指数和与给定阈值的有关的用于方位角方差的两种可能性)。

[0195] 该操作可以以下面的代码表示。

```
static short quantize_coherence(IVAS_MASA_QDIRECTION* q_direction,
    unsigned char coding_subbands,
    unsigned char no_directions,
    short all_coherence_zero,
    short max_bits_coherence,
    IVAS_MASA_METADATA_FRAME* metadata,
    short write_flag,
    int * first_pos)
{
    short i, j, k;
    float dct_coh[MASA_MAXIMUM_CODING_SUBBANDS][MASA_SUBFRAMES];
    unsigned short idx_dct[MASA_SUBFRAMES*MASA_MAXIMUM_CODING_SUBBANDS];
    short nbits;
    int no_cb;
    short no_cb_vec[MASA_MAXIMUM_CODING_SUBBANDS];
    short bits_surround_coh;
```

```
[0196]
    if (all_coherence_zero == 1)
    {
        nbits = 0;
        return nbits;
    }
    else
    {
        for (i = 0; i < no_directions; i++)
        {
            k = 0;
            no_cb = 1;
            for (j = 0; j < coding_subbands; j++)
            {
                /* DCT 变换 */
                dct4_transform(q_direction[i].spread_coherence[j], dct_coh[j]);
                if (write_flag)
                {
                    /* 量化第一DCT参数 */
```

```

        dct_coh[j][0] = quantize_DCT_0_coh(dct_coh[j][0], j,
coherence_cb0, DELTA_AZI_DCT0, NO_CV_COH, &q_direction[i], &idx_dct[k],
&no_cb_vec[j]);
    }

    no_cb *= len_cb_dct0[q_direction->energy_ratio_index[j][0]];
    idx_dct[k + coding_subbands] = quantize_sq(dct_coh[j][1],
&coherence_cb1[j * NO_CV_COH1], NO_CV_COH1, &dct_coh[j][1]);
    k++;

    /* 选取第二DCT参数进行量化 */
    /*vec_dct_coh1[j] = dct_coh[j][1];*/
    if (j > 2)
    {
        dct_coh[j][2] = 0.0f; /* dct_coh[j][2]; */
    }
    else
    {
        dct_coh[j][2] = 0.0f;
    }

    dct_coh[j][3] = 0.0f;
}
if (write_flag)
{
    for (j = 0; j < coding_subbands; j++)
    {
        /* 逆 DCT 变换*/
        invdct4_transform(dct_coh[j],
q_direction[i].spread_coherence[j]);
    }
}

/* 编码指数并写位流 */
nbits = ceilf(logf((float)no_cb)*INV_LOG_2);
if (write_flag)
{
    nbits = encode_coherence_indexesDCT0(idx_dct, coding_subbands,
no_cb_vec, metadata, *first_pos);
}
else
{
    *first_pos = metadata->bit_pos;
    metadata->bit_pos += nbits;
    nbits += encode_coherence_indexesDCT1(&idx_dct[coding_subbands],
coding_subbands, no_cb_vec, metadata);
}

}

if (write_flag == 0)
{
    bits_surround_coh = max_bits_coherence - nbits;
    if (bits_surround_coh < MIN_BITS_SURR_COH)
    {
        bits_surround_coh = 0;
    }
    else
    {
        /* 编码环绕相干 */
        bits_surround_coh = encode_surround_coherence(bits_surround_coh,
q_direction, coding_subbands, no_directions, all_coherence_zero, metadata);
    }
}

```

[0197]

```

    }

    /* 输出位数量 */
    return nbits + bits_surround_coh;
}
else
{
    return nbits;
}
}
}

static short encode_coherence_indexesDCT0(unsigned short* idx_dct, short len,
short* no_cb_vec, IVAS_MASA_METADATA_FRAME* metadata, int first_pos)
{
    short nbits = 0;
    short i;
    int no_cb;
    unsigned short idx;

    /* 计算利用联合编码的dct0分量的位*/
    no_cb = no_cb_vec[0];
    for (i = 1; i < len; i++)
    {
        no_cb *= no_cb_vec[i];
    }
    nbits = ceilf(logf((float)no_cb)*INV_LOG_2);
    /* 创建组合的指数 */
    idx = create_combined_index(idx_dct, len, no_cb_vec);
    /* 写组合的指数 */
    first_pos = write_in_bit_buff(metadata->bit_buffer, idx, first_pos, nbits);

    return nbits;
}

static short encode_coherence_indexesDCT1(unsigned short* idx_dct, short len,
short* no_cb_vec, IVAS_MASA_METADATA_FRAME* metadata)
{
    short nbits = 0;
    short i;
    short GR_ord;
    short av;
    short data, bits_GR;
    unsigned short mr_idx_dct[MASA_MAXIMUM_CODING_SUBBANDS];

    GR_ord = 0;
    bits_GR = mean_removed_GR(idx_dct, len, 0, &GR_ord, &av, metadata,
mr_idx_dct);
    for (i = 0; i < len; i++)
    {
        data = GR_data(mr_idx_dct[i], GR_ord, &bits_GR, 0);
        nbits += bits_GR;
        metadata->bit_pos = write_in_bit_buff(metadata->bit_buffer, data,
metadata->bit_pos, bits_GR);
    }
    nbits += len_huf[av];
    metadata->bit_pos = write_in_bit_buff(metadata->bit_buffer,
huff_code_av[av], metadata->bit_pos, len_huf[av]);
}

```

[0198]

```

        return nbits;
    }

    static short mean_removed_GR(unsigned short* idx,
                                short len,
                                short adapt_GR,
                                short* GR_ord,
                                short* p_av,
                                IVAS_MASA_METADATA_FRAME* metadata,
                                unsigned short * mr_idx)
    {
        short av, i, nbits;
        short sh_idx[5];

        av = (short)roundf(sum_s((short*) idx, len) / (float)len);
        *p_av = av;
        for (i = 0; i < len; i++)
        {
            sh_idx[i] = idx[i] - av;
        }

        for (i = 0; i < len; i++)
        {
            if (sh_idx[i] < 0)
            {
                sh_idx[i] = -2*sh_idx[i];
            }
            else if (sh_idx[i] > 0)
            {
                sh_idx[i] = sh_idx[i] * 2 - 1;
            }
            else
            {
                sh_idx[i] = 0;
            }
            mr_idx[i] = (unsigned short)sh_idx[i];
        }

        nbits = GR_bits(mr_idx, len, *GR_ord, adapt_GR, GR_ord);
        return nbits;
    }
}

```

[0199]

[0200] 关于图8,示出了根据一些实施例的从对相干值进行提取和解码的视角来看作为解码器133的一部分的示例元数据提取器137。

[0201] 在一些实施例中,编码的数据流被传递给解多路复用器。解多路复用器提取编码的方向指数、能量比指数和相干指数,并且在一些实施例中还可以提取其他元数据和传送音频信号(未示出)。

[0202] 能量比指数可以由能量比解码器解码,通过执行由能量比编码器实现的能量比编码的逆转来生成针对帧的能量比。此外,能量比指数可以被传递给相干DCT向量发生器(并且在一些实施例中传递给码本确定器815)。

[0203] 方向指数可以由方向解码器解码,该方向解码器被配置为执行由方向编码器实现的方向值的编码的逆转。在一些实施例中,在对方向值进行解码后,确定方位角值的方差,并将其输出给相干DCT向量发生器(并且在一些实施例中输出给码本确定器815)。

[0204] 元数据提取器137在一些实施例中包括相干DCT向量发生器801(并且在一些实施

例中包括到码本确定器815)。相干DCT向量发生器801被配置为接收编码的相干值800,并且此外接收编码的能量比812和(解码的)方位角值的方差814。基于这些值选择或确定码本(例如码本确定器815可以与来自相干编码器209的码本确定器415相同)。

[0205] 在确定了码本之后,于是使用相干编码器中使用的编码方法的逆转对所接收的编码的相干指数进行解码,从而为扩展相干值和环绕相干值生成合适的DCT相干向量802。然后将DCT相干向量802传递给逆离散余弦变换器803。

[0206] 元数据提取器137在一些实施例中包括逆离散余弦变换器803。逆离散余弦变换器803被配置为接收(解码的)DCT相干向量802并生成相干向量804,该相干向量804被输出给向量解码器805。

[0207] 元数据提取器137在一些实施例中包括向量解码器805。向量解码器805被配置为接收解码的相干向量804并从中提取用于子带的相干参数806。

[0208] 关于图9示出了用于对扩展相干参数进行解码的方法的流程图。

[0209] 第一操作是获得(例如接收或检索)编码的扩展相干值,如图9中通过步骤901所示。

[0210] 在获得了编码的扩展相干值后,于是下一个操作是针对(每个)子带:读取第一DCT扩展相干参数指数(初级DCT参数),如图9中通过步骤903所示。

[0211] 尽管在图9中没有与获得编码的扩展相干值一起示出,但编码的环绕相干值、编码的能量比和编码的方位角和仰角值被获得。

[0212] 通过应用在编码器中执行的编码过程的逆转,对编码的能量比和编码的方位角和仰角值进行解码。首先对能量比进行解码。基于能量比值知道用于扩展相干DCT指数的位数。被发送用于对扩展相干的零阶DCT参数进行编码的指数首先被读取,并且仅在对方位角值进行解码之后才能被解码。

[0213] 此外,编码的环绕相干值基于应用编码器中编码过程的逆转进行解码。例如,这涉及基于能量比值选择合适的码本。

[0214] 下一个操作是基于量化的能量比和解码的方位角量化方差来确定用于第一DCT扩展相干参数的码本。确定了码本后,对第一DCT扩展相干参数指数进行解码,如图9中通过步骤905所示。

[0215] 下一个操作是确定正在解码的当前子带是否小于或等于编码器中使用的子带值(I_N),如图9中通过步骤907所示。

[0216] 在正在解码的当前子带小于或等于编码器中使用的子带值(I_N)的情况下,于是读取下一个(第一次级)DCT扩展相干参数,并使用编码器中实现的编码的逆转对其进行解码,如图9中通过步骤909所示。

[0217] 在正在解码的当前子带大于编码器中使用的子带值(I_N)的情况下,于是读取两个(第一次级和第二次级)DCT扩展相干参数,并使用编码器中实现的编码的逆转对其进行解码,如图9中通过步骤911所示。

[0218] 在对两个(或三个)DCT参数进行解码后,下一个操作是对参数执行逆DCT以生成解码的向量,如图9中通过步骤913所示。

[0219] 然后,可以将解码的向量读取作为用于该子带的时频块扩展相干值。下一个操作是检查是否所有的子带都已被解码,如图9中通过步骤915所示。

[0220] 当有另一个子带要解码时,操作可以循环回到步骤903。

[0221] 当所有的子带都被解码时,那么下一帧解码可以开始,如图9中通过步骤917所示(换句话说操作循环回到步骤901)。

[0222] 关于图10,示出了可用作分析或合成装置的示例电子设备。该设备可以是任何合适的电子设备或装置。例如在一些实施例中,设备1400是移动设备、用户装备、平板计算机、计算机、音频回放装置等。

[0223] 在一些实施例中,设备1400包括至少一个处理器或中央处理单元1407。处理器1407可以被配置为执行各种程序代码诸如方法,诸如本文中所述。

[0224] 在一些实施例中,设备1400包括存储器1411。在一些实施例中,至少一个处理器1407耦合到存储器1411。存储器1411可以是任何合适的存储手段。在一些实施例中,存储器1411包括用于存储可在处理器1407上实现的程序代码的程序代码区段。此外在一些实施例中,存储器1411可以进一步包括存储数据区段,其用于存储数据,例如根据本文中所述的实施例已处理或要处理的数据。存储在程序代码区段内的所实现的程序代码和存储在存储数据区段内的数据可以每当需要的时候经由存储器-处理器耦合由处理器1407检索。

[0225] 在一些实施例中,设备1400包括用户接口1405。在一些实施例中,用户接口1405可以耦合到处理器1407。在一些实施例中,处理器1407可以控制用户接口1405的操作并接收来自用户接口1405的输入。在一些实施例中,用户接口1405可以使用户能够向设备1400输入命令,例如经由键盘。在一些实施例中,用户接口1405可以使用户能够从设备1400获得信息。例如,用户接口1405可以包括被配置为向用户显示来自设备1400的信息的显示器。在一些实施例中,用户接口1405可以包括触摸屏或触摸界面,其既能够使信息能够被输入到设备1400,又能够进一步向设备1400的用户显示信息。在一些实施例中,用户接口1405可以是用于与如本文中所述的位置确定器通信的用户接口。

[0226] 在一些实施例中,设备1400包括输入/输出端口1409。在一些实施例中,输入/输出端口1409包括收发器。在此类实施例中,收发器可以耦合到处理器1407,并且被配置为实现与其他装置或电子设备的通信,例如经由无线通信网络。在一些实施例中,收发器或任何合适的收发器或发射器和/或接收器手段可以被配置为经由电线或有线耦合与其他电子设备或装置通信。

[0227] 收发器可以通过任何合适的已知通信协议与进一步的装置通信。例如在一些实施例中,收发器可以使用合适的通用移动通信系统(UMTS)协议、无线局域网(WLAN)协议(诸如,例如IEEE 802.X)、合适的短距离射频通信协议诸如蓝牙,或红外数据通信路径(IRDA)。

[0228] 收发器输入/输出端口1409可以被配置为接收信号,并且在一些实施例中,通过使用执行合适代码的处理器1407来确定如本文中所述的参数。此外,该装置可以生成合适的混缩信号和参数输出以传输到合成设备。

[0229] 在一些实施例中,设备1400可被用作合成设备的至少一部分。因此,输入/输出端口1409可以被配置为接收混缩信号,并且在一些实施例中,接收如本文中所述在采集设备或处理设备处确定的参数,并且通过使用执行合适代码的处理器1407来生成合适的音频信号格式输出。输入/输出端口1409可以耦合到任何合适的音频输出,例如耦合到多声道扬声器系统和/或耳机或类似的音频输出。

[0230] 一般来说,本发明的各种实施例可以在硬件或专用电路、软件、逻辑或其任何组合

中实现。例如,一些方面可以在硬件中实现,而其他方面可以在固件或软件中实现,该固件或软件可以由控制器、微处理器或其他计算设备执行,尽管本发明不限于此。尽管本发明的各方面可以示出和描述为框图、流程图或使用一些其他图形表示来示出和描述,但很好理解,本文中描述的这些块、装置、系统、技术或方法可以作为非限制性示例在硬件、软件、固件、专用电路或逻辑、通用硬件或控制器或其他计算设备,或其一些组合中实现。

[0231] 本发明的实施例可以由可由移动设备的数据处理器(诸如在处理器实体中)执行的计算机软件实现,或者由硬件实现,或者由软件和硬件的组合实现。在这方面进一步应该注意,如在图中的逻辑流程的任何块可以表示程序步骤,或互连的逻辑电路、块和功能,或程序步骤和逻辑电路、块和功能的组合。软件可以存储在物理介质上,诸如存储器芯片,或处理器内实现的存储器块,磁性介质,诸如硬盘或软盘,以及光学介质,诸如,例如DVD及其数据变体、CD。

[0232] 存储器可以具有适合本地技术环境的任何类型,并且可以使用任何合适的数据存储技术来实现,诸如基于半导体的存储设备、磁性存储设备和系统、光学存储设备和系统、固定存储器和可移除存储器。数据处理器可以具有适合本地技术环境的任何类型,并且可以包括以下中的一种或多种作为非限制性示例:通用计算机、专用计算机、微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、门级电路和基于多核处理器架构的处理器。

[0233] 本发明的实施例可以在各种部件诸如集成电路模块中实践。集成电路的设计大体上是高度自动化的过程。复杂且强大的软件工具可用于将逻辑级设计转换为准备在半导体衬底上蚀刻和形成的半导体电路设计。

[0234] 程序,诸如由加利福尼亚州山景城的Synopsys公司和加利福尼亚州圣何塞的Cadence设计公司提供的那些程序,使用既定的设计规则以及预先存储的设计模块库,自动地在半导体芯片上路由导体和定位部件。一旦完成了用于半导体电路的设计,则呈标准化的电子格式(例如,Opus、GDSII等)的所得设计可以被传输到半导体制造设施或“制造厂”用于制造。

[0235] 上述描述已通过示例性和非限制性的示例提供了本发明的示例性实施例的完整和信息性描述。然而,鉴于上述描述,当结合附图和所附权利要求阅读时,各种修改和适配对于相关领域技术人员来说可能变得明显。然而,对本发明的教导的所有此类和类似的修改仍将落在如所附权利要求中定义的本发明的范围内。

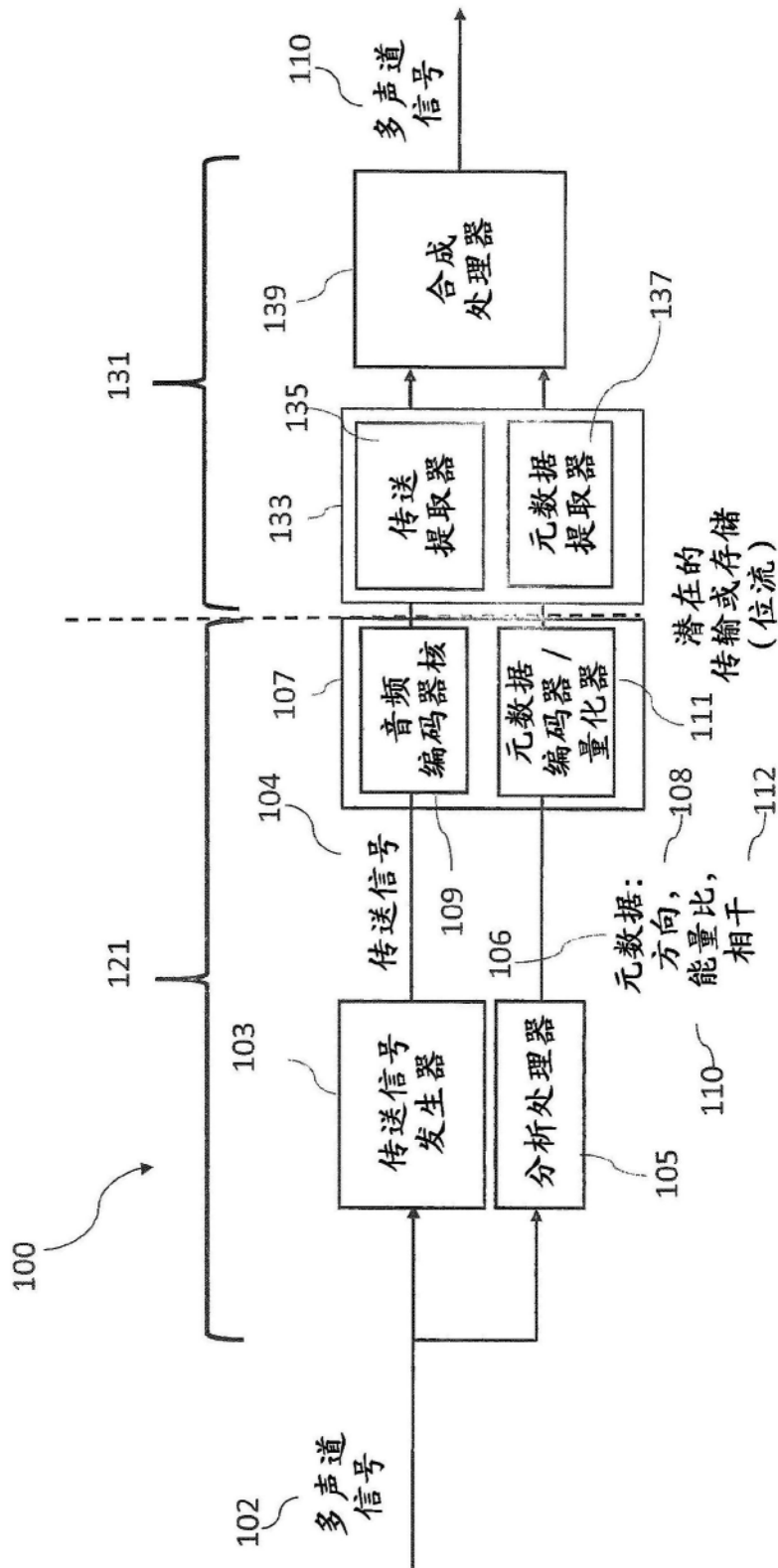


图1

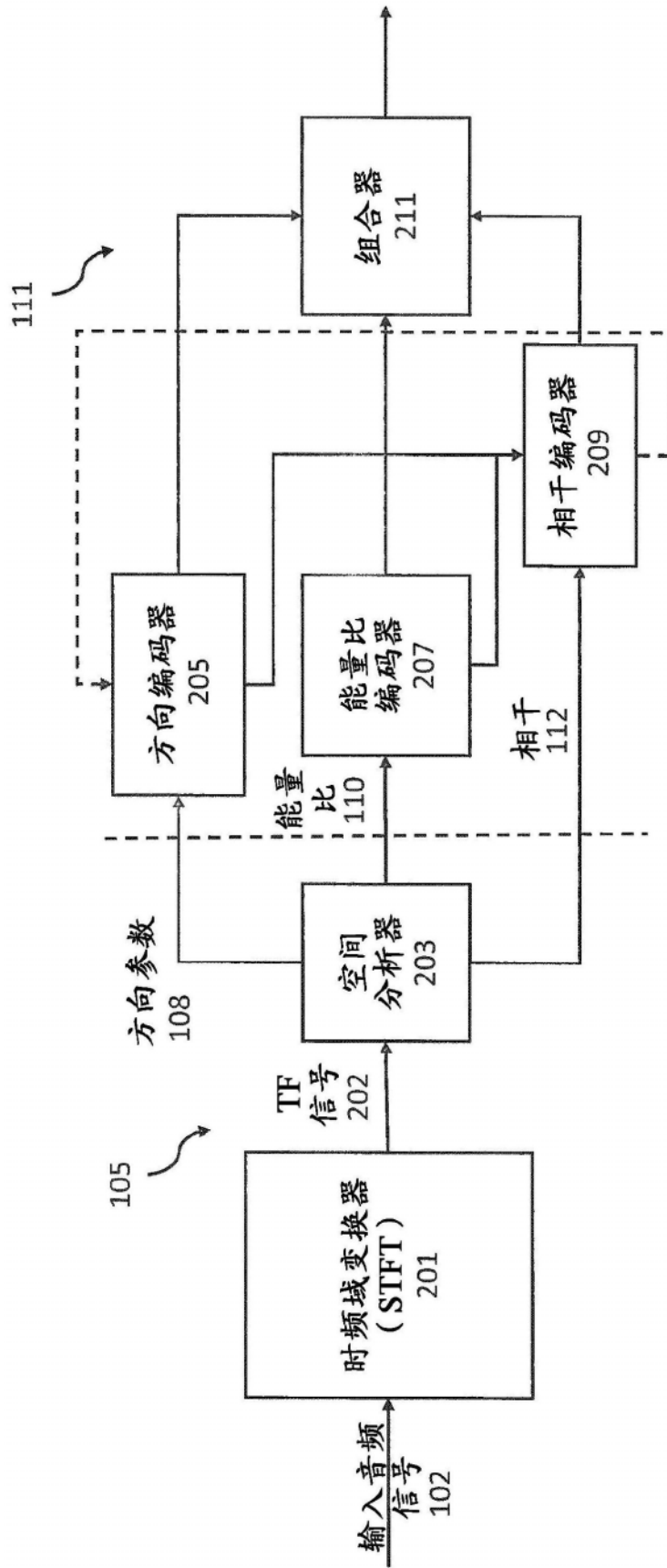


图2

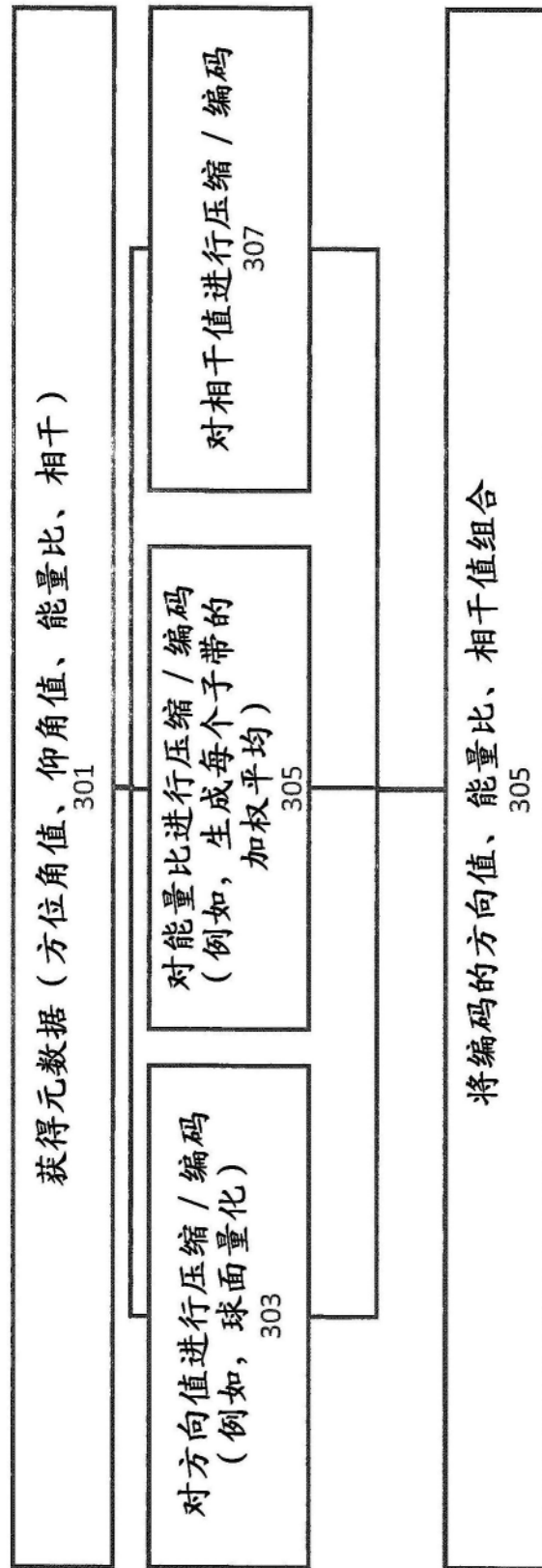


图3

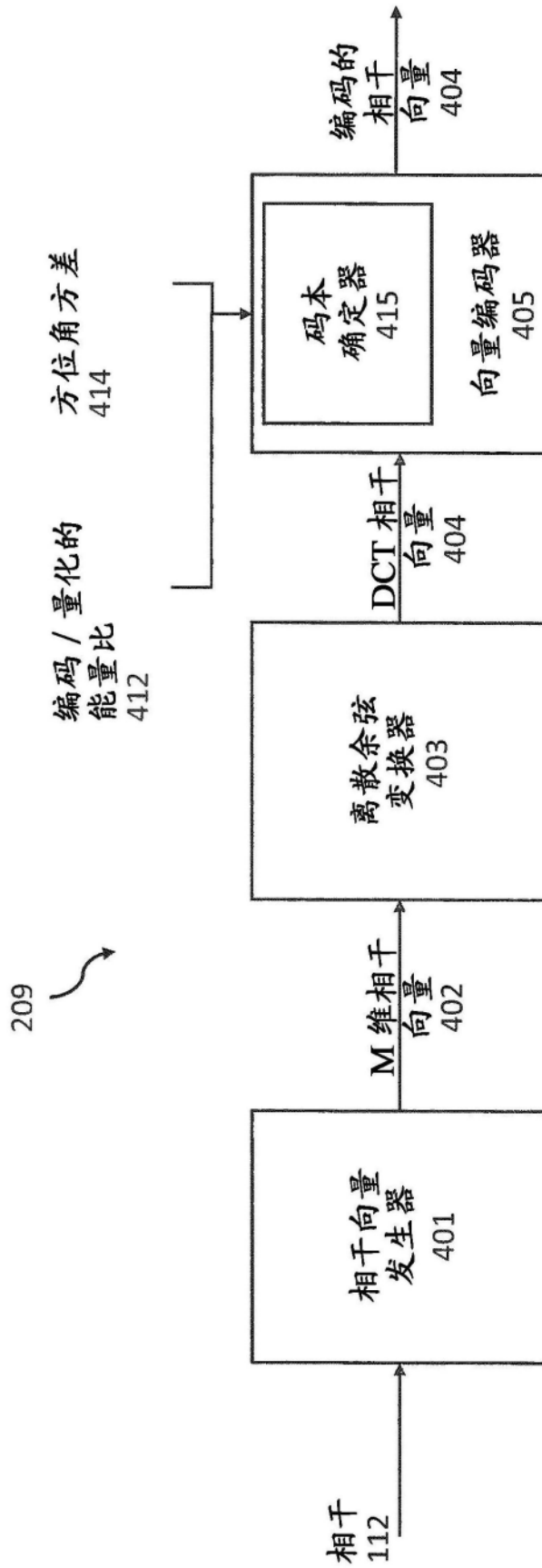


图4

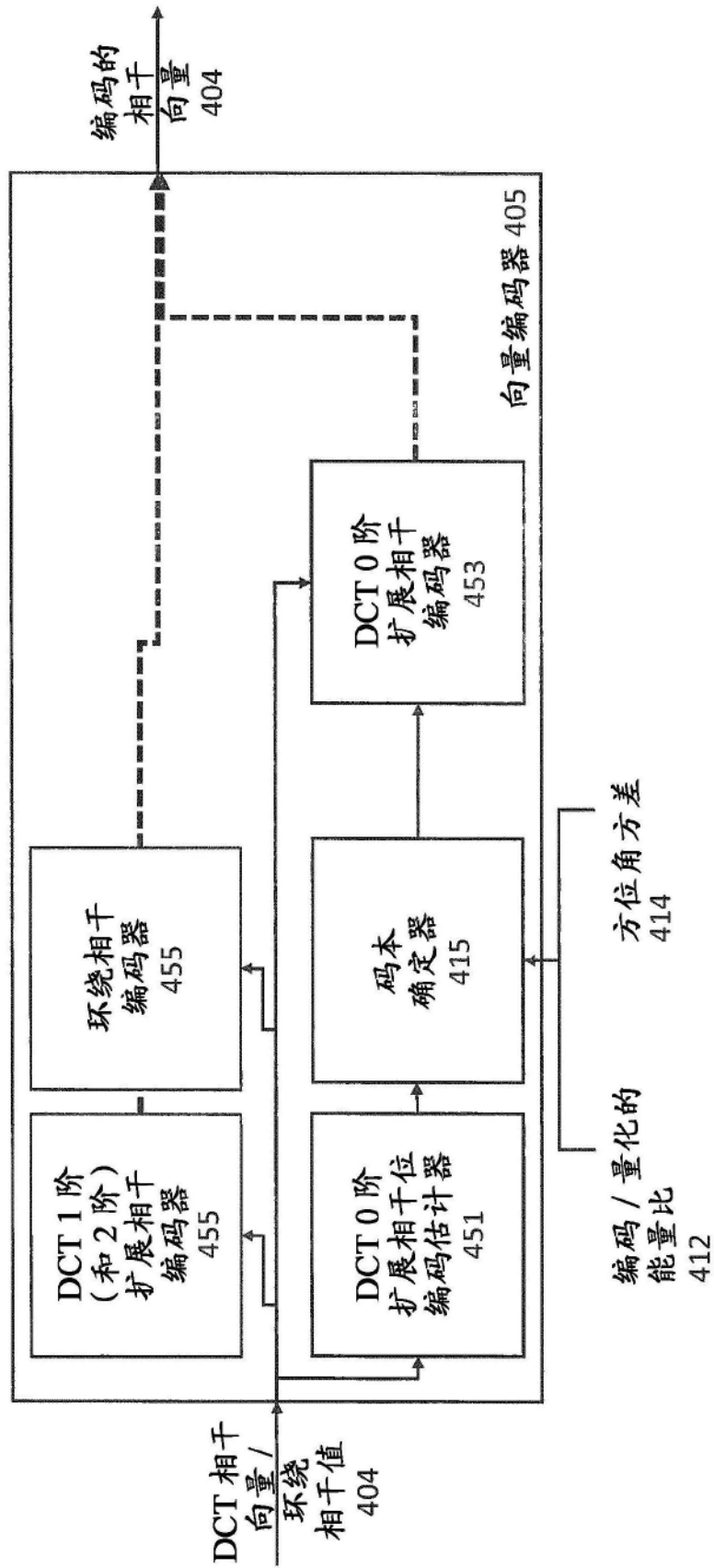


图5

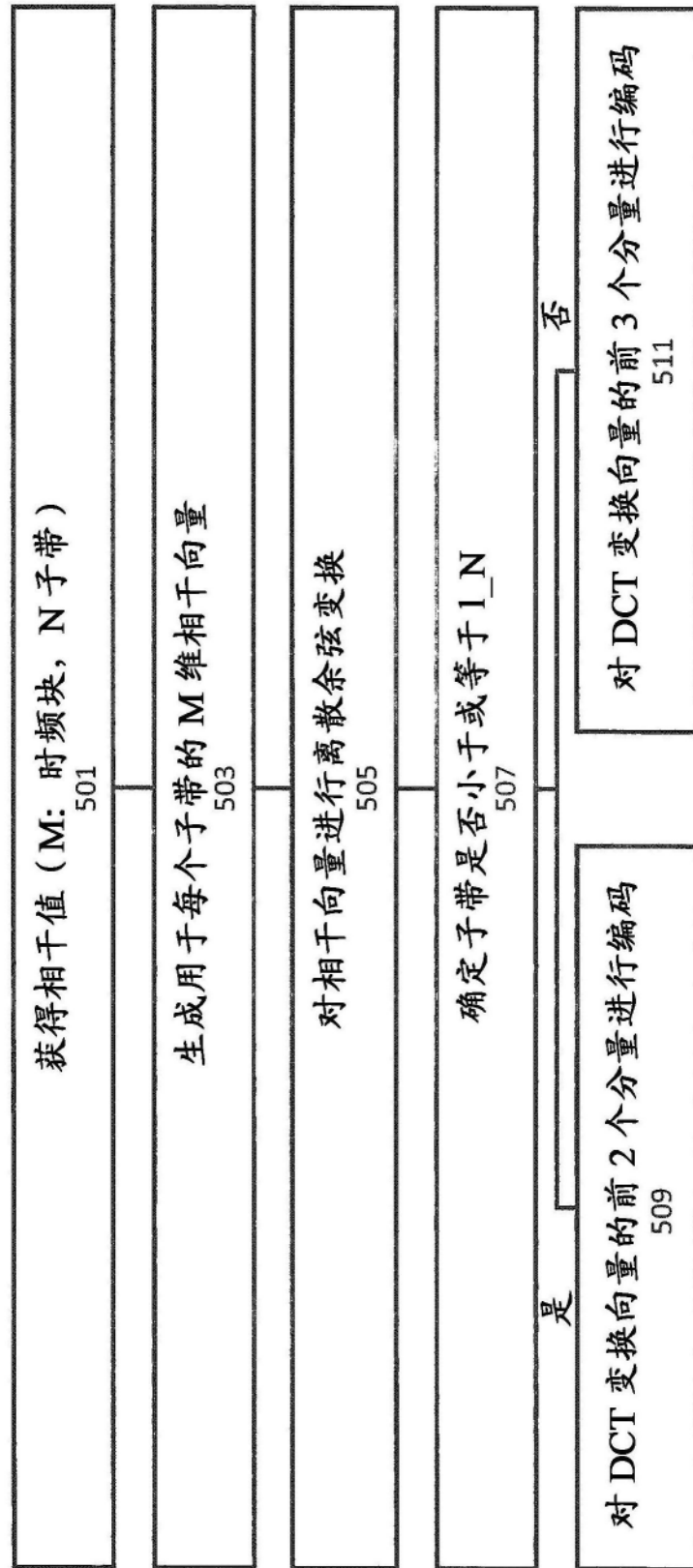


图6

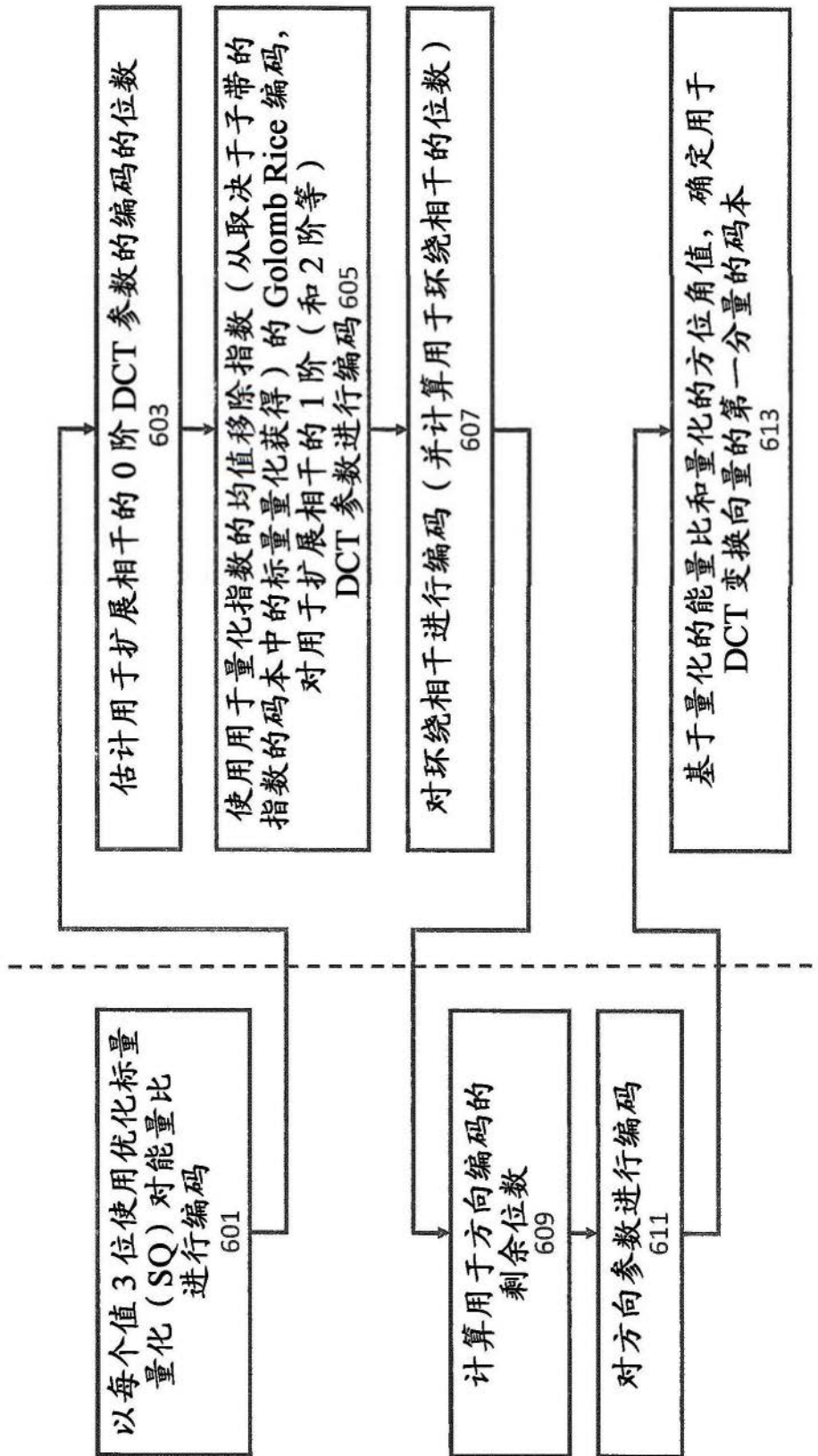


图7

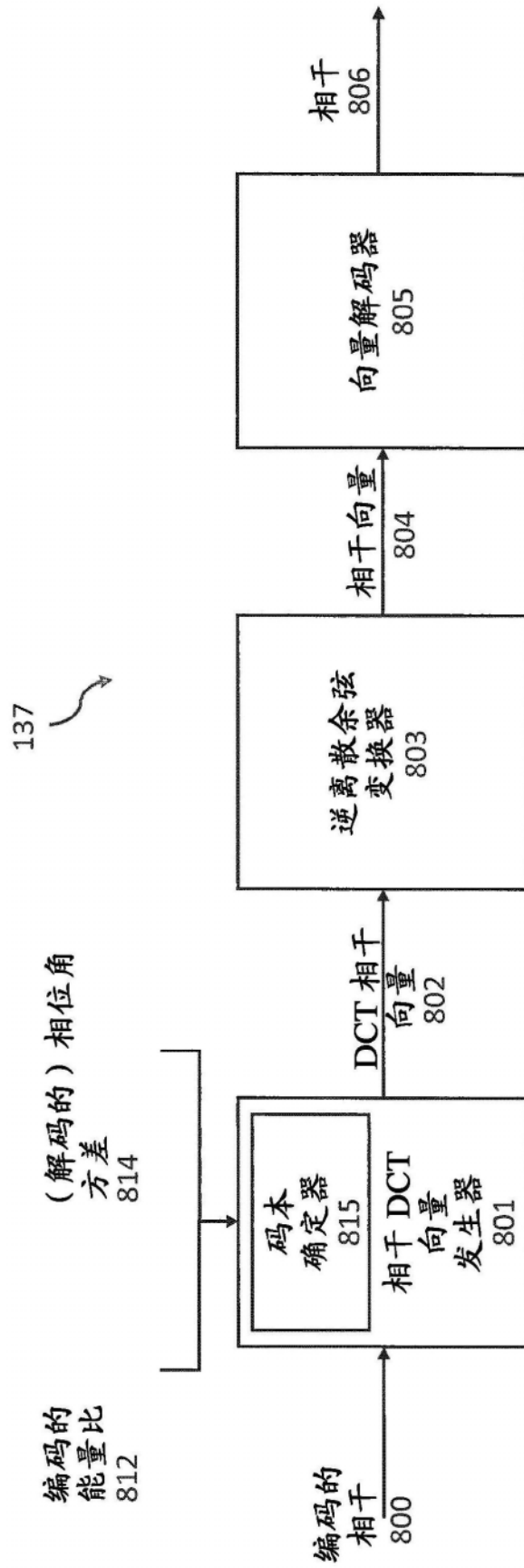


图8

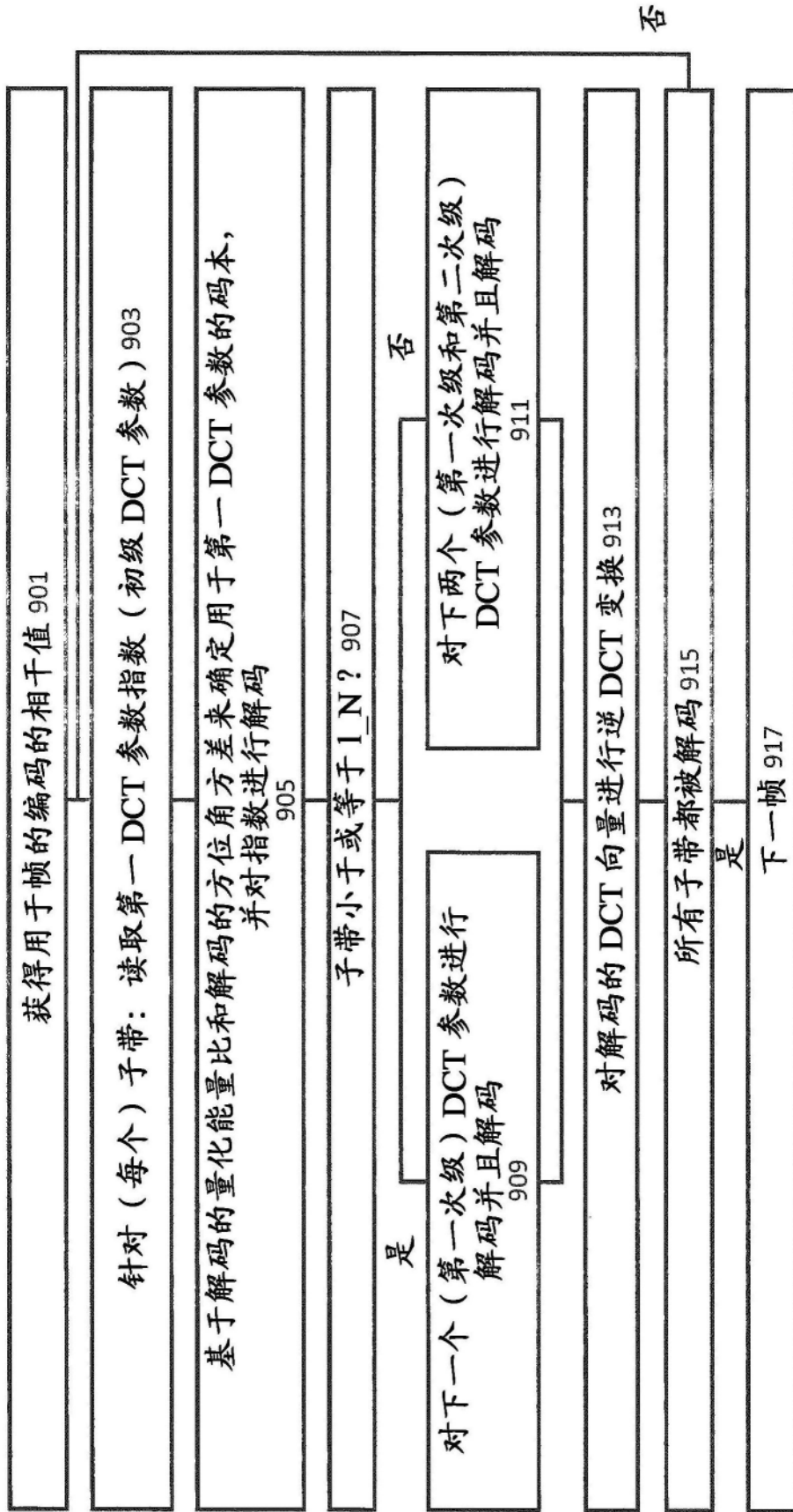


图9

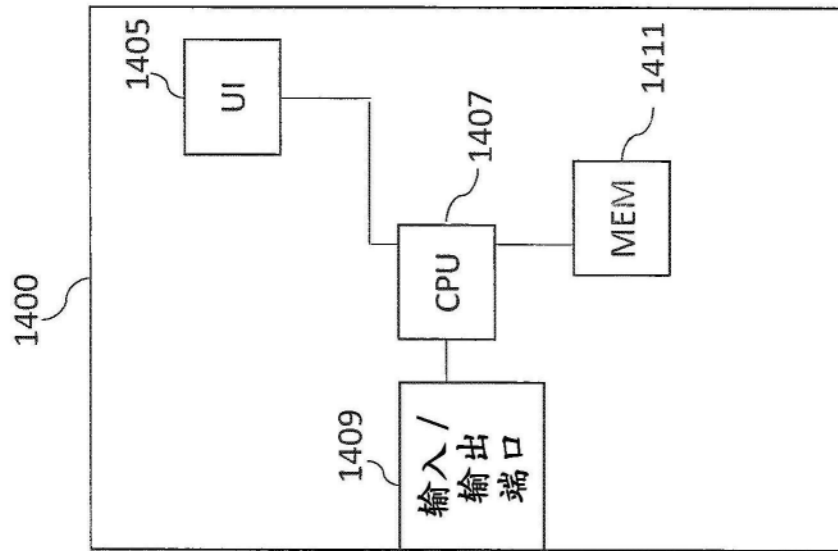


图10