



(12) 发明专利

(10) 授权公告号 CN 111105805 B

(45) 授权公告日 2025. 03. 04

(21) 申请号 201911231996.6

安德烈斯·霍瑟 亚琴·昆兹

(22) 申请日 2014.07.11

(74) 专利代理机构 中科专利商标代理有限责任公司 11021

(65) 同一申请的已公布的文献号  
申请公布号 CN 111105805 A

专利代理师 黄亮

(43) 申请公布日 2020.05.05

(51) Int.Cl.

(30) 优先权数据

G10L 19/008 (2013.01)

13177376.4 2013.07.22 EP

13189305.9 2013.10.18 EP

(56) 对比文件

CN 101689368 A, 2010.03.31

CN 102388417 A, 2012.03.21

(62) 分案原申请数据

201480041694.1 2014.07.11

审查员 周群

(73) 专利权人 弗劳恩霍夫应用研究促进会  
地址 德国慕尼黑

(72) 发明人 萨沙·迪克 克里斯汀·厄泰尔  
克里斯汀·赫姆瑞希  
约翰内斯·希尔珀特

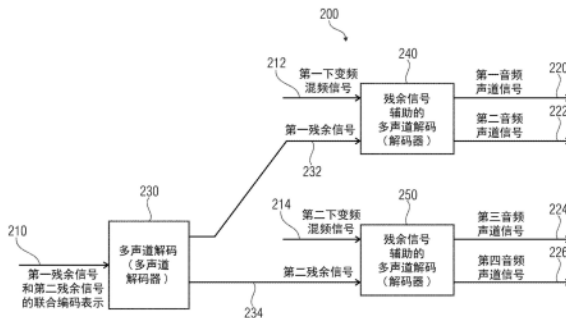
权利要求书6页 说明书30页 附图21页

(54) 发明名称

音频编码器、音频解码器、方法及计算机可读介质

(57) 摘要

用于基于已编码表示来提供至少四个音频声道信号的音频解码器被配置为使用多声道解码,基于第一残余信号及第二残余信号的联合编码表示来提供该第一残余信号及该第二残余信号。该音频解码器被配置为使用残余信号辅助的多声道解码,基于第一下混合信号及该第一残余信号来提供第一音频声道信号及第二音频声道信号。该音频解码器被配置为使用残余信号辅助的多声道解码,基于第二下混合信号及该第二残余信号来提供第三音频声道信号及第四音频声道信号。音频编码器基于对应的考虑。



1. 一种音频解码器(200;300;600;1300;1600;2000),用于基于已编码表示(210;310、360;610、682;1310、1312;1610)提供至少四个音频声道信号(220、222、224、226;320、322、324、326;620、622、624、626;1320、1322、1324、1326),

其中,所述音频解码器被配置为使用多声道解码(230;330;680;1360),基于第一残余信号及第二残余信号的联合编码表示(210;310;682;1312)来提供所述第一残余信号(232;332;684;1362)及所述第二残余信号(234;334;686;1364);

其中,所述音频解码器被配置为使用残余信号辅助的多声道解码(240;340;640;1370),基于第一下混合信号(212;312;632;1342)及所述第一残余信号来提供第一音频声道信号(220;320;642;1372)及第二音频声道信号(222;322;644;1374);以及

其中,所述音频解码器被配置为使用残余信号辅助的多声道解码(250;350;650;1380),基于第二下混合信号(214;314;634;1344)及所述第二残余信号来提供第三音频声道信号(224;324;656;1382)及第四音频声道信号(226;326;658;1384)。

2. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用多声道解码(370;630;1340),基于所述第一下混合信号及所述第二下混合信号的联合编码表示(360;610;1310)来提供所述第一下混合信号(212;312;632;1342)及所述第二下混合信号(214;314;634;1344)。

3. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用基于预测的多声道解码,基于所述第一残余信号及所述第二残余信号的联合编码表示来提供所述第一残余信号及所述第二残余信号。

4. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用残余信号辅助的多声道解码,基于所述第一残余信号及所述第二残余信号的联合编码表示来提供所述第一残余信号及所述第二残余信号。

5. 根据权利要求3所述的音频解码器,其中,所述基于预测的多声道解码被配置为:估计预测参数,所述预测参数描述使用先前帧的信号分量导出的信号分量对提供当前帧的残余信号的贡献。

6. 根据权利要求3所述的音频解码器,其中,所述基于预测的多声道解码被配置为:基于所述第一残余信号及所述第二残余信号的下混合信号且基于所述第一残余信号及所述第二残余信号的公共残余信号来获得所述第一残余信号及所述第二残余信号。

7. 根据权利要求6所述的音频解码器,其中,所述基于预测的多声道解码被配置为:以第一符号应用公共残余信号,以获得所述第一残余信号,且以第二符号应用公共残余信号,以获得所述第二残余信号,所述第二符号与所述第一符号相反。

8. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用在MDCT域中操作的多声道解码,基于所述第一残余信号及所述第二残余信号的联合编码表示来提供所述第一残余信号及所述第二残余信号。

9. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用USAC复杂立体声预测,基于所述第一残余信号及所述第二残余信号的联合编码表示来提供所述第一残余信号及所述第二残余信号。

10. 根据权利要求1所述的音频解码器,

其中,所述音频解码器被配置为使用基于参数的、残余信号辅助的多声道解码,基于所

述第一下混合信号及所述第一残余信号来提供所述第一音频声道信号及所述第二音频声道信号;以及

其中,所述音频解码器被配置为使用基于参数的、残余信号辅助的多声道解码,基于所述第二下混合信号及所述第二残余信号来提供所述第三音频声道信号及所述第四音频声道信号。

11.根据权利要求10所述的音频解码器,其中,所述基于参数的、残余信号辅助的多声道解码被配置为:估计描述两个声道之间的所需相关性及/或两个声道之间的阶差的一个或多个参数,以基于所述第一下混合信号和所述第二下混合信号中相应的一个下混合信号以及所述第一残余信号和所述第二残余信号中对应的一个残余信号来提供相应的两个或两个以上音频声道信号。

12.根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用在QMF域中操作的、残余信号辅助的多声道解码,基于所述第一下混合信号及所述第一残余信号来提供所述第一音频声道信号及所述第二音频声道信号;以及

所述音频解码器被配置为:使用在QMF域中操作的、残余信号辅助的多声道解码,基于所述第二下混合信号及所述第二残余信号来提供所述第三音频声道信号及所述第四音频声道信号。

13.根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用MPEG环绕声2-1-2解码或统一立体声解码,基于所述第一下混合信号及所述第一残余信号来提供所述第一音频声道信号及所述第二音频声道信号;以及

所述音频解码器被配置为:使用MPEG环绕声2-1-2解码或统一立体声解码,基于所述第二下混合信号及所述第二残余信号来提供所述第三音频声道信号及所述第四音频声道信号。

14.根据权利要求1所述的音频解码器,其中,所述第一残余信号及所述第二残余信号与音频场景的不同水平位置或与所述音频场景的不同方位角位置相关联。

15.根据权利要求1所述的音频解码器,其中,所述第一音频声道信号及所述第二音频声道信号与音频场景的垂直相邻的位置相关联,以及

所述第三音频声道信号及所述第四音频声道信号与所述音频场景的垂直相邻的位置相关联。

16.根据权利要求1所述的音频解码器,其中,所述第一音频声道信号及所述第二音频声道信号与音频场景的第一水平位置或方位角位置相关联,以及

所述第三音频声道信号及所述第四音频声道信号与所述音频场景的第二水平位置或方位角位置相关联,所述第二水平位置或方位角位置不同于所述第一水平位置或方位角位置。

17.根据权利要求1所述的音频解码器,其中,所述第一残余信号与音频场景的左侧相关联,且所述第二残余信号与音频场景的右侧相关联。

18.根据权利要求17所述的音频解码器,

其中,所述第一音频声道信号及所述第二音频声道信号与所述音频场景的左侧相关联,以及

所述第三音频声道信号及所述第四音频声道信号与所述音频场景的右侧相关联。

19. 根据权利要求18所述的音频解码器,其中,所述第一音频声道信号与所述音频场景的左下位置相关联,

所述第二音频声道信号与所述音频场景的左上位置相关联,

所述第三音频声道信号与所述音频场景的右下位置相关联,以及

所述第四音频声道信号与所述音频场景的右上位置相关联。

20. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用多声道解码,基于所述第一下混合信号及所述第二下混合信号的联合编码表示来提供所述第一下混合信号及所述第二下混合信号,所述第一下混合信号与音频场景的左侧相关联,且所述第二下混合信号与所述音频场景的右侧相关联。

21. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用基于预测的多声道解码,基于所述第一下混合信号及所述第二下混合信号的联合编码表示来提供所述第一下混合信号及所述第二下混合信号。

22. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为使用残余信号辅助的、基于预测的多声道解码,基于所述第一下混合信号及所述第二下混合信号的联合编码表示来提供所述第一下混合信号及所述第二下混合信号。

23. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:基于所述第一音频声道信号及所述第三音频声道信号执行第一多声道带宽扩展(660;1390),以及

所述音频解码器被配置为:基于所述第二音频声道信号及所述第四音频声道信号执行第二多声道带宽扩展(670;1394)。

24. 根据权利要求23所述的音频解码器,其中,所述音频解码器被配置为:基于所述第一音频声道信号及所述第三音频声道信号以及一个或多个带宽扩展参数(1338)执行所述第一多声道带宽扩展,以获得与音频场景的第一公共水平面或第一公共高度相关联的两个或两个以上带宽扩展的音频声道信号(620、624;1320、1324),以及

所述音频解码器被配置为:基于所述第二音频声道信号及所述第四音频声道信号以及一个或多个带宽扩展参数(1358)执行所述第二多声道带宽扩展,以获得与所述音频场景的第二公共水平面或第二公共高度相关联的两个或两个以上带宽扩展的音频声道信号(622、626;1322、1326)。

25. 根据权利要求1所述的音频解码器,其中,所述第一残余信号及所述第二残余信号的联合编码表示包括声道对单元,所述声道对单元包括所述第一残余信号及所述第二残余信号的下混合信号以及所述第一残余信号及所述第二残余信号的公共残余信号。

26. 根据权利要求1所述的音频解码器,其中,所述音频解码器被配置为:使用多声道解码,基于所述第一下混合信号及所述第二下混合信号的联合编码表示来提供所述第一下混合信号及所述第二下混合信号,

所述第一下混合信号及所述第二下混合信号的联合编码表示包括声道对单元,所述声道对单元包括所述第一下混合信号及所述第二下混合信号的下混合信号以及所述第一下混合信号及所述第二下混合信号的公共残余信号。

27. 一种音频编码器(100;1100;1200;1500;2100),用于基于至少四个音频声道信号(110、112、114、116;1110、1112、1114、1116;1210、1212、1214、1216;2216、2226、2218、2228)提供已编码表示(130;1144、1154;1220、1222;2272、2282),

其中,所述音频编码器被配置为使用残余信号辅助的多声道编码(140;1120;1230;2230)对至少第一音频声道信号及第二音频声道信号进行联合编码,以获得第一下混合信号(120;1122;1232;2234)及第一残余信号(142;1124;1234;2236);以及

其中,所述音频编码器被配置为使用残余信号辅助的多声道编码(150;1130;1240;2240)对至少第三音频声道信号及第四音频声道信号进行联合编码,以获得第二下混合信号(122;1132;1242;2244)及第二残余信号(152;1134;1244;2246);以及

其中,所述音频编码器被配置为使用多声道编码(160;1150;1260;2260)对所述第一残余信号及所述第二残余信号进行联合编码,以获得所述第一残余信号和所述第二残余信号的联合编码表示(130;1154;1262;2264)。

28.根据权利要求27所述的音频编码器,其中,所述音频编码器被配置为:使用多声道编码(1140;1250;2250)对所述第一下混合信号及所述第二下混合信号进行联合编码,以获得所述第一下混合信号和所述第二下混合信号的联合编码表示(1144;1252;2254)。

29.根据权利要求28所述的音频编码器,其中,所述音频编码器被配置为:使用基于预测的多声道编码对所述第一残余信号及所述第二残余信号进行联合编码,以及

所述音频编码器被配置为:使用基于预测的多声道编码对所述第一下混合信号及所述第二下混合信号进行联合编码。

30.根据权利要求27所述的音频编码器,其中,所述音频编码器被配置为:使用基于参数的、残余信号辅助的多声道编码对至少所述第一音频声道信号及所述第二音频声道信号进行联合编码,以及

所述音频编码器被配置为:使用基于参数的、残余信号辅助的多声道编码对至少所述第三音频声道信号及所述第四音频声道信号进行联合编码。

31.根据权利要求27所述的音频编码器,其中,所述第一音频声道信号及所述第二音频声道信号与音频场景的垂直相邻的位置相关联,以及

所述第三音频声道信号及所述第四音频声道信号与所述音频场景的垂直相邻的位置相关联。

32.根据权利要求27所述的音频编码器,其中,所述第一音频声道信号及所述第二音频声道信号与音频场景的第一水平位置或方位角位置相关联,以及

所述第三音频声道信号及所述第四音频声道信号与所述音频场景的第二水平位置或方位角位置相关联,所述第二水平位置或方位角位置不同于所述第一水平位置或方位角位置。

33.根据权利要求27所述的音频编码器,其中,所述第一残余信号与音频场景的左侧相关联,且所述第二残余信号与所述音频场景的右侧相关联。

34.根据权利要求33所述的音频编码器,

其中,所述第一音频声道信号及所述第二音频声道信号与所述音频场景的左侧相关联,以及

其中,所述第三音频声道信号及所述第四音频声道信号与所述音频场景的右侧相关联。

35.根据权利要求34所述的音频编码器,其中,所述第一音频声道信号与所述音频场景的左下位置相关联,

所述第二音频声道信号与所述音频场景的左上位置相关联,  
所述第三音频声道信号与所述音频场景的右下位置相关联,以及  
所述第四音频声道信号与所述音频场景的右上位置相关联。

36. 根据权利要求27所述的音频编码器,其中,所述音频编码器被配置为:使用多声道编码对所述第一下混合信号及所述第二下混合信号进行联合编码,以获得所述第一下混合信号和所述第二下混合信号的联合编码表示,所述第一下混合信号与音频场景的左侧相关联,且所述第二下混合信号与所述音频场景的右侧相关联。

37. 一种用于基于已编码表示来提供至少四个音频声道信号的方法(800),所述方法包括:

使用多声道解码,基于第一残余信号及第二残余信号的联合编码表示来提供(810)所述第一残余信号及所述第二残余信号;

使用残余信号辅助的多声道解码,基于第一下混合信号及所述第一残余信号来提供(820)第一音频声道信号及第二音频声道信号;以及

使用残余信号辅助的多声道解码,基于第二下混合信号及所述第二残余信号来提供(830)第三音频声道信号及第四音频声道信号。

38. 一种用于基于至少四个音频声道信号来提供已编码表示的方法(700),所述方法包括:

使用残余信号辅助的多声道编码对至少第一音频声道信号及第二音频声道信号进行联合编码(710),以获得第一下混合信号及第一残余信号;

使用残余信号辅助的多声道编码对至少第三音频声道信号及第四音频声道信号进行联合编码(720),以获得第二下混合信号及第二残余信号;以及

使用多声道编码对所述第一残余信号及所述第二残余信号进行联合编码(730),以获得所述第一残余信号和所述第二残余信号的已编码表示。

39. 一种包括计算机程序的计算机程序产品,当在计算机上执行所述计算机程序时,所述计算机程序用于执行根据权利要求37或38的方法。

40. 一种音频解码器(200;300;600;1300;1600;2000),用于基于已编码表示(210;310;360;610;682;1310、1312;1610)提供至少四个音频声道信号(220、222、224、226;320、322、324、326;620、622、624、626;1320、1322、1324、1326),

其中,所述音频解码器被配置为使用多声道解码(230;330;680;1360),基于第一残余信号及第二残余信号的联合编码表示(210;310;682;1312)来提供所述第一残余信号(232;332;684;1362)及所述第二残余信号(234;334;686;1364);

其中,所述音频解码器被配置为使用残余信号辅助的多声道解码(240;340;640;1370),基于第一下混合信号(212;312;632;1342)及所述第一残余信号来提供第一音频声道信号(220;320;642;1372)及第二音频声道信号(222;322;644;1374);以及

其中,所述音频解码器被配置为使用残余信号辅助的多声道解码(250;350;650;1380),基于第二下混合信号(214;314;634;1344)及所述第二残余信号来提供第三音频声道信号(224;324;656;1382)及第四音频声道信号(226;326;658;1384);

其中,所述音频解码器被配置为:基于所述第一音频声道信号及所述第三音频声道信号执行第一多声道带宽扩展(660;1390),以及

其中,所述音频解码器被配置为:基于所述第二音频声道信号及所述第四音频声道信号执行第二多声道带宽扩展(670;1394);

其中,所述音频解码器被配置为:基于所述第一音频声道信号及所述第三音频声道信号以及一个或多个带宽扩展参数(1338)执行所述第一多声道带宽扩展,以获得与音频场景的第一公共水平面或第一公共高度相关联的两个或两个以上带宽扩展的音频声道信号(620、624;1320、1324),以及

所述音频解码器被配置为:基于所述第二音频声道信号及所述第四音频声道信号以及一个或多个带宽扩展参数(1358)执行所述第二多声道带宽扩展,以获得与所述音频场景的第二公共水平面或第二公共高度相关联的两个或两个以上带宽扩展的音频声道信号(622、626;1322、1326)。

41.一种用于基于已编码表示来提供至少四个音频声道信号的方法(800),所述方法包括:

使用多声道解码,基于第一残余信号及第二残余信号的联合编码表示来提供(810)所述第一残余信号及所述第二残余信号;

使用残余信号辅助的多声道解码,基于第一下混合信号及所述第一残余信号来提供(820)第一音频声道信号及第二音频声道信号;以及

使用残余信号辅助的多声道解码,基于第二下混合信号及所述第二残余信号来提供(830)第三音频声道信号及第四音频声道信号;

其中,所述方法包括:基于所述第一音频声道信号及所述第三音频声道信号执行第一多声道带宽扩展(660;1390),以及

其中,所述方法包括:基于所述第二音频声道信号及所述第四音频声道信号执行第二多声道带宽扩展(670;1394);

其中,基于所述第一音频声道信号及所述第三音频声道信号以及一个或多个带宽扩展参数(1338)执行所述第一多声道带宽扩展,以获得与音频场景的第一公共水平面或第一公共高度相关联的两个或两个以上带宽扩展的音频声道信号(620、624;1320、1324),以及

其中,基于所述第二音频声道信号及所述第四音频声道信号以及一个或多个带宽扩展参数(1358)执行所述第二多声道带宽扩展,以获得与所述音频场景的第二公共水平面或第二公共高度相关联的两个或两个以上带宽扩展的音频声道信号(622、626;1322、1326)。

42.一种包括计算机程序的计算机程序产品,当在计算机上执行所述计算机程序时,所述计算机程序用于执行根据权利要求41的方法。

## 音频编码器、音频解码器、方法及计算机可读介质

[0001] 本申请是国际申请日为2014年7月11日, 优先权日为2013年7月22日的国际申请PCT/EP2014/064915进入中国国家阶段的对应的申请号为201480041694.1的中国发明专利申请“使用联合编码残余信号的音频解码器、音频编码器、方法和计算机可读介质”的分案申请。

### 技术领域

[0002] 根据本发明的实施例涉及用于基于已编码表示来提供至少四个音频声道信号的音频解码器。

[0003] 根据本发明的另一实施例涉及用于基于至少四个音频声道信号来提供已编码表示的音频编码器。

[0004] 根据本发明的另一实施例涉及用于基于已编码表示来提供至少四个音频声道信号的方法及用于基于至少四个音频声道信号来提供已编码表示的方法。

[0005] 根据本发明的另一实施例涉及用于执行所述方法之一的计算机程序。

[0006] 一般而言, 根据本发明的实施例涉及n个声道的联合编码。

### 背景技术

[0007] 近年来, 对音频内容的储存及发送的需求一直在稳定地增加。此外, 对音频内容的储存及发送的质量要求也一直在稳定地增加。因此, 用于音频内容的编码及解码的概念已得到增强。例如, 已开发了所谓的“先进音频编码”(AAC), 在例如国际标准ISO/IEC 13818-7:2003中描述了该“先进音频编码”。此外, 已创建一些空间延伸, 例如所谓的“MPEG环绕声”, 在例如国际标准ISO/IEC 23003-1:2007中对其进行进行了描述。此外, 在国际标准ISO/IEC 23003-2:2010中描述了用于编码及解码音频信号的空间信息的额外改进, 该国际标准涉及所谓的空间音频对象编码(SAOC)。

[0008] 此外, 在国际标准ISO/IEC 23003-3:2012中定义了灵活音频编码/解码概念, 灵活音频编码/解码概念提供以良好的编码效率编码一般音频信号及语言信号两者且处理多声道音频信号的可能性, 该国际标准描述所谓的“统一语音及音频编码”(USAC)概念。

[0009] 在MPEG USAC [1]中, 使用具有频带受限残余信号或全频带残余信号的复杂预测、MPS 2-1-1或统一立体声来执行两个声道的联合立体声编码。

[0010] MPEG环绕声[2]分层地组合OTT框及TTT框, 以在发送残余信号或不发送残余信号的情况下进行多声道音频的联合编码。

[0011] 然而, 希望提供用于三维音频场景的有效编码及解码的甚至更先进的概念。

### 发明内容

[0012] 根据本发明的实施例创建一种用于基于已编码表示来提供至少四个音频声道信号的音频解码器。该音频解码器被配置为: 使用多声道解码, 基于第一残余信号及第二残余信号的联合编码表示来提供该第一残余信号及该第二残余信号。该音频解码器还被配置

为:使用残余信号辅助的多声道解码,基于第一下混合信号及该第一残余信号来提供第一音频声道信号及第二音频声道信号。该音频解码器还被配置为:使用残余信号辅助的多声道解码,基于第二下混合信号及该第二残余信号来提供第三音频声道信号及第四音频声道信号。

[0013] 根据本发明的实施例基于发现可通过根据残余信号的联合编码表示导出两个残余信号来利用四个或甚至更多音频声道信号之间的依从性,该两个残余信号中每一个被用于使用残余信号辅助的多声道解码来提供两个或两个以上音频声道信号。换言之,已发现通常存在所述残余信号的一些相似性,使得可通过使用多声道解码从联合编码表示导出这两个残余信号来减少在解码至少四个音频声道信号时有助于改进音频质量的用于编码所述残余信号的比特率,这利用了所述残余信号之间的相似性及/或依从性。

[0014] 在优选实施例中,该音频解码器被配置为使用多声道解码,基于该第一下混合信号及该第二下混合信号的联合编码表示来提供该第一下混合信号及该第二下混合信号。因此,创建音频解码器的分层结构,其中使用分离的多声道解码导出在用于提供至少四个音频声道信号的残余信号辅助的多声道解码中使用的下混合信号及残余信号两者。此概念尤其有效,因为两个下混合信号通常包括可在多声道编码/解码中使用的相似性,且因为两个残余信号通常还包括可在多声道编码/解码中使用的相似性。因而,使用此概念通常可获得良好的编码效率。

[0015] 在优选实施例中,该音频解码器被配置为使用基于预测的多声道解码,基于该第一残余信号及该第二残余信号的该联合编码表示来提供该第一残余信号及该第二残余信号。基于预测的多声道解码的使用通常带来相当良好的残余信号重建质量。如果第一残余信号表示音频场景的左侧且第二残余信号表示音频场景的右侧,则此状况例如有利的,因为人类听觉对于音频场景的左侧与右侧之间的差异通常相当敏感。

[0016] 在优选实施例中,该音频解码器被配置为使用残余信号辅助的多声道解码,基于该第一残余信号及该第二残余信号的联合编码表示来提供该第一残余信号及该第二残余信号。已发现,如果第一残余信号及第二残余信号是使用还接收残余信号(且通常还接收下混合信号,该下混合信号组合第一残余信号及第二残余信号)的多声道解码来提供的,则可实现第一残余信号及第二残余信号的尤其良好的质量。因而,存在解码阶段的级联,在该级联中,基于输入的下混合信号及输入的残余信号来提供两个残余信号(用于提供第一音频声道信号及第二音频声道信号的第一残余信号,及用于提供第三音频声道信号及第四音频声道信号的第二残余信号),其中该输入的残余信号还可指定为(该第一残余信号及该第二残余信号的)公共残余信号。因而,第一残余信号及第二残余信号事实上是“中间”残余信号,该“中间”残余信号是使用多声道解码从对应的下混合信号及对应的“公共”残余信号导出的。

[0017] 在优选实施例中,该基于预测的多声道解码被配置为估计预测参数,该预测参数描述使用先前帧的信号分量导出的信号分量对提供当前帧的残余信号(即,第一残余信号及第二残余信号)的贡献。该基于预测的多声道解码的使用带来残余信号(第一残余信号及第二残余信号)的尤其良好的质量。

[0018] 在优选实施例中,该基于预测的多声道解码被配置为基于(对应的)下混合信号及(对应的)“公共”残余信号获得该第一残余信号及该第二残余信号,其中该基于预测的多声

道解码被配置为以第一符号应用公共残余信号,以获得该第一残余信号,且以第二符号应用公共残余信号,以获得该第二残余信号,该第二符号与该第一符号相反。已发现此基于预测的多声道解码带来重建该第一残余信号及该第二残余信号的良好效率。

[0019] 在优选实施例中,该音频解码器被配置为使用在修改型离散余弦变换域(MDCT域)中操作的多声道解码,基于该第一残余信号及该第二残余信号的联合编码表示来提供该第一残余信号及该第二残余信号。已发现可通过有效的方式来实现该概念,因为可用来提供该第一残余信号及该第二残余信号的联合编码表示的音频解码优选地在MDCT域中操作。因此,可通过在MDCT域中应用提供该第一残余信号及该第二残余信号的多声道解码来避免中间转换。

[0020] 在优选实施例中,该音频解码器被配置为使用USAC复杂立体声预测(例如,如以上引用的USAC标准中所提及的),基于该第一残余信号及该第二残余信号的联合编码表示来提供该第一残余信号及该第二残余信号。已发现此USAC复杂立体声预测带来该第一残余信号及该第二残余信号的良好解码结果。此外,将USAC复杂立体声预测使用于第一残余信号及第二残余信号的解码还使得可使用在统一语音及音频编码(USAC)中已可用的解码区块来简单地实现该概念。因此,可容易地重新配置统一语言及音频编码解码器来执行在此论述的解码概念。

[0021] 在优选实施例中,该音频解码器被配置为使用基于参数的、残余信号辅助的多声道解码,基于该第一下混合信号及该第一残余信号来提供该第一音频声道信号及该第二音频声道信号。类似地,该音频解码器被配置为使用基于参数的、残余信号辅助的多声道解码,基于该第二下混合信号及该第二残余信号来提供该第三音频声道信号及该第四音频声道信号。已发现该多声道解码极其适合于基于第一下混合信号、第一残余信号、第二下混合信号及第二残余信号进行的音频声道信号导出。此外,已发现可使用已存在于典型多声道音频解码器中的处理区块来以较小的努力实现该基于参数的、残余信号辅助的多声道解码。

[0022] 在优选实施例中,该基于参数的、残余信号辅助的多声道解码被配置为估计描述两个声道之间的所需相关性及/或两个声道之间的阶差的一个或多个参数,以便基于相应下混合信号及相应的对应残余信号来提供两个或两个以上音频声道信号。已发现此基于参数的、残余信号辅助的多声道解码极其适于级联多声道解码的第二阶段(其中,优选地,使用基于预测的多声道解码提供第一下混合信号及第二下混合信号以及第一残余信号及第二残余信号)。

[0023] 在优选实施例中,该音频解码器被配置为使用在QMF域中操作的、残余信号辅助的多声道解码,基于该第一下混合信号及该第一残余信号来提供该第一音频声道信号及该第二音频声道信号。类似地,该音频解码器优选地被配置为使用在QMF域中操作的、残余信号辅助的多声道解码,基于该第二下混合信号及该第二残余信号来提供该第三音频声道信号及该第四音频声道信号。因此,在QMF域中操作分层多声道解码的第二阶段,该第二阶段极其适于典型的后期处理,该第二阶段还通常在QMF域中执行,使得可避免中间转换。

[0024] 在优选实施例中,该音频解码器被配置为使用MPEG环绕声2-1-2解码或统一立体声解码,基于该第一下混合信号及该第一残余信号来提供该第一音频声道信号及该第二音频声道信号。类似地,该音频解码器优选地被配置为使用MPEG环绕声2-1-2解码或统一立体

声解码,基于该第二下混合信号及该第二残余信号来提供该第三音频声道信号及该第四音频声道信号。已发现此解码概念尤其适合于分层解码的第二阶段。

[0025] 在优选实施例中,该第一残余信号及该第二残余信号与音频场景的不同水平位置(或等效地,方位角位置)相关联。已发现,在分层多声道处理的第一阶段中将与不同水平位置(或方位角位置)相关联的残余信号分离是尤其有利的,因为如果在分层多声道解码的第一阶段中执行在知觉上重要的左/右分离,则可获得尤其良好的听觉印象。

[0026] 在优选实施例中,该第一音频声道信号及该第二声道信号与该音频场景的垂直相邻的位置(或等效地,该音频场景的相邻的高度位置)相关联。此外,该第三音频声道信号及该第四音频声道信号优选地与该音频场景的垂直相邻的位置(或等效地,该音频场景的相邻的高度位置)相关联。已发现,如果在分层音频解码的第二阶段中执行较高信号与较低信号之间的分离(该分离通常包括比第一阶段稍小的分离精确度),则可实现良好的解码结果,因为在与音频来源的水平位置相比时,人类听觉系统对于音频来源的垂直位置不太敏感。

[0027] 在优选实施例中,该第一音频声道信号及该第二音频声道信号与音频场景的第一水平位置(或等效地,方位角位置)相关联,且该第三音频声道信号及该第四音频声道信号与该音频场景的第二水平位置(或等效地,方位角位置)相关联,该第二水平位置(或等效地,方位角位置)不同于该第一水平位置(或等效地,方位角位置)。

[0028] 优选地,该第一残余信号与音频场景的左侧相关联,且该第二残余信号与音频场景的右侧相关联。因此,在分层音频解码的第一阶段中执行左右分离。

[0029] 在优选实施例中,该第一音频声道信号及该第二音频声道信号与音频场景的左侧相关联,且该第三音频声道信号及该第四音频声道信号与该音频场景的右侧相关联。

[0030] 在另一优选实施例中,该第一音频声道信号与音频场景的左下侧相关联,该第二音频声道信号与该音频场景的左上侧相关联,该第三音频声道信号与该音频场景的右下侧相关联,且该第四音频声道信号与该音频场景的右上侧相关联。此音频声道信号的关联性带来尤其良好的编码结果。

[0031] 在优选实施例中,该音频解码器被配置为使用多声道解码,基于该第一下混合信号及该第二下混合信号的联合编码表示来提供该第一下混合信号及该第二下混合信号,其中该第一下混合信号与音频场景的左侧相关联,且该第二下混合信号与该音频场景的右侧相关联。已发现,即使下混合信号与音频场景的不同侧相关联,还可使用多声道编码来以良好的编码效率编码下混合信号。

[0032] 在优选实施例中,该音频解码器被配置为使用基于预测的多声道解码或甚至使用残余信号辅助的、基于预测的多声道解码,基于该第一下混合信号及该第二下混合信号的联合编码表示来提供该第一下混合信号及该第二下混合信号。已发现,对此类多声道解码概念的使用提供了尤其良好的解码结果。此外,可在一些音频解码器中重新使用现有解码功能。

[0033] 在优选实施例中,该音频解码器被配置为基于该第一音频声道信号及该第三音频声道信号执行第一多声道带宽扩展。此外,该音频解码器可被配置为基于该第二音频声道信号及该第四音频声道信号执行第二(通常分离的)多声道带宽扩展。已发现,基于与音频场景的不同侧相关联的两个音频声道信号(其中不同的残余信号通常与该音频场景的不同

侧相关联)来执行可能的带宽扩展是有利的。

[0034] 在优选实施例中,该音频解码器被配置为基于该第一音频声道信号及该第三音频声道信号以及一个或多个带宽扩展参数来执行第一多声道带宽扩展,以获得与音频场景的第一公共水平面(或等效地,第一公共高度)相关联的两个或两个以上带宽扩展的音频声道信号。此外,该音频解码器优选地被配置为基于该第二音频声道信号及该第四音频声道信号以及一个或多个带宽扩展参数来执行第二多声道带宽扩展,以获得与该音频场景的第二公共水平面(或等效地,第二公共高度)相关联的两个或两个以上带宽扩展的音频声道信号。已发现,此解码方案导致良好的音频质量,因为在此布置中,多声道带宽扩展可考虑到立体声特性,该立体声特性对于听觉印象为重要的。

[0035] 在优选实施例中,该第一残余信号及该第二残余信号的联合编码表示包括声道对单元,该声道对单元包括该第一残余信号及该第二残余信号的下混合信号以及该第一残余信号及该第二残余信号的公共残余信号。已发现,使用声道对单元进行的该第一残余信号及该第二残余信号的下混合信号以及该第一残余信号及该第二残余信号的公共残余信号的编码是有利的,因为该第一残余信号及该第二残余信号的下混合信号以及该第一残余信号及该第二残余信号的公共残余信号通常共享多个特性。因此,对声道对单元的使用通常减少信号发送开销且因此使得可进行有效编码。

[0036] 在另一优选实施例中,该音频解码器被配置为使用多声道解码,基于该第一下混合信号及该第二下混合信号的联合编码表示来提供该第一下混合信号及该第二下混合信号,其中该第一下混合信号及该第二下混合信号的联合编码表示包括声道对单元。该声道对单元包括该第一下混合信号及该第二下混合信号的下混合信号以及该第一下混合信号及该第二下混合信号的公共残余信号。此实施例基于与以上所述实施例相同的考虑。

[0037] 根据本发明的另一实施例创建一种用于基于至少四个音频声道信号来提供已编码表示的音频编码器。该音频编码器被配置为使用残余信号辅助的多声道编码对至少第一音频声道信号及第二音频声道信号进行联合编码,以获得第一下混合信号及第一残余信号。该音频编码器被配置为使用残余信号辅助的多声道编码对至少第三音频声道信号及第四音频声道信号进行联合编码,以获得第二下混合信号及第二残余信号。此外,该音频编码器被配置为使用多声道编码对该第一残余信号及该第二残余信号进行联合编码,以获得该残余信号的联合编码表示。此音频编码器基于与以上所述音频解码器相同的考虑。

[0038] 此外,此音频编码器的可选改进及该音频编码器的优选配置实质上与以上论述的音频解码器的改进及优选配置并行。因此,对以上论述进行参考。

[0039] 根据本发明的另一实施例创建一种用于基于已编码表示来提供至少四个音频声道信号的方法,该方法实质上执行以上所述音频编码器的功能,且该方法可由以上论述的任一特征及功能补充。

[0040] 根据本发明的另一实施例创建一种用于基于至少四个音频声道信号来提供已编码表示的方法,该方法实质上实现以上所述音频解码器的功能。

[0041] 根据本发明的另一实施例创建一种用于执行以上提及的方法的计算机程序。

## 附图说明

[0042] 随后将参考附图来描述根据本发明的实施例,在附图中:

- [0043] 图1示出了根据本发明的实施例的音频编码器的示意框图；
- [0044] 图2示出了根据本发明的实施例的音频解码器的示意框图；
- [0045] 图3示出了根据本发明的另一实施例的音频解码器的示意框图；
- [0046] 图4示出了根据本发明的实施例的音频编码器的示意框图；
- [0047] 图5示出了根据本发明的实施例的音频解码器的示意框图；
- [0048] 图6A和图6B示出了根据本发明的另一实施例的音频解码器的示意框图；
- [0049] 图7示出了根据本发明的实施例的用于基于至少四个音频声道信号来提供已编码表示的方法的流程图；
- [0050] 图8示出了根据本发明的实施例的用于基于已编码表示来提供至少四个音频声道信号的方法的流程图；
- [0051] 图9示出了根据本发明的实施例的用于基于至少四个音频声道信号来提供已编码表示的方法的流程图；以及
- [0052] 图10示出了根据本发明的实施例的用于基于已编码表示来提供至少四个音频声道信号的方法的流程图；
- [0053] 图11示出了根据本发明的实施例的音频编码器的示意框图；
- [0054] 图12示出了根据本发明的另一实施例的音频编码器的示意框图；
- [0055] 图13展示根据本发明的实施例的音频解码器的示意框图；
- [0056] 图14A示出了比特流的语法表示,该语法表示可与根据图13的音频编码器一起使用；
- [0057] 图14B示出了参数qceIndex的不同的值的表格表示；
- [0058] 图15示出了可使用根据本发明的概念的3D音频编码器的示意框图；
- [0059] 图16示出了可使用根据本发明的概念的3D音频解码器的示意框图；以及
- [0060] 图17示出了格式转换器的示意框图。
- [0061] 图18示出了根据本发明的实施例的四声道单元(QCE)的拓扑结构的图解表示；
- [0062] 图19示出了根据本发明的实施例的音频解码器的示意框图；
- [0063] 图20示出了根据本发明的实施例的QCE解码器的详细示意框图；以及
- [0064] 图21示出了根据本发明的实施例的四声道编码器的详细示意框图。

## 具体实施方式

### [0065] 1. 根据图1的音频编码器

[0066] 图1示出了音频编码器的示意框图,该音频编码器全部以100指定。音频编码器100被配置为基于至少四个音频声道信号提供已编码表示。音频编码器100被配置为接收第一音频声道信号110、第二音频声道信号112、第三音频声道信号114及第四音频声道信号116。此外,音频编码器100被配置为提供第一下混合信号120的已编码表示及第二下混合信号122的已编码表示,以及残余信号的联合编码表示130。音频编码器100包括残余信号辅助的多声道编码器140,该残余信号辅助的多声道编码器被配置为使用残余信号辅助的多声道编码来对第一音频声道信号110及第二音频声道信号112进行联合编码,以获得第一下混合信号120及第一残余信号142。音频信号编码器100还包括残余信号辅助的多声道编码器150,该残余信号辅助的多声道编码器被配置为使用残余信号辅助的多声道编码对至少第

三音频声道信号114及第四音频声道信号116进行联合编码,以获得第二下混合信号122及第二残余信号152。音频解码器100还包括多声道编码器160,该多声道编码器被配置为使用多声道编码对第一残余信号142及第二残余信号152进行联合编码,以获得残余信号142、152的联合编码表示130。

[0067] 关于音频编码器100的功能,应注意音频编码器100执行分层编码,其中使用残余信号辅助的多声道编码140对第一音频声道信号110及第二音频声道信号112进行联合编码,其中提供第一下混合信号120及第一残余信号142两者。第一残余信号142可例如描述第一音频声道信号110与第二音频声道信号112之间的差异,和/或可描述不能由第一下混合信号120及可选参数表示的一些或任何信号特征,该可选参数可由残余信号辅助的多声道编码器140提供。换言之,第一残余信号142可以是考虑到可基于第一下混合信号120及任何可能的参数获得的解码结果的精炼的残余信号,该任何可能的参数可由残余信号辅助的多声道编码器140提供。例如,在与高阶信号特性(类似例如,相关性特性、协方差特性、阶差特性,等等)的纯粹重建相比时,第一残余信号142可至少考虑到音频解码器侧的第一音频声道信号110及第二音频声道信号112的部分波形重建。类似地,残余信号辅助的多声道编码器150基于第三音频声道信号114及第四音频声道信号116提供第二下混合信号122及第二残余信号152两者,使得第二残余信号考虑到在音频解码器的侧第三音频声道信号114及第四音频声道信号116的信号重建的精炼。第二残余信号152可因此充当与第一残余信号142相同的功能。然而,如果音频声道信号110、112、114、116包括一些相关性,则第一残余信号142及第二残余信号152通常还在某种程度上相关。因此,使用多声道编码器160进行的第一残余信号142及第二残余信号152的联合编码通常包括高效率,因为相关的信号的多声道编码通常通过利用依从性来降低比特率。因此,可利用良好的精确度来对第一残余信号142及第二残余信号152进行编码,同时保持残余信号的联合编码表示130的比特率合理地小。

[0068] 简而言之,根据图1的实施例提供分层多声道编码,其中可通过使用残余信号辅助的多声道编码器140、150实现良好的重现质量,且其中可通过联合编码第一残余信号142及第二残余信号152保持适度的比特率需求。

[0069] 音频编码器100的另一可选改进是可能的。将参考图4、图11及图12描述这些改进中的一些。然而,应注意,音频编码器100还可适配为与本文所述的音频解码器并行,其中音频编码器的功能通常与音频解码器的功能相反。

[0070] 2. 根据图2的音频解码器

[0071] 图2示出了音频解码器的示意框图,该音频解码器全部以200指定。

[0072] 音频解码器200被配置为接收已编码表示,该已编码表示包括第一残余信号及第二残余信号的联合编码表示210。音频解码器200还接收第一下混合信号212及第二下混合信号214的表示。音频解码器200被配置为提供第一音频声道信号220、第二音频声道信号222、第三音频声道信号224及第四音频声道信号226。

[0073] 音频解码器200包括多声道解码器230,该多声道解码器被配置为基于第一残余信号232及第二残余信号234的联合编码表示210来提供第一残余信号232及第二残余信号234。音频解码器200还包括(第一)残余信号辅助的多声道解码器240,该残余信号辅助的多声道解码器被配置为使用多声道解码,基于第一下混合信号212及第一残余信号232来提供第一音频声道信号220及第二音频声道信号222。音频解码器200还包括(第二)残余信号辅

助的多声道解码器250,该残余信号辅助的多声道解码器被配置为基于第二下混合信号214及第二残余信号234提供第三音频声道信号224及第四音频声道信号226。

[0074] 关于音频解码器200的功能,应注意,音频信号解码器200基于(第一)公共残余信号辅助的多声道解码器240来提供第一音频声道信号220及第二音频声道信号222,其中由第一残余信号232提高多声道解码的解码质量(在与非残余信号辅助的解码相比时)。换言之,第一下混合信号212提供关于第一音频声道信号220及第二音频声道信号222的“粗略”信息,其中,例如,第一音频声道信号220与第二音频声道信号222之间的差异可由(可选)参数并由第一残余信号232描述,该(可选)参数可由残余信号辅助的多声道解码器240接收。因此,第一残余信号232可例如考虑到第一音频声道信号220及第二音频声道信号222的部分波形重建。

[0075] 类似地,(第二)残余信号辅助的多声道解码器250基于第二下混合信号214提供第三音频声道信号224及第四音频声道信号226,其中第二下混合信号214可例如“粗略地”描述第三音频声道信号224及第四音频声道信号226。此外,第三音频声道信号224与第四音频声道信号226之间的差异可例如由(可选的)参数并由第二残余信号234描述,该(可选的)参数可由(第二)残余信号辅助的多声道解码器250接收。因此,第二残余信号234的估计可例如考虑到第三音频声道信号224及第四音频声道信号226的部分波形重建。因此,第二残余信号234可考虑到第三音频声道信号224及第四音频声道信号226的重建质量的增强。

[0076] 然而,第一残余信号232及第二残余信号234是从第一残余信号及第二残余信号的联合编码表示210导出的。由多声道解码器230执行的这种多声道解码考虑到高解码效率,因为第一音频声道信号220、第二音频声道信号222、第三音频声道信号224及第四音频声道信号226通常类似或“相关”。因此,第一残余信号232及第二残余信号234通常也类似或“相关”,可通过使用多声道解码从联合编码表示210导出第一残余信号232及第二残余信号234来利用这种情况。

[0077] 因此,有可能通过基于残余信号232、234的联合编码表示210解码残余信号,且通过将残余信号中每一个用于两个或两个以上音频声道信号的解码来获得具有适度比特率的高解码质量。

[0078] 总而言之,音频解码器200通过提供高质量音频声道信号220、222、224、226来考虑到高编码效率。

[0079] 应注意,随后将参考图3、图5、图6A、图6B及图13来描述可在音频解码器200中可选地实现的附加特征及功能。然而,应注意,音频编码器200可在无任何附加修改的情况下包括以上提及的优点。

[0080] 3. 根据图3的音频解码器

[0081] 图3示出了根据本发明的另一实施例的音频解码器的示意框图。图3的音频解码器全部以300指定。音频解码器300类似于根据图2的音频解码器200,使得以上的解释也适用。然而,音频解码器300在与音频解码器200相比时补充了附加特征和功能,如下文中将解释。

[0082] 音频解码器300被配置为接收第一残余信号及第二残余信号的联合编码表示310。此外,音频解码器300被配置为接收第一下混合信号及第二下混合信号的联合编码表示360。此外,音频解码器300被配置为提供第一音频声道信号320、第二音频声道信号322、第三音频声道信号324及第四音频声道信号326。音频解码器300包括多声道解码器330,该多

声道解码器被配置为接收第一残余信号及第二残余信号的联合编码表示310,且基于该联合编码表示提供第一残余信号332及第二残余信号334。音频解码器300还包括(第一)残余信号辅助的多声道解码340,该(第一)残余信号辅助的多声道解码接收第一残余信号332及第一下混合信号312,且提供第一音频声道信号320及第二音频声道信号322。音频解码器300还包括(第二)残余信号辅助的多声道解码350,该残余信号辅助的多声道解码器被配置为接收第二残余信号334及第二下混合信号314,且提供第三音频声道信号324及第四音频声道信号326。

[0083] 音频解码器300还包括另一多声道解码器370,该另一多声道解码器被配置为接收第一下混合信号及第二下混合信号的联合编码表示360,且基于该联合编码表示提供第一下混合信号312及第二下混合信号314。

[0084] 在下文中,将描述音频解码器300的其他一些特定细节。然而,应注意,实际的音频解码器无需实现所有这些附加特征和功能的组合。相反,下文中所述的特征及功能可单独地添加至音频解码器200(或任何其他音频解码器),以逐步改进音频解码器200(或任何其他音频解码器)。

[0085] 在优选实施例中,音频解码器300接收第一残余信号及第二残余信号的联合编码表示310,其中联合编码表示310可包括第一残余信号332及第二残余信号334的下混合信号,以及第一残余信号332及第二残余信号334的公共残余信号。另外,联合编码表示310可例如包括一个或多个预测参数。因此,多声道解码器330可以是基于预测的残余信号辅助的多声道解码器。例如,多声道解码器330可以是如例如国际标准ISO/IEC 23003-3:2012的“复杂立体声预测”部分中所述的USAC复杂立体声预测。例如,多声道解码器330可被配置为估计预测参数,该预测参数描述使用先前帧的信号分量导出的信号分量对提供当前帧的第一残余信号332及第二残余信号334的贡献。此外,多声道解码器330可被配置为以第一符号应用公共残余信号(该公共残余信号包括在联合编码表示310中),以获得第一残余信号332,以及以与第一符号相反的第二符号应用公共残余信号(该公共残余信号包括在联合编码表示310中),以获得第二残余信号334。因而,公共残余信号可至少部分地描述第一残余信号332与第二残余信号334之间的差异。然而,多声道解码器330可估计下混合信号、公共残余信号及一个或多个预测参数(这些参数都包括在联合编码表示310中),以获得第一残余信号332及第二残余信号334,如以上引用的国际标准ISO/IEC 23003-3:2012中所述。此外,应注意,第一残余信号332可与第一水平位置(或方位角位置)(例如,左水平位置)相关联,且第二残余信号334可与音频场景的第二水平位置(或方位角位置)(例如右水平位置)相关联。

[0086] 第一下混合信号及第二下混合信号的联合编码表示360优选地包括第一下混合信号及第二下混合信号的下混合信号、第一下混合信号及第二下混合信号的公共残余信号及一个或多个预测参数。换言之,存在第一下混合信号312及第二下混合信号314下混频成的“公共”下混合信号,且存在可至少部分描述第一下混合信号312与第二下混合信号314之间的差异的“公共”残余信号。多声道解码器370优选地是基于预测的残余信号辅助的多声道解码器,例如,USAC复杂立体声预测解码器。换言之,提供第一下混合信号312及第二下混合信号314的多声道解码器370可实质上与提供第一残余信号332及第二残余信号334的多声道解码器330相同,使得以上解释及参考文献也适用。此外,应注意,第一下混合信号312优

选地与音频场景的第一水平位置或方位角位置(例如,左水平位置或方位角位置)相关联,且第二下混合信号314优选地与音频场景的第二水平位置或方位角位置(例如,右水平位置或方位角位置)相关联。因此,第一下混合信号312及第一残余信号332可与相同的第一水平位置或方位角位置(例如,左水平位置)相关联,且第二下混合信号314及第二残余信号334可与相同的第二水平位置或方位角位置(例如,右水平位置)相关联。因此,多声道解码器370及多声道解码器330两者可执行水平划分(或水平分离或水平分布)。

[0087] 残余信号辅助的多声道解码器340优选地可以是基于参数的,且可因此接收描述两个声道之间(例如,第一音频声道信号320与第二音频声道信号322之间)的所需相关性及/或该两个声道之间的阶差的一个或多个参数342。例如,残余信号辅助的多声道解码器340可基于具有残余信号扩展的MPEG环绕声编码(如例如ISO/IEC 23003-1:2007中所述),或“统一立体声解码”解码器(如例如ISO/IEC 23003-3,第7.11章(解码器)及附录B.21(编码器的描述以及术语“统一立体声”的定义)中所述)。因此,残余信号辅助的多声道解码器340可提供第一音频声道信号320及第二音频声道信号322,其中第一音频声道信号320及第二音频声道信号322与音频场景的垂直相邻的位置相关联。例如,第一音频声道信号可与音频场景的左下位置相关联,且第二音频声道信号可与音频场景的左上位置相关联(使得第一音频声道信号320及第二音频声道信号322例如与音频场景的相同水平位置或方位角位置相关联,或与相隔不超过30度的方位角位置相关联)。换言之,残余信号辅助的多声道解码器340可执行垂直划分(或分布,或分离)。

[0088] 残余信号辅助的多声道解码器350的功能可与残余信号辅助的多声道解码器340的功能相同,其中第三音频声道信号可例如与音频场景的右下位置相关联,且第四音频声道信号可例如与音频场景的右上位置相关联。换言之,第三音频声道信号及第四音频声道信号可与音频场景的垂直相邻的位置相关联,且可与音频场景的相同的水平位置或方位角位置相关联,其中残余信号辅助的多声道解码器350执行垂直划分(或分离,或分布)。

[0089] 总而言之,根据图3的音频解码器300执行分层音频解码,其中在第一阶段(多声道解码器330、多声道解码器370)中执行左右划分,且其中在第二阶段(残余信号辅助的多声道解码器340、350)中执行上下划分。此外,还使用联合编码表示310对残余信号332、334进行编码,而且(使用联合编码表示360)对下混合信号312、314进行编码。因而,将不同声道之间的相关性用于下混合信号312、314的编码(及解码)及残余信号332、334的编码(及解码)两者。因此,实现了高编码效率,且还利用了信号之间的相关性。

[0090] 4. 根据图4的音频编码器

[0091] 图4示出了根据本发明的另一实施例的音频编码器的示意框图。根据图4的音频编码器全部以400指定。音频编码器400被配置为接收四个音频声道信号,即第一音频声道信号410、第二音频声道信号412、第三音频声道信号414及第四音频声道信号416。此外,音频编码器400被配置为基于音频声道信号410、412、414及416提供已编码表示,其中该已编码表示包括两个下混合信号的联合编码表示420,以及公共带宽扩展参数的第一集合422及公共带宽扩展参数的第二集合424的已编码表示。音频编码器400包括第一带宽扩展参数提取器430,该第一带宽扩展参数提取器被配置为基于第一音频声道信号410及第三音频声道信号414获得公共带宽提取参数的第一集合422。音频编码器400还包括第二带宽扩展参数提取器440,该第二带宽扩展参数提取器被配置为基于第二音频声道信号412及第四音频声道

信号416获得公共带宽扩展参数的第二集合424。

[0092] 此外,音频编码器400包括(第一)多声道编码器450,该(第一)多声道编码器被配置为使用多声道编码对至少第一音频声道信号410及第二音频声道信号412进行联合编码,以获得第一下混合信号452。此外,音频编码器400还包括(第二)多声道编码器460,该(第二)多声道编码器被配置为使用多声道编码对至少第三音频声道信号414及第四音频声道信号416进行联合编码,以获得第二下混合信号462。此外,音频编码器400还包括(第三)多声道编码器470,该(第三)多声道编码器被配置为使用多声道编码第一下混合信号452及第二下混合信号462进行联合编码,以获得下混合信号的联合编码表示420。

[0093] 关于音频编码器400的功能,应注意,音频编码器400执行分层多声道编码,其中第一音频声道信号410及第二音频声道信号412在第一阶段中组合,且第三音频声道信号414及第四音频声道信号416也在第一阶段中组合,以藉此获得第一下混合信号452及第二下混合信号462。然后在第二阶段中对第一下混合信号452及第二下混合信号462进行联合编码。然而,应注意,第一带宽扩展参数提取器430基于在分层多声道编码的第一阶段中由不同的多声道编码器450、460处理的音频声道信号410、414来提供公共带宽提取参数的第一集合422。类似地,第二带宽扩展参数提取器440基于在第一处理阶段中由不同的多声道编码器450、460处理的不同音频声道信号412、416来提供公共带宽提取参数的第二集合424。此特定的处理顺序带来以下优点:该带宽扩展参数的集合422、424基于仅在分层编码的第二阶段中(即,在多声道编码器470中)组合的声道。这是有利的,因为在分层编码的第一阶段中组合这种音频声道是所希望的,该音频声道的关系关于声源位置知觉并非极其相关的。相反,第一下混合信号与第二下混合信号之间的关系主要决定声源位置知觉是值得推荐的,因为与相应音频声道信号410、412、414、416之间的关系相比,第一下混合信号452与第二下混合信号462之间的关系可更好维持。换言之,已发现,希望公共带宽扩展参数的第一集合422基于对下混合信号452、462的差异作出贡献的两个音频声道(音频声道信号),且公共带宽扩展参数的第二集合424是基于还对下混合信号452、462的差异作出贡献的音频声道信号412、416来提供的,这是由上述分层多声道编码中的音频声道信号的处理来实现的。因此,当与第一下混合信号452和第二下混合信号462之间的声道关系相比时,公共带宽扩展参数的第一集合422基于类似的声道关系,其中第一下混合信号与第二下混合信号之间的声道关系通常在音频解码器侧产生的空间印象中占据优势。因此,带宽扩展参数的第一集合422的提供以及带宽扩展参数的第二集合424的提供极其适于音频解码器侧产生的空间听觉印象。

[0094] 5. 根据图5的音频解码器

[0095] 图5示出了根据本发明的另一实施例的音频解码器的示意框图。根据图5的音频解码器全部以500指定。

[0096] 音频解码器500被配置为接收第一下混合信号及第二下混合信号的联合编码表示510。此外,音频解码器500被配置为提供第一带宽扩展的声道信号520、第二带宽扩展的声道信号522、第三带宽扩展的声道信号524及第四带宽扩展的声道信号526。

[0097] 音频解码器500包括(第一)多声道解码器530,该(第一)多声道解码器被配置为使用多声道解码,基于第一下混合信号及第二下混合信号的联合编码表示510来提供第一下混合信号532及第二下混合信号534。音频解码器500还包括(第二)多声道解码器540,该(第

二)多声道解码器被配置为使用多声道解码,基于第一下混合信号532来提供至少第一音频声道信号542及第二音频声道信号544。音频解码器500还包括(第三)多声道解码器550,该(第三)多声道解码器被配置为使用多声道解码,基于第二下混合信号544来提供至少第三音频声道信号556及第四音频声道信号558。此外,音频解码器500包括(第一)多声道带宽扩展560,该(第一)多声道带宽扩展被配置为基于第一音频声道信号542及第三音频声道信号556执行多声道带宽扩展,以获得第一带宽扩展的声道信号520及第三带宽扩展的声道信号524。此外,音频解码器包括(第二)多声道带宽扩展570,该(第二)多声道带宽扩展被配置为基于第二音频声道信号544及第四音频声道信号558执行多声道带宽扩展,以获得第二带宽扩展的声道信号522及第四带宽扩展的声道信号526。

[0098] 关于音频解码器500的功能,应注意,音频解码器500执行分层多声道解码,其中第一下混合信号532与第二下混合信号534之间的划分在分层解码的第一阶段中执行,且在分层解码的第二阶段中从第一下混合信号532导出第一音频声道信号542及第二音频声道信号544,且在分层解码的第二阶段中从第二下混合信号550导出第三音频声道信号556及第四音频声道信号558。然而,第一多声道带宽扩展560及第二多声道带宽扩展570两者各自接收从第一下混合信号532导出的一个音频声道信号,及从第二下混合信号534导出的一个音频声道信号。因为较好的声道分离通常由(第一)多声道解码530实现(作为分层多声道解码的第一阶段执行),当与分层解码的第二阶段相比时,可看出每一多声道带宽扩展560、570接收被很好地分离的输入信号(因为输入信号源自于很好地声道分离的第一下混合信号532及第二下混合信号534)。因而,多声道带宽扩展560、570可考虑立体声特性,该立体声特性对于听觉印象是重要的,且该立体声特性由第一下混合信号532与第二下混合信号534之间的关系很好地表示,且该多声道带宽扩展可因此提供良好的听觉印象。

[0099] 换言之,音频解码器的“交叉”结构考虑到良好的多声道带宽扩展,这考虑了声道之间的立体声关系,其中,多声道带宽扩展阶段560、570中每一个从(第二阶段)多声道解码器540、550两者接收输入信号。

[0100] 然而,应注意,音频解码器500可由本文关于根据图2、图3、根据6及图13的音频解码器所述的特征及功能中的任一项来补充,其中有可能将相应特征引入音频解码器500中以逐步改进音频解码器的性能。

[0101] 6. 根据图6A和图6B的音频解码器

[0102] 图6A和图6B示出了根据本发明的另一实施例的音频解码器的示意框图。根据图6A和图6B的音频解码器全部以600指定。根据图6A和图6B的音频解码器600类似于根据图5的音频解码器500,使得以上解释也适用。然而,音频解码器600已由还可单独地或通过组合方式引入至音频解码器500中以用于改进的一些特征及功能补充。

[0103] 音频解码器600被配置为接收第一下混合信号及第二下混合信号的联合编码表示610,且提供第一带宽扩展的信号620、第二带宽扩展的信号622、第三带宽扩展的信号624及第四带宽扩展的信号626。音频解码器600包括多声道解码器630,该多声道解码器被配置为接收第一下混合信号及第二下混合信号的联合编码表示610,且基于该联合编码表示来提供第一下混合信号632及第二下混合信号634。音频解码器600另一包括多声道解码器640,该多声道解码器被配置为接收第一下混合信号632,且基于该第一下混合信号来提供第一音频声道信号542及第二音频声道信号544。音频解码器600还包括多声道解码器650,该多

声道解码器被配置为接收第二下混合信号634,且提供第三音频声道信号656及第四音频声道信号658。音频解码器600还包括(第一)多声道带宽扩展660,该(第一)多声道带宽扩展被配置为接收第一音频声道信号642及第三音频声道信号656,且基于该第一音频声道信号及该第三音频声道信号来提供第一带宽扩展的声道信号620及第三带宽扩展的声道信号624。此外,(第二)多声道带宽扩展670接收第二音频声道信号644及第四音频声道信号658,且基于该第二音频声道信号及该第四音频声道信号来提供第二带宽扩展的声道信号622及第四带宽扩展的声道信号626。

[0104] 音频解码器600还包括另一多声道解码器680,该另一多声道解码器被配置为接收第一残余信号及第二残余信号的联合编码表示682,且该另一多声道解码器基于该联合编码表示来提供用于由多声道解码器640使用的的第一残余信号684及用于由多声道解码器650使用的第二残余信号686。

[0105] 多声道解码器630优选地是基于预测的残余信号辅助的多声道解码器。例如,多声道解码器630可实质上与以上所述的多声道解码器370相同。例如,多声道解码器630可以是如以上所述且如以上引用的USAC标准中所述的USAC复杂立体声预测解码器。因此,第一下混合信号及第二下混合信号的联合编码表示610可例如包括第一下混合信号及第二下混合信号的(公共)下混合信号、第一下混合信号及第二下混合信号的(公共)残余信号,及一个或多个预测参数,该一个或多个预测参数由多声道解码器630估计。

[0106] 此外,应注意,第一下混合信号632可例如与音频场景的第一水平位置或方位角位置(例如,左水平位置)相关联,且第二下混合信号634可例如与音频场景的第二水平位置或方位角位置(例如,右水平位置)相关联。

[0107] 此外,多声道解码器680可例如是基于预测的残余信号相关联的多声道解码器。多声道解码器680可实质上与以上所述多声道解码器330相同。例如,多声道解码器680可以是USAC复杂立体声预测解码器,如以上所提及。因此,第一残余信号及第二残余信号的联合编码表示682可包括第一残余信号及第二残余信号的(公共)下混合信号、第一残余信号及第二残余信号的(公共)残余信号,及一个或多个预测参数,该一个或多个预测参数由多声道解码器680估计。此外,应注意,第一残余信号684可与音频场景的第一水平位置或方位角位置(例如,左水平位置)相关联,且第二残余信号686可与音频场景的第二水平位置或方位角位置(例如,右水平位置)相关联。

[0108] 多声道解码器640可例如是基于参数的多声道解码,类似如以上所述且如所引用的标准中所述的例如MPEG环绕声多声道解码。然而,在存在(可选的)多声道解码器680及(可选的)第一残余信号684的情况下,多声道解码器640可以是基于参数的、残余信号辅助的多声道解码器,类似例如统一立体声解码器。因而,多声道解码器640可实质上与以上所述的多声道解码器340相同,且多声道解码器640可例如接收以上所述的参数342。

[0109] 类似地,多声道解码器650可实质上与多声道解码器640相同。因此,多声道解码器650可例如是基于参数的,且可选地是残余信号辅助的(在存在可选的多声道解码器680的情况下)。

[0110] 此外,应注意,第一音频声道信号642及第二音频声道信号644优选地与音频场景的垂直相邻的空间位置相关联。例如,第一音频声道信号642与音频场景的左下位置相关联,且第二音频声道信号644与音频场景的左上位置相关联。因此,多声道解码器640执行由

第一下混合信号632(且,可选地,由第一残余信号684)描述的音频内容的垂直划分(或分离,或分布)。类似地,第三音频声道信号656及第四音频声道信号658与音频场景的垂直相邻的位置相关联,且优选地与音频场景的相同水平位置或方位角位置相关联。例如,第三音频声道信号656优选地与音频场景的右下位置相关联,且第四音频声道信号658优选地与音频场景的右上位置相关联。因而,多声道解码器650执行由第二下混合信号634(且,可选地,由第二残余信号686)描述的音频内容的垂直划分(或分离,或分布)。

[0111] 然而,第一多声道带宽扩展660接收第一音频声道信号642及第三音频声道656,该第一音频声道信号及该第三音频声道与音频场景的左下位置及右下位置相关联。因此,第一多声道带宽扩展660基于与音频场景的相同水平面(例如,下水平面)或高度以及音频场景的不同侧(左/右)相关联的两个音频声道信号来执行多声道带宽扩展。因此,当执行带宽扩展时,多声道带宽扩展可考虑立体声特性(例如,人类立体声知觉)。类似地,第二多声道带宽扩展670还可考虑立体声特性,因为第二多声道带宽扩展对音频场景的相同水平面(例如,上水平面)或高度但在不同水平位置(不同侧)(左/右)处的音频声道信号进行操作。

[0112] 进一步总结,分层音频解码器600包括以下结构:在第一阶段(多声道解码630、680)中执行左/右划分(或分离,或分布),在第二阶段(多声道解码640、650)中执行垂直划分(分离或分布),且多声道带宽扩展对一对左/右信号进行操作(多声道带宽扩展660、670)。解码路径的此“交叉”允许可在分层音频解码器的第一处理阶段中执行对于听觉印象尤其重要(例如,比上/下划分更重要)的左/右分离,且还可对一对左右音频声道信号执行多声道带宽扩展,此举又导致尤其良好的听觉印象。上/下划分是作为左右分离与多声道带宽扩展之间的中间阶段来执行,这使得可导出四个音频声道信号(或带宽扩展的声道信号),而不显著地降级听觉印象。

[0113] 7. 根据图7的方法

[0114] 图7示出了用于基于至少四个音频声道信号来提供已编码表示的方法700的流程图。

[0115] 方法700包括使用残余信号辅助的多声道编码来对至少第一音频声道信号及第二音频声道信号进行联合编码710,以获得第一下混合信号及第一残余信号。方法还包括使用残余信号辅助的多声道编码来对至少第三音频声道信号及第四音频声道信号进行联合编码720,以获得第二下混合信号及第二残余信号。方法还包括使用多声道编码来对第一残余信号及第二残余信号进行联合编码730,以获得残余信号的已编码表示。然而,应注意,方法700可由本文中关于音频编码器及音频解码器所述的特征及功能中的任一项来补充。

[0116] 8. 根据图8的方法

[0117] 图8示出了用于基于已编码表示来提供至少四个音频声道信号的方法800的流程图。

[0118] 方法800包括使用多声道解码,基于第一残余信号及第二残余信号的联合编码表示来提供810第一残余信号及第二残余信号。方法800还包括使用残余信号辅助的多声道解码,基于第一下混合信号及第一残余信号来提供820第一音频声道信号及第二音频声道信号。方法还包括使用残余信号辅助的多声道解码,基于第二下混合信号及第二残余信号来提供830第三音频声道信号及第四音频声道信号。

[0119] 此外,应注意,方法800可由本文中关于音频解码器及音频编码器所述的特征及功

能中的任一项来补充。

[0120] 9. 根据图9的方法

[0121] 图9示出了用于基于至少四个音频声道信号来提供已编码表示的方法900的流程图。

[0122] 方法900包括基于第一音频声道信号及第三音频声道信号来获得910公共带宽扩展参数的第一集合。方法900还包括基于第二音频声道信号及第四音频声道信号来获得920公共带宽扩展参数的第二集合。方法还包括使用多声道编码来对至少第一音频声道信号及第二音频声道信号进行联合编码,以获得第一下混合信号,且使用多声道编码来对至少第三音频声道信号及第四音频声道信号进行联合编码940,以获得第二下混合信号。方法还包括使用多声道编码来对第一下混合信号及第二下混合信号进行联合编码950,以获得该下混合信号的已编码表示。

[0123] 应注意,可以通过任意顺序或并行地执行方法900的不包括特定互相依从性的步骤中的一些。此外,应注意,方法900可由本文中关于音频编码器及音频解码器所述的特征及功能中的任一项来补充。

[0124] 10. 根据图10的方法

[0125] 图10示出了用于基于已编码表示来提供至少四个音频声道信号的方法1000的流程图。

[0126] 方法1000包括:使用多声道解码,基于第一下混合信号及第二下混合信号的联合编码表示来提供1010第一下混合信号及第二下混合信号;使用多声道解码,基于第一下混合信号来提供1020至少第一音频声道信号及第二音频声道信号;使用多声道解码,基于第二下混合信号来提供1030至少第三音频声道信号及第四音频声道信号;基于第一音频声道信号及第三音频声道信号来执行1040多声道带宽扩展,以获得第一带宽扩展的声道信号及第三带宽扩展的声道信号;以及基于第二音频声道信号及第四音频声道信号来执行1050多声道带宽扩展,以获得第二带宽扩展的声道信号及第四带宽扩展的声道信号。

[0127] 应注意,可以通过任意顺序或并行地执行方法1000的步骤中的一些。此外,应注意,方法1000可由本文中关于音频编码器及音频解码器所述的特征及功能中的任一项来补充。

[0128] 11. 根据图11、图12及图13的实施例

[0129] 在下文中,将描述根据本发明的一些附加实施例及底层考虑。

[0130] 图11示出了根据本发明的实施例的音频编码器1100的示意框图。音频编码器1100被配置为接收左下声道信号1110、左上声道信号1112、右下声道信号1114及右上声道信号1116。

[0131] 音频编码器1100包括第一多声道音频编码器(或编码)1120,该第一多声道音频编码器(或编码)是MPEG环绕声2-1-2音频编码器(或编码)或统一立体声音频编码器(或编码),且该第一多声道音频编码器(或编码)接收左下声道信号1110及左上声道信号1112。第一多声道音频编码器1120提供左下混合信号1122及(可选地)左残余信号1124。此外,音频编码器1100包括第二多声道编码器(或编码)1130,该第二多声道编码器(或编码)是MPEG环绕声2-1-2编码器(或编码)或统一立体声编码器(或编码),该第二多声道编码器(或编码)接收右下声道信号1114及右上声道信号1116。第二多声道音频编码器1130提供右下混

合信号1132及(可选地)右残余信号1134。音频编码器1100还包括立体声编码器(或编码)1140,该立体声编码器(或编码)接收左下混合信号1122及右下混合信号1132。此外,作为复杂预测立体声编码的第一立体声编码1140从心理声学模型接收心理声学模型信息1142。例如,心理模型信息1142可描述不同的频带或子频带、心理声学掩蔽效应等的心理声学相关性。立体声编码1140提供声道对单元(CPE)“下混频”,该声道对单元(CPE)“下混频”以1144指定并以联合编码形式描述左下混合信号1122及右下混合信号1132。此外,音频编码器1100可选地包括第二立体声编码器(或编码)1150,该第二立体声编码器(或编码)被配置为接收可选的左残余信号1124及可选的右残余信号1134,以及心理声学模型信息1142。作为复杂预测立体声编码的第二立体声编码1150被配置为提供声道对单元(CPE)“残余”,该声道对单元(CPE)“残余”以联合编码形式表示左残余信号1124及右残余信号1134。

[0132] 编码器1100(以及本文所述其他音频编码器)基于通过分层地组合可用的USAC立体声工具来利用水平信号依从性及垂直信号依从性的思想(即,在USAC编码中可用的编码概念)。使用具有频带受限残余信号或全频带残余信号(以1124及1134指定)的MPEG环绕声2-1-2或统一立体声(以1120及1130指定)来组合垂直相邻的声道对。每一垂直声道对的输出是下混合信号1122、1132,且对于统一立体声是残余信号1124、1134。为了满足对双耳无掩蔽的知觉要求,通过使用MDCT域中的复杂预测(编码器1140)来对下混合信号1122、1132两者进行水平组合和联合编码,这包括左右编码及中侧编码的可能性。相同的方法可应用于水平组合的残余信号1124、1134。此概念在图11中示出。

[0133] 参考图11解释的分层结构可通过启用两个立体声工具(例如,两个USAC立体声工具)及在两者之间重新分拆声道来实现。因而,没有必需的附加预处理/后期处理步骤,且用于发送工具的有效载荷的比特流语法保持不变(例如,在与USAC标准相比时大体上不变)。此思想导致图12中所示的编码器结构。

[0134] 图12示出了根据本发明的实施例的音频编码器1200的示意框图。音频编码器1200被配置为接收第一声道信号1210、第二声道信号1212、第三声道信号1214及第四声道信号1216。音频编码器1200被配置为提供用于第一声道对单元的比特流1220以及用于第二声道对单元的比特流1222。

[0135] 音频编码器1200包括第一多声道编码器1230,该第一多声道编码器是MPEG环绕声2-1-2编码器或统一立体声编码器,且该第一多声道编码器接收第一声道信号1210及第二声道信号1212。此外,第一多声道编码器1230提供第一下混合信号1232、MPEG环绕声有效载荷1236及(可选地)第一残余信号1234。音频编码器1200还包括第二多声道编码器1240,该第二多声道编码器是MPEG环绕声2-1-2编码器或统一立体声编码器,且该第二多声道编码器接收第三声道信号1214及第四声道信号1216。第二多声道编码器1240提供第一下混合信号1242、MPEG环绕声有效载荷1246及(可选地)第二残余信号1244。

[0136] 音频编码器1200还包括第一立体声编码1250,该第一立体声编码是复杂预测立体声编码。第一立体声编码1250接收第一下混合信号1232及第二下混合信号1242。第一立体声编码1250提供第一下混合信号1232及第二下混合信号1242的联合编码表示1252,其中联合编码表示1252可包括(第一下混合信号1232及第二下混合信号1242的)(公共)下混合信号以及(第一下混合信号1232及第二下混合信号1242的)公共残余信号的表示。此外,(第一)复杂预测立体声编码1250提供复杂预测有效载荷1254,该复杂预测有效载荷通常包括

一个或多个复杂预测系数。此外,音频编码器1200还包括第二立体声编码1260,该第二立体声编码是复杂预测立体声编码。第二立体声编码1260接收第一残余信号1234及第二残余信号1244(或零输入值,如果不存在由多声道编码器1230、1240提供的残余信号)。第二立体声编码1260提供第一残余信号1234及第二残余信号1244的联合编码表示1262,该联合编码表示可例如包括(第一残余信号1234及第二残余信号1244的)(公共)下混合信号及(第一残余信号1234及第二残余信号1244的)公共残余信号。此外,复杂预测立体声编码1260提供复杂预测有效载荷1264,该复杂预测有效载荷通常包括一个或多个预测系数。

[0137] 此外,音频编码器1200包括心理声学模型1270,该心理声学模型提供控制第一复杂预测立体声编码1250及第二复杂预测立体声编码1260的信息。例如,由心理声学模型1270提供的信息可描述哪些频带或频格具有高的心理声学相关性且应以高精度编码。然而,应注意,使用心理声学模型1270提供的信息是可选的。

[0138] 此外,音频编码器1200包括第一编码器及复用器1280,该第一编码器及复用器从第一复杂预测立体声编码1250接收联合编码表示1252,从第一复杂预测立体声编码1250接收复杂预测有效载荷1254且从第一多声道音频编码器1230接收MPEG环绕声有效载荷1236。此外,第一编码及复用1280可从心理声学模型1270接收信息,该信息描述例如考虑心理声学掩蔽效应等,哪个编码精确度应该应用于哪些频带或子频带。因此,第一编码及复用1280提供第一声道对单元比特流1220。

[0139] 此外,音频编码器1200包括第二编码及复用1290,该第二编码及复用被配置为接收由第二复杂预测立体声编码1260提供的联合编码表示1262、由第二复杂预测立体声编码1260提供的复杂预测有效载荷1264及由第二多声道音频编码器1240提供的MPEG环绕声有效载荷1246。此外,第二编码及复用1290可从心理声学模型1270接收信息。因此,第二编码及复用1290提供第二声道对单元比特流1222。

[0140] 关于音频编码器1200的功能,参考以上解释,且还参考关于根据图2、图3、图5及图6A和图6B的音频编码器的解释。

[0141] 此外,应注意,此概念可扩展至将多个MPEG环绕声频格用于水平相关的声道、垂直相关的声道或其他几何相关的声道的联合编码以及将下混合信号及残余信号组合成复杂预测立体声对,考虑其几何学性质及知觉性质。这导致广义的解码器结构。

[0142] 在下文中,将描述四声道单元的实现。在三维音频编码系统中,使用用以形成四声道单元(QCE)的四个声道的分层组合。QCE由两个USAC声道对单元(CPE)组成(或提供两个USAC声道对单元,或接收两个USAC声道对单元)。使用MPS 2-1-2或统一立体声来组合垂直声道对。在第一声道对单元CPE中对下混频声道进行联合编码。如果应用残余编码,则在第二声道对单元CPE中对残余信号进行联合编码,否则将第二CPE中的信号设定为零。两个声道对单元CPE将复杂预测用于联合立体声编码,包括左右编码及中侧编码的可能性。为保留信号的高频率部分的知觉立体声性质,在应用SBR之前,通过附加的重新分拣步骤将立体声SBR(频谱带宽复制)应用于左上/右上声道对与左下/右下通路对之间。

[0143] 将参考图13描述可能的解码器结构,图13示出了根据本发明的实施例的音频解码器的示意框图。音频解码器1300被配置为接收表示第一声道对单元的第一比特流1310及表示第二声道对单元的第二比特流1312。然而,第一比特流1310及第二比特流1312可包括在公共的总比特流中。

[0144] 音频解码器1300被配置为提供第一带宽扩展的声道信号1320、第二带宽扩展的声道信号1322、第三带宽扩展的声道信号1324和第四带宽扩展的声道信号1326,第一带宽扩展的声道信号1320可例如表示音频场景的左下位置,第二带宽扩展的声道信号1322可例如表示音频场景的左上位置;第三带宽扩展的声道信号1324可例如与音频场景的右下位置相关联;以及第四带宽扩展的声道信号1326可例如与音频场景的右上位置相关联。

[0145] 音频解码器1300包括第一比特流解码1330,该第一比特流解码被配置为接收用于第一声道对单元的比特流1310,且基于该比特流来提供两个下混合信号的联合编码表示、复杂预测有效载荷1334、MPEG环绕声有效载荷1336及频谱带宽复制有效载荷1338。音频解码器1300还包括第一复杂预测立体声解码1340,该第一复杂预测立体声解码被配置为接收联合编码表示1332及复杂预测有效载荷1334,且基于该联合编码表示及该复杂预测有效载荷来提供第一下混合信号1342及第二下混合信号1344。类似地,音频解码器1300包括第二比特流解码1350,该第二比特流解码被配置为接收用于第二声道单元的比特流1312,且基于该比特流来提供两个残余信号的联合编码表示1352、复杂预测有效载荷1354、MPEG环绕声有效载荷1356及频谱带宽复制位负载1358。音频解码器还包括第二复杂预测立体声解码1360,该第二复杂预测立体声解码基于联合编码表示1352及复杂预测有效载荷1354来提供第一残余信号1362及第二残余信号1364。

[0146] 此外,音频解码器1300包括第一MPEG环绕声型多声道解码1370,该第一MPEG环绕声型多声道解码是MPEG环绕声2-1-2解码或统一立体声解码。第一MPEG环绕声型多声道解码1370接收第一下混合信号1342、第一残余信号1362(可选)及MPEG环绕声有效载荷1336,且基于该第一下混合信号、该第一残余信号及该MPEG环绕声有效载荷来提供第一音频声道信号1372及第二音频声道信号1374。音频解码器1300还包括第二MPEG环绕声型多声道解码1380,该第二MPEG环绕声型多声道解码是MPEG环绕声2-1-2多声道解码或统一立体声多声道解码。第二MPEG环绕声型多声道解码1380接收第二下混合信号1344及第二残余信号1364(可选),以及MPEG环绕声有效载荷1356,且基于该第二下混合信号、该第二残余信号及MPEG环绕声有效载荷来提供第三音频声道信号1382及第四音频声道信号1384。音频解码器1300还包括第一立体声频谱带宽复制1390,该第一立体声频谱带宽复制被配置为接收第一音频声道信号1372及第三音频声道信号1382,以及频谱带宽复制有效载荷1338,且基于该第一音频声道信号、该第三音频声道信号及该频谱带宽复制有效载荷来提供第一带宽扩展的声道信号1320及第三带宽扩展的声道信号1324。此外,音频解码器包括第二立体声频谱带宽复制1394,该第二立体声频谱带宽复制被配置为接收第二音频声道信号1374及第四音频声道信号1384,以及频谱带宽复制有效载荷1358,且基于该第二音频声道信号、该第四音频声道信号及该频谱带宽复制有效载荷来提供第二带宽扩展的声道信号1322及第四带宽扩展的声道信号1326。

[0147] 关于音频解码器1300的功能,参考以上论述,且还参考根据图2、图3、图5及图6A和图6B的音频解码器的论述。

[0148] 在下文中,将参考图14A及图14B来描述可用于本文所述音频编码/解码的比特流的示例。应注意,比特流可例如是统一语音及音频编码(USAC)中使用的比特流的扩展,该统一语音及音频编码(USAC)描述于以上提及的标准(ISO/IEC 23003-3:2012)中。例如,可将MPEG环绕声有效载荷1236、1246、1336、1356及复杂预测有效载荷1254、1264、1334、1354作

为传统声道对单元(即,对于根据USAC标准的声道对单元)发送。对于以信号方式发送四声道单元QCE的使用,USAC声道对配置可扩展两比特,如图14A中所示。换言之,可将以“qceIndex”指定的两个比特添加至USAC比特流单元“UsacChannelPairElementConfig()”。可例如如图14B的表格中所示地定义由比特“qceindex”表示的参数的意义。

[0149] 例如,形成QCE的两个声道对单元可作为连续单元发送,首先包含下混频声道及用于第一MPS框的MPS有效载荷的CPE,其次包含残余信号(或用于MPS 2-1-2编码的零音频信号)及用于第二MPS框的MPS有效载荷的CPE。

[0150] 换言之,当与用于发送四声道单元QCE的常规USAC比特流相比时,仅存在小的信令开销。

[0151] 然而,自然还可使用不同的比特流格式。

[0152] 12. 编码/解码环境

[0153] 在下文中,将描述可应用根据本发明的概念的音频编码/解码环境。

[0154] 可在其中使用根据本发明的概念的3D音频编解码器系统基于用于声道及对象信号的解码的MPEG-D USAC编解码器。为提高编码大量对象的效率,已适配了MPEG SAOC技术。三种类型的渲染器执行将对象渲染至声道、将声道渲染至耳机或将声道渲染至不同扬声器设置的任务。当显式地发送对象信号或使用SAOC参数化编码对象信号时,将对应的对象元数据信息经压缩且复用为3D音频比特流。

[0155] 图15示出了这种音频编码器的示意框图,以及图16示出了这种音频解码器的示意框图。换言之,图15及图16示出了3D音频系统的不同算法框。

[0156] 参考图15,现将解释一些细节,图15示出了3D音频编码器1500的示意框图。编码器1500包括可选的预渲染器/混合器1510,该可选的预渲染器/混合器接收一个或多个声道信号1512及一个或多个对象信号1514,且基于该一个或多个声道信号及该一个或多个对象信号来提供一个或多个声道信号1516及一个或多个对象信号1518、1520。音频编码器还包括USAC编码器1530及(可选地)SAOC编码器1540。SAOC编码器1540被配置为基于提供至SAOC编码器的一个或多个对象1520来提供一个或多个SAOC传送声道1542及SAOC边带信息1544。此外,USAC编码器1530被配置为从预渲染器/混合器接收包括声道及预渲染对象的声道信号1516,从预渲染器/混合器接收一个或多个对象信号1518且接收一个或多个SAOC传送声道1542及SAOC边带信息1544,且基于上述各项来提供已编码表示1532。此外,音频编码器1500还包括对象元数据编码器1550,该对象元数据编码器被配置为接收对象元数据1552(该对象元数据可由预渲染器/混合器1510估计)且对对象元数据编码以获得编码对象元数据1554。编码元数据还由USAC编码器1530接收,且用来提供已编码表示1532。

[0157] 以下将描述关于音频编码器1500的各个组件的一些细节。

[0158] 现在参考图16,将描述音频解码器1600。音频解码器1600被配置为接收已编码表示1610,且基于该已编码表示来提供多声道扬声器信号1612、耳机信号1614及/或替代格式(例如,5.1格式)的扬声器信号1616。

[0159] 音频解码器1600包括USAC解码器1620,且基于已编码表示1610来提供一个或多个声道信号1622、一个或多个预渲染对象信号1624、一个或多个对象信号1626、一个或多个SAOC传送声道1628、SAOC边带信息1630及压缩对象元数据信息1632。音频解码器1600还包括对象渲染器1640,该对象渲染器被配置为基于对象信号1626及对象元数据信息1644来提

供一个或多个渲染对象信号1642,其中,由对象元数据解码器1650基于压缩对象元数据信息1632提供对象元数据信息1644。音频解码器1600还包括(可选地)SAOC解码器1660,该SAOC解码器被配置为接收SAOC传送声道1628及SAOC边带信息1630,且基于该SAOC传送声道及该SAOC边带信息来提供一个或多个渲染对象信号1662。音频解码器1600还包括混合器1670,该混合器被配置为接收声道信号1622、预渲染对象信号1624、渲染对象信号1642及渲染对象信号1662,且基于上述各项来提供多个混合声道信号1672,该多个混合声道信号可例如构成多声道扬声器信号1612。音频解码器1600可例如还包括双耳渲染1680,该双耳渲染被配置为接收混合声道信号1672且基于该混合声道信号来提供耳机信号1614。此外,音频解码器1600可包括格式转换1690,该格式转换被配置为接收混合声道信号1672及重现布局信息1692,且基于该混合声道信号及该重现布局信息来为替代扬声器设置提供扬声器信号1616。

[0160] 在下文中,将描述关于音频编码器1500及音频解码器1600的组件的一些细节。

[0161] 预渲染器/混合器

[0162] 预渲染器/混合器1510可选地用于在编码之前将声道加对象输入场景转换成声道场景。在功能上,该预渲染器/混合器可与以下所述的对象渲染器/混合器相同。对象的预渲染可例如确保编码器输入处的确定信号熵,该确定信号熵基本上独立于同时有效的对象信号的数目。在对象的预渲染中,无需对象元数据发送。谨慎的(discreet)的对象信号被渲染至编码器所配置使用的声道布局。从相关联的对象元数据(OAM)1552获得针对每一声道的对象的权重。

[0163] USAC核心编解码器

[0164] 用于扬声器声道信号、谨慎的对象信号、对象下混合信号及预渲染信号的核心编解码器1530、1620基于MPEG-D USAC技术。通过基于输入声道及对象指派的几何学信息及语义信息来创建声道及对象映射信息,该核心编解码器处理大量信号的编码。该映射信息描述输入声道及对象如何映射至USAC声道单元(CPE、SCE、LFE)及对应的信息如何发送至解码器。所有附加有效载荷(如SAOC数据或对象元数据)已通过扩展单元且已在编码器速率控制中予以考虑。

[0165] 对象的编码可能以不同的方式,取决于对渲染器的速率/失真要求及交互性要求。以下对象编码变型为可能的:

[0166] 1. 预渲染对象:在编码之前将对象信号预渲染且混合为22.2声道信号。后续编码链参见22.2声道信号。

[0167] 2. 谨慎的对象波形式:将对象作为单音波形式供应至编码器。除声道信号外,编码器使用单声道单元SCE来传递对象。在接收器侧渲染且混合解码对象。压缩对象元数据信息沿侧发送至接收器/渲染器。

[0168] 3. 参数对象波形式:通过SAOC参数描述对象性质及其彼此的关系。使用USAC来编码对象信号的下混频。参数信息沿侧发送。取决于对象的数目及整体数据速率来选择下混频声道的数目。压缩对象元数据信息发送至SAOC渲染器。

[0169] SAOC

[0170] 用于对象信号的SAOC编码器1540及SAOC解码器1660基于MPEG SAOC技术。系统能够基于较小数目的发送声道及附加参数数据(对象阶差OLD、对象间相关性IOC、下混频增益

DMG) 来重新创建、修改且渲染许多音频对象。附加参数数据展现出比单独发送所有对象所需的数据速率显著降低的数据速率,使得编码极其有效。SAOC编码器将对象/声道信号(例如单音波形)作为输入,且输出参数信息(该参数信息被封装在3D音频比特流1532、1610中)及SAOC传送声道(使用单声道单元编码且发送)。

[0171] SAOC解码器1600根据解码的SAOC传送声道1628及参数信息1630重建对象/声道信号,且基于重现布局、解压的对象元数据信息以及可选地基于用户交互信息来产生输出音频场景。

[0172] 对象元数据编解码器

[0173] 对于每一对象,通过对象性质在时间和空间中的量化来有效地编码对对象在3D空间中的几何位置及容积进行规定的相关元数据。压缩的对象元数据cOAM 1554、1632作为边带信息发送至接收器。

[0174] 对象渲染器/混合器

[0175] 对象渲染器利用压缩的对象元数据来根据给定重现格式产生对象波形。每一对象根据其元数据渲染至某些输出声道。该框的输出来自于部分结果的和。如果对基于声道的内容及谨慎的对象/参数对象进行解码,则在输出所产生的波形之前(或在将该所产生的波形馈送至后期处理器模块(例如双耳渲染器或扬声器渲染器模块)之前),混合基于声道的波形及渲染对象波形经。

[0176] 双耳渲染器

[0177] 双耳渲染器模块1680产生多声道音频材料的双耳下混频,使得每一输入声道都由虚拟声源表示。在QMF域中按帧执行处理。双耳化基于测量的双耳空间脉冲响应。

[0178] 扬声器渲染器/格式转换

[0179] 扬声器渲染器1690在发送声道配置与所需重现格式之间转换。该扬声器渲染器因此在下文中被称为“格式转换器”。格式转换器执行至较低数目的输出声道的转换,即,该格式转换器创建下混频。系统自动产生针对输入格式及输出格式的给定组合的最优下混频矩阵,且在下混频处理中应用该矩阵。格式转换器考虑到标准扬声器配置且考虑到具有非标准扬声器位置的随机配置。

[0180] 图17示出了格式转换器的示意框图。如图可看出,格式转换器1700接收混合器输出信号1710,例如,混合声道信号1672,且提供扬声器信号1712,例如,扬声器信号1616。格式转换器包括下混频配置器1730和QMF域中的下混频处理1720,其中下混频配置器基于混合器输出布局信息1732及重现布局信息1734来提供用于下混频处理1720的配置信息。

[0181] 此外,应注意,以上所述概念,例如音频编码器100、音频解码器200或300、音频编码器400、音频解码器500或600、方法700、800、900或1000、音频编码器1100或1200及音频解码器1300可在音频编码器1500内及/或音频解码器1600内使用。例如,先前提及的音频编码器/解码器可用于与不同空间位置相关联的声道信号的编码或解码。

[0182] 13. 替代性实施例

[0183] 在下文中,将描述一些附加实施例。

[0184] 现参考图18至图21,将解释根据本发明的附加实施例。

[0185] 应注意,所谓的“四声道单元”(QCE)可被视为音频解码器的工具,该音频解码器可用于例如解码三维音频内容。

[0186] 换言之,四声道单元(QCE)是用于水平分布及垂直分布声道的更有效编码的四声道联合编码方法。QCE由两个连续CPE组成,且通过分层地组合在水平方向上具有复杂立体声预测工具的可能性且在垂直方向上具有基于MPEG环绕声的立体声工具的可能性的联合立体声工具来形成。这是通过启用两个立体声工具及在应用工具之间调换输出声道来实现的。在水平方向上执行立体声SBR来保留高频率的左右关系。

[0187] 图18示出了QCE的拓扑结构。应注意,图18的QCE极其类似于图11的QCE,使得可参考以上解释。然而,应注意,在图18的QCE中,在执行复杂立体声预测时并非必须使用心理学模型(可选地,虽然这种使用当然时可能的)。此外,可看出,基于左下声道及右下声道来执行第一立体声频谱带宽复制(立体声SBR),且基于左上声道及右上声道来执行第二立体声频谱带宽复制(立体声SBR)。

[0188] 在下文中,将提供一些术语及定义,该术语及定义可应用于一些实施例中。

[0189] 数据单元qceIndex指示CPE的QCE模式。关于比特流变量qceIndex的意义,参考图14B。应注意,qceIndex描述UsacChannelPairElement()类型的两个后续单元是否被当作四声道单元(QCE)。在图14B中给出不同的QCE模式。qceIndex对于形成一个QCE的两个后续单元而言应该相同。

[0190] 在下文中,将定义一些帮助单元,该帮助单元可在根据本发明的一些实现中使用:

[0191] cplx\_out\_dm\_x\_L[] 复杂预测立体声解码之后的第一CPE的第一声道

[0192] cplx\_out\_dm\_x\_R[] 复杂预测立体声解码之后的第一CPE的第二声道

[0193] cplx\_out\_res\_L[] 复杂预测立体声解码之后的第二CPE(如果qceIndex=1,则为零)

[0194] cplx\_out\_res\_R[] 复杂预测立体声解码之后的第二CPE的第二声道(如果qceIndex=1,则为零)

[0195] mps\_out\_L\_1[] 第一MPS框的第一输出声道

[0196] mps\_out\_L\_2[] 第一MPS框的第二输出声道

[0197] mps\_out\_R\_1[] 第二MPS框的第一输出声道

[0198] mps\_out\_R\_2[] 第二MPS框的第二输出声道

[0199] sbr\_out\_L\_1[] 第一立体声SBR框的第一输出声道

[0200] sbr\_out\_R\_1[] 第一立体声SBR框的第二输出声道

[0201] sbr\_out\_L\_2[] 第二立体声SBR框的第一输出声道

[0202] sbr\_out\_R\_2[] 第二立体声SBR框的第二输出声道

[0203] 在下文中,将解释在根据本发明的实施例中执行的解码处理。

[0204] UsacChannelPairElementConfig()中的语法单元(或比特流单元,或数据单元)qceIndex指示CPE是否属于QCE且是否使用残余编码。在qceIndex不等于0的情况下,当前CPE与其后续单元一起形成QCE,该后续单元应该是具有相同qceIndex的CPE。立体声SBR始终用于QCE,因而语法项stereoConfigIndex应为3且bsStereoSbr应为1。

[0205] 在qceIndex == 1的情况下,仅用于MPEG环绕声及SBR的有效载荷且无相关音频信号数据包含在第二CPE中,且语法单元bsResidualCoding设定为0。

[0206] 由qceIndex == 2指示存在第二CPE中残余信号。在此情况下,语法单元bsResidualCoding设定为1。

[0207] 然而,还可使用一些不同的且可能简化的信号传输方案。

[0208] 如ISO/IEC 23003-3第7.7小节中所述地执行具有复杂立体声预测的可能性的联合立体声的解码。第一CPE的所产生的输出是MPS下混合信号cplx\_out\_dmx\_L[ ]及cplx\_out\_dmx\_R[ ]。如果使用残余编码(还即,qceIndex == 2),则第二CPE的输出是MPS残余信号cplx\_out\_res\_L[ ]、cplx\_out\_res\_R[ ],若无残余信号已发送(即,qceIndex == 1),则插入零信号。

[0209] 在应用MPEG环绕声解码之前,调换第一组件(cplx\_out\_dmx\_R[ ])的第二声道和第二组件(cplx\_out\_res\_L[ ])的第一声道。

[0210] 如ISO/IEC 23003-3第7.11小节中所述地执行MPEG环绕声的解码。如果使用残余编码,然而在一些实施例中,与常规的MPEG环绕声解码相比,可修改解码。如ISO/IEC 23003-3第7.11.2.7小节(图23)中所定义的使用SBR的无残余MPEG环绕声的解码来进行修改,以使立体声SBR还用于bsResidualCoding == 1,从而导致图19中所示的解码器示意图。图19示出了用于bsResidualCoding == 0且bsStereoSbr == 1的音频编码器的示意框图。

[0211] 如图19中可看出,USAC核心解码器2010将下混合信号(DMX)2012提供至MPS(MPEG环绕声)解码器2020,该MPS(MPEG环绕声)解码器提供第一解码音频信号2022及第二解码音频信号2024。立体声SBR解码器2030接收第一解码音频信号2022及第二解码音频信号2024,且基于该第一解码音频信号及该第二解码音频信号来提供左带宽扩展的音频信号2032及右带宽扩展的音频信号2034。

[0212] 在应用立体声SBR之前,对第一组件(mps\_out\_L\_2[ ])的第二声道及第二组件(mps\_out\_R\_1[ ])的第一声道进行调换以允许左右立体声SBR。在立体声SBR的应用之后,对第一组件(sbr\_out\_R\_1[ ])的第二输出声道及第二组件(sbr\_out\_L\_2[ ])的第一声道再次进行调换,以恢复输入声道顺序。

[0213] 在图20中例示出QCE解码器结构,图20示出了QCE解码器示意图。

[0214] 应注意,图20的示意框图极其类似于图13的示意框图,使得还可参考以上解释。此外,应注意,在图20中已添加一些信号标示,其中,参考本部分中的定义。此外,示出了声道的最终重新分拣,该最终重新分拣是在立体声SBR之后执行。

[0215] 图21示出了根据本发明的实施例的四声道编码器2200的示意框图。换言之,在图21中例示出可被视为核心编码器工具的四声道编码器(四声道单元)。

[0216] 四声道编码器2200包括第一立体声SBR 2210,该第一立体声SBR接收第一左声道输入信号2212及第二左声道输入信号2214,且该第一立体声SBR基于该第一左声道输入信号及该第二左声道输入信号来提供第一-SBR有效载荷2215、第一左声道SBR输出信号2216及第一右声道SBR输出信号2218。此外,四声道编码器2200包括第二立体声SBR,该第二立体声SBR接收第二左声道输入信号2222及第二右声道输入信号2224,且该第二立体声SBR基于该第二左声道输入信号及该第二右声道输入信号来提供第一-SBR有效载荷2225、第一左声道SBR输出信号2226及第一右声道SBR输出信号2228。

[0217] 四声道编码器2200包括第一MPEG环绕声型(MPS 2-1-2或统一立体声)多声道编码器2230,该第一MPEG环绕声型(MPS 2-1-2或统一立体声)多声道编码器接收第一左声道SBR输出信号2216及第二左声道SBR输出信号2226,且该第一MPEG环绕声型(MPS 2-1-2或统一立体声)多声道编码器基于该第一左声道SBR输出信号及该第二左声道SBR输出信号来提供

第一MPS有效载荷2232、左声道MPEG环绕声下混合信号2234及(可选地)左声道MPEG环绕声残余信号2236。四声道编码器2200还包括第二MPEG环绕声型(MPS 2-1-2或统一立体声)多声道编码器2240,该第二MPEG环绕声型(MPS 2-1-2或统一立体声)多声道编码器接收第一右声道SBR输出信号2218及第二右声道SBR输出信号2228,且该第二MPEG环绕声型(MPS 2-1-2或统一立体声)多声道编码器基于该第一右声道SBR输出信号及该第二右声道SBR输出信号来提供第一MPS有效载荷2242、右声道MPEG环绕声下混合信号2244及(可选地)右声道MPEG环绕声残余信号2246。

[0218] 四声道编码器2200包括第一复杂预测立体声编码2250,该第一复杂预测立体声编码接收左声道MPEG环绕声下混合信号2234及右声道MPEG环绕声下混合信号2244,且该第一复杂预测立体声编码基于该左声道MPEG环绕声下混合信号及该右声道MPEG环绕声下混合信号来提供复杂预测有效载荷2252以及左声道MPEG环绕声下混合信号2234和右声道MPEG环绕声下混合信号2244的联合编码表示2254。四声道编码器2200包括第二复杂预测立体声编码2260,该第二复杂预测立体声编码接收左声道MPEG环绕声残余信号2236及右声道MPEG环绕声残余信号2246,该第二复杂预测立体声编码基于该左声道MPEG环绕声残余信号及该右声道MPEG环绕声残余信号来提供复杂预测有效载荷2262以及左声道MPEG环绕声下混合信号2236和右声道MPEG环绕声下混合信号2246的联合编码表示2264。

[0219] 四声道编码器还包括第一比特流编码2270,该第一比特流编码接收联合编码表示2254、复杂预测有效载荷2252、MPS有效载荷2232及SBR有效载荷2215,且基于以上各项来提供表示第一声道对单元的比特流部分。四声道编码器还包括第二比特流编码2280,该第二比特流编码接收联合编码表示2264、复杂预测有效载荷2262、MPS有效载荷2242及SBR有效载荷2225,且基于以上各项来提供表示第一声道对单元的比特流部分。

[0220] 14. 实现方案的备选

[0221] 虽然在设备的上下文中已描述一些方案,但是明显地,这些方案还表示对应的方法的描述,其中框或装置对应于方法步骤或方法步骤的特征。类似地,在方法步骤的上下文中,所述的方案还表示对应的设备的对应的框或项或特征的描述。方法步骤中的一些或全部可由(使用)硬件设备来执行,该硬件设备如例如微处理器、可编程计算机或电子电路。在一些实施例中,最重要的方法步骤中的某一个或多个步骤可由此设备来执行。

[0222] 发明性编码音频信号可储存在数字储存介质上,或可经由诸如无线传输介质或有线传输介质的传输介质来发送,该传输介质诸如因特网。

[0223] 取决于某些实现要求,本发明的实施例可实现在硬件中或软件中。可使用数字储存介质来执行实现,该数字储存介质例如软盘、DVD、蓝光、CD、ROM、PROM、EPROM、EEPROM或闪存,该数字储存介质上储存有电子可读的控制信号,该电子可读的控制信号与可编程计算机系统合作(或能够与可编程计算机系统合作),使得可执行相应方法。因此,数字储存介质可以是计算机可读的。

[0224] 根据本发明的一些实施例,包括具有电子可读的控制信号的数据载体,该电子可读的控制信号能够与可编程计算机系统合作,使得可执行本文所述方法之一。

[0225] 通常,本发明的实施例可实行为具有程序代码的计算机程序产品,当计算机程序产品在计算机上执行时,该程序代码可操作于执行方法之一。程序代码可例如储存在机器可读载体上。

[0226] 其他实施例包括用于执行本文所述方法之一的计算机程序,该计算机程序储存在机器可读载体上。

[0227] 换言之,发明性方法的实施例因此是具有程序代码的计算机程序,当在计算机上执行计算机程序时,所述程序代码用于执行本文所述方法之一。

[0228] 发明性方法的另一实施例因此是数据载体(或数字储存介质,或计算机可读介质),该数据载体包括记录在该数据载体上的用于执行本文所述方法之一的计算机程序。数据载体、数字储存介质或记录介质通常是有形的和/或非暂时性的。

[0229] 发明性方法的另一实施例因此是表示用于执行本文所述方法之一的计算机程序的数据流或信号序列。数据流或信号序列可例如被配置为经由数据通信连接(例如经由因特网)传递。

[0230] 另一实施例包括处理装置,例如计算机或可编程逻辑设备,该处理装置被配置或适配为执行本文所述方法之一。

[0231] 另一实施例包括计算机,该计算机上安装有用于执行本文所述方法之一的计算机程序。

[0232] 根据本发明的另一实施例包括设备或系统,该设备或系统被配置为将用于执行本文所述方法之一的计算机程序传递(例如,电子地或光学地)至接收器。接收器可例如是计算机、移动设备、存储设备等。设备或系统可例如包括用于将计算机程序传递至接收器的文件服务器。

[0233] 在一些实施例中,可编程逻辑设备(例如现场可编程门阵列)可用来执行本文所述方法的功能中的一些或全部。在一些实施例中,现场可编程门阵列可与微处理器合作,以执行本文所述方法之一。通常,方法优选地由任何硬件设备执行。

[0234] 以上所述实施例对于本发明的原理仅是示意性的。将理解,本领域技术人员将显而易见本文所述布置及细节的修改及变化。因此,意图是仅受即将出现的专利权利要求的范围而不是通过本文实施例的描述及解释的方式呈现的特定细节来限制。

[0235] 15. 结论

[0236] 在下文中,将提供一些结论。

[0237] 根据本发明的实施例基于以下考虑:为说明垂直分布的声道与水平分布的声道之间的信号依从性,可通过分层地组合联合立体声编码工具来对四个声道进行联合编码。例如,使用具有频带受限残余编码或全频带残余编码的MPS 2-1-2及/或统一立体声来组合垂直声道对。为了满足对双耳无掩蔽的知觉要求,例如通过在MDCT域中使用复杂预测来对输出下混频进行联合编码,这包括左右编码及中侧编码的可能性。如果残余信号存在,则使用相同方法来水平地组合该残余信号。

[0238] 此外,应注意,根据本发明的实施例克服先前技术的缺点中的一些或全部。根据本发明的实施例适于3D音频情境,其中扬声器声道分布在如果干高度的层中,从而导致水平声道对及垂直声道对。已发现,如USAC中定义的仅两个声道的联合编码不足以考虑声道之间的空间关系及知觉关系。然而,根据本发明的实施例克服了该问题。

[0239] 此外,在附加预处理/后期处理步骤中应用常规的MPEG环绕声,使得在无联合立体声编码的可能性的情况下单独发送残余信号,例如,以探索左基础音残余信号与右基础音残余信号之间的依从性。相反,根据本发明的实施例考虑到通过利用这种依从性进行的有

效编码/解码。

[0240] 进一步总结,根据本发明的实施例创建如本文所述的用于编码及解码的设备、方法或计算机程序。

[0241] 16. 补充注释

[0242] 下面将描述前述实施例的示例作为补充注释。然而,前述示例性实施例中举例说明的本发明和对其进行的修改不限于以下内容。

[0243] 示例1. 一种音频解码器(200;300;600;1300;1600;2000),用于基于已编码表示(210;310;360;610;682;1310、1312;1610)提供至少四个音频声道信号(220、222、224、226;320、322、324、326;620、622、624、626;1320、1322、1324、1326),

[0244] 其中,所述音频解码器被配置为使用多声道解码(230;330;680;1360),基于第一残余信号及第二残余信号的联合编码表示(210;310;682;1312)来提供所述第一残余信号(232;332;684;1362)及所述第二残余信号(234;334;686;1364);

[0245] 其中,所述音频解码器被配置为使用残余信号辅助的多声道解码(240;340;640;1370),基于第一下混合信号(212;312;632;1342)及所述第一残余信号来提供第一音频声道信号(220;320;642;1372)及第二音频声道信号(222;322;644;1374);以及

[0246] 其中,所述音频解码器被配置为使用残余信号辅助的多声道解码(250;350;650;1380),基于第二下混合信号(214;314;634;1344)及所述第二残余信号来提供第三音频声道信号(224;324;656;1382)及第四音频声道信号(226;326;658;1384)。

[0247] 示例2. 根据示例1所述的音频解码器,其中,所述音频解码器被配置为:使用多声道解码(370;630;1340),基于所述第一下混合信号及所述第二下混合信号的联合编码表示(360;610;1310)来提供所述第一下混合信号(212;312;632;1342)及所述第二下混合信号(214;314;634;1344)。

[0248] 示例3. 根据示例1或2所述的音频解码器,其中,所述音频解码器被配置为:使用基于预测的多声道解码,基于所述第一残余信号及所述第二残余信号的联合编码表示来提供所述第一残余信号及所述第二残余信号。

[0249] 示例4. 根据示例1至3中任一项所述的音频解码器,其中,所述音频解码器被配置为:使用残余信号辅助的多声道解码,基于所述第一残余信号及所述第二残余信号的联合编码表示来提供所述第一残余信号及所述第二残余信号。

[0250] 示例5. 根据示例3所述的音频解码器,其中,所述基于预测的多声道解码被配置为:估计预测参数,所述预测参数描述使用先前帧的信号分量导出的信号分量对提供当前帧的残余信号的贡献。

[0251] 示例6. 根据示例3至5中任一项所述的音频解码器,其中,所述基于预测的多声道解码被配置为:基于所述第一残余信号及所述第二残余信号的下混合信号且基于所述第一残余信号及所述第二残余信号的公共残余信号来获得所述第一残余信号及所述第二残余信号。

[0252] 示例7. 根据示例6所述的音频解码器,其中,所述基于预测的多声道解码被配置为:以第一符号应用公共残余信号,以获得所述第一残余信号,且以第二符号应用公共残余信号,以获得所述第二残余信号,所述第二符号与所述第一符号相反。

[0253] 示例8. 根据示例1至7中任一项所述的音频解码器,其中,所述音频解码器被配置

为:使用在MDCT域中操作的多声道解码,基于所述第一残余信号及所述第二残余信号的联合编码表示来提供所述第一残余信号及所述第二残余信号。

[0254] 示例9. 根据示例1至8中任一项所述的音频解码器,其中,所述音频解码器被配置为:使用USAC复杂立体声预测,基于所述第一残余信号及所述第二残余信号的联合编码表示来提供所述第一残余信号及所述第二残余信号。

[0255] 示例10. 根据示例1至9中任一项所述的音频解码器,

[0256] 其中,所述音频解码器被配置为使用基于参数的、残余信号辅助的多声道解码,基于所述第一下混合信号及所述第一残余信号来提供所述第一音频声道信号及所述第二音频声道信号;以及

[0257] 其中,所述音频解码器被配置为使用基于参数的、残余信号辅助的多声道解码,基于所述第二下混合信号及所述第二残余信号来提供所述第三音频声道信号及所述第四音频声道信号。

[0258] 示例11. 根据示例10所述的音频解码器,其中,所述基于参数的、残余信号辅助的多声道解码被配置为:估计描述两个声道之间的所需相关性及/或两个声道之间的阶差的一个或多个参数,以基于所述下混合信号中相应的一个下混合信号以及所述残余信号中对应的一个残余信号来提供所述两个或两个以上音频声道信号。

[0259] 示例12. 根据示例1至11中任一项所述的音频解码器,其中,所述音频解码器被配置为:使用在QMF域中操作的、残余信号辅助的多声道解码,基于所述第一下混合信号及所述第一残余信号来提供所述第一音频声道信号及所述第二音频声道信号;以及

[0260] 所述音频解码器被配置为:使用在QMF域中操作的、残余信号辅助的多声道解码,基于所述第二下混合信号及所述第二残余信号来提供所述第三音频声道信号及所述第四音频声道信号。

[0261] 示例13. 根据示例1至12中任一项所述的音频解码器,其中,所述音频解码器被配置为:使用MPEG环绕声2-1-2解码或统一立体声解码,基于所述第一下混合信号及所述第一残余信号来提供所述第一音频声道信号及所述第二音频声道信号;以及

[0262] 所述音频解码器被配置为:使用MPEG环绕声2-1-2解码或统一立体声解码,基于所述第二下混合信号及所述第二残余信号来提供所述第三音频声道信号及所述第四音频声道信号。

[0263] 示例14. 根据示例1至13中任一项所述的音频解码器,其中,所述第一残余信号及所述第二残余信号与音频场景的不同水平位置或与所述音频场景的不同方位角位置相关联。

[0264] 示例15. 根据示例1至14中任一项所述的音频解码器,其中,所述第一音频声道信号及所述第二音频声道信号与音频场景的垂直相邻的位置相关联,以及

[0265] 所述第三音频声道信号及所述第四音频声道信号与所述音频场景的垂直相邻的位置相关联。

[0266] 示例16. 根据示例1至15中任一项所述的音频解码器,其中,所述第一音频声道信号及所述第二音频声道信号与音频场景的第一水平位置或方位角位置相关联,以及

[0267] 所述第三音频声道信号及所述第四音频声道信号与所述音频场景的第二水平位置或方位角位置相关联,所述第二水平位置或方位角位置不同于所述第一水平位置或所述

第一方位角位置。

[0268] 示例17. 根据示例1至16中任一项所述的音频解码器,其中,所述第一残余信号与音频场景的左侧相关联,且所述第二残余信号与音频场景的右侧相关联。

[0269] 示例18. 根据示例17所述的音频编码器,

[0270] 其中,所述第一音频声道信号及所述第二音频声道信号与所述音频场景的左侧相关联,以及

[0271] 所述第三音频声道信号及所述第四音频声道信号与所述音频场景的右侧相关联。

[0272] 示例19. 根据示例18所述的音频解码器,其中,所述第一音频声道信号与所述音频场景的左下位置相关联,

[0273] 所述第二音频声道信号与所述音频场景的左上位置相关联,

[0274] 所述第三音频声道信号与所述音频场景的右下位置相关联,以及

[0275] 所述第四音频声道信号与所述音频场景的右上位置相关联。

[0276] 示例20. 根据示例1至19中任一项所述的音频解码器,其中,所述音频解码器被配置为:使用多声道解码,基于所述第一下混合信号及所述第二下混合信号的联合编码表示来提供所述第一下混合信号及所述第二下混合信号,所述第一下混合信号与音频场景的左侧相关联,且所述第二下混合信号与所述音频场景的右侧相关联。

[0277] 示例21. 根据示例1至20中任一项所述的音频解码器,其中,所述音频解码器被配置为:使用基于预测的多声道解码,基于所述第一下混合信号及所述第二下混合信号的联合编码表示来提供所述第一下混合信号及所述第二下混合信号。

[0278] 示例22. 根据示例1至21中任一项所述的音频解码器,其中,所述音频解码器被配置为使用残余信号辅助的、基于预测的多声道解码,基于所述第一下混合信号及所述第二下混合信号的联合编码表示来提供所述第一下混合信号及所述第二下混合信号。

[0279] 示例23. 根据示例1至22中任一项所述的音频解码器,其中,所述音频解码器被配置为:基于所述第一音频声道信号及所述第三音频声道信号执行第一多声道带宽扩展(660;1390),以及

[0280] 所述音频解码器被配置为:基于所述第二音频声道信号及所述第四音频声道信号执行第二多声道带宽扩展(670;1394)。

[0281] 示例24. 根据示例23所述的音频解码器,其中,所述音频解码器被配置为:基于所述第一音频声道信号及所述第三音频声道信号以及一个或多个带宽扩展参数(1338)执行所述第一多声道带宽扩展,以获得与音频场景的第一公共水平面或第一公共高度相关联的两个或两个以上带宽扩展的音频声道信号(620、624;1320、1324),以及

[0282] 所述音频解码器被配置为:基于所述第二音频声道信号及所述第四音频声道信号以及一个或多个带宽扩展参数(1358)执行所述第二多声道带宽扩展,以获得与所述音频场景的第二公共水平面或第二公共高度相关联的两个或两个以上带宽扩展的音频声道信号(622、626;1322、1326)。

[0283] 示例25. 根据示例1至24中任一项所述的音频解码器,其中,所述第一残余信号及所述第二残余信号的联合编码表示包括声道对单元,所述声道对单元包括所述第一残余信号及所述第二残余信号的下混合信号以及所述第一残余信号及所述第二残余信号的公共残余信号。

[0284] 示例26. 根据示例1至25中任一项所述的音频解码器,其中,所述音频解码器被配置为:使用多声道解码,基于所述第一下混合信号及所述第二下混合信号的联合编码表示来提供所述第一下混合信号及所述第二下混合信号,

[0285] 所述第一下混合信号及所述第二下混合信号的联合编码表示包括声道对单元,所述声道对单元包括所述第一下混合信号及所述第二下混合信号的下混合信号以及所述第一下混合信号及所述第二下混合信号的公共残余信号。

[0286] 示例27. 一种音频编码器(100;1100;1200;1500;2100),用于基于至少四个音频声道信号(110、112、114、116;1110、1112、1114、1116;1210、1212、1214、1216;2216、2226、2218、2228)提供已编码表示(130;1144、1154;1220、1222;2272、2282),

[0287] 其中,所述音频编码器被配置为使用残余信号辅助的多声道编码(140;1120;1230;2230)对至少第一音频声道信号及第二音频声道信号进行联合编码,以获得第一下混合信号(120;1122;1232;2234)及第一残余信号(142;1124;1234;2236);以及

[0288] 其中,所述音频编码器被配置为使用残余信号辅助的多声道编码(150;1130;1240;2240)对至少第三音频声道信号及第四音频声道信号进行联合编码,以获得第二下混合信号(122;1132;1242;2244)及第二残余信号(152;1134;1244;2246);以及

[0289] 其中,所述音频编码器被配置为使用多声道编码(160;1150;1260;2260)对所述第一残余信号及所述第二残余信号进行联合编码,以获得所述残余信号的联合编码表示(130;1154;1262;2264)。

[0290] 示例28. 根据示例27所述的音频编码器,其中,所述音频编码器被配置为:使用多声道编码(1140;1250;2250)对所述第一下混合信号及所述第二下混合信号进行联合编码,以获得下混合信号的联合编码表示(1144;1252;2254)。

[0291] 示例29. 根据示例28所述的音频编码器,其中,所述音频编码器被配置为:使用基于预测的多声道编码对所述第一残余信号及所述第二残余信号进行联合编码,以及

[0292] 所述音频编码器被配置为:使用基于预测的多声道编码对所述第一下混合信号及所述第二下混合信号进行联合编码。

[0293] 示例30. 根据示例27至29中任一项所述的音频编码器,其中,所述音频编码器被配置为:使用基于参数的、残余信号辅助的多声道编码对至少所述第一音频声道信号及所述第二音频声道信号进行联合编码,以及

[0294] 所述音频编码器被配置为:使用基于参数的、残余信号辅助的多声道编码对至少所述第三音频声道信号及所述第四音频声道信号进行联合编码。

[0295] 示例31. 根据示例27至30中任一项所述的音频编码器,其中,所述第一音频声道信号及所述第二音频声道信号与音频场景的垂直相邻的位置相关联,以及

[0296] 所述第三音频声道信号及所述第四音频声道信号与所述音频场景的垂直相邻的位置相关联。

[0297] 示例32. 根据示例27至31中任一项所述的音频编码器,其中,所述第一音频声道信号及所述第二音频声道信号与音频场景的第一水平位置或方位角位置相关联,以及

[0298] 所述第三音频声道信号及所述第四音频声道信号与所述音频场景的第二水平位置或方位角位置相关联,所述第二水平位置或方位角位置不同于所述第一水平位置或方位角位置。

[0299] 示例33. 根据示例27至32中任一项所述的音频编码器,其中,所述第一残余信号与音频场景的左侧相关联,且所述第二残余信号与所述音频场景的右侧相关联。

[0300] 示例34. 根据示例33所述的音频编码器,

[0301] 其中,所述第一音频声道信号及所述第二音频声道信号与所述音频场景的左侧相关联,以及

[0302] 其中,所述第三音频声道信号及所述第四音频声道信号与所述音频场景的右侧相关联。

[0303] 示例35. 根据示例34所述的音频解码器,其中,所述第一音频声道信号与所述音频场景的左下位置相关联,

[0304] 所述第二音频声道信号与所述音频场景的左上位置相关联,

[0305] 所述第三音频声道信号与所述音频场景的右下位置相关联,以及

[0306] 所述第四音频声道信号与所述音频场景的右上位置相关联。

[0307] 示例36. 根据示例27至35中任一项所述的音频编码器,其中,所述音频编码器被配置为:使用多声道编码对所述第一下混合信号及所述第二下混合信号进行联合编码,以获得下混合信号的联合编码表示,所述第一下混合信号与音频场景的左侧相关联,且所述第二下混合信号与所述音频场景的右侧相关联。

[0308] 示例37. 一种用于基于已编码表示来提供至少四个音频声道信号的方法(800),所述方法包括:

[0309] 使用多声道解码,基于第一残余信号及第二残余信号的联合编码表示来提供(810)所述第一残余信号及所述第二残余信号;

[0310] 使用残余信号辅助的多声道解码,基于第一下混合信号及所述第一残余信号来提供(820)第一音频声道信号及第二音频声道信号;以及

[0311] 使用残余信号辅助的多声道解码,基于第二下混合信号及所述第二残余信号来提供(830)第三音频声道信号及第四音频声道信号。

[0312] 示例38. 一种用于基于至少四个音频声道信号来提供已编码表示的方法(700),所述方法包括:

[0313] 使用残余信号辅助的多声道编码对至少第一音频声道信号及第二音频声道信号进行联合编码(710),以获得第一下混合信号及第一残余信号;

[0314] 使用残余信号辅助的多声道编码对至少第三音频声道信号及第四音频声道信号进行联合编码(720),以获得第二下混合信号及第二残余信号;以及

[0315] 使用多声道编码对所述第一残余信号及所述第二残余信号进行联合编码(730),以获得所述残余信号的已编码表示。

[0316] 示例39. 一种计算机程序,当在计算机上执行所述计算机程序时,所述计算机程序用于执行根据示例37或38的方法。

[0317] 参考文献:

[0318] [1] ISO/IEC 23003-3: 2012 - Information Technology - MPEG Audio Technologies, Part 3: Unified Speech and Audio Coding;

[0319] [2] ISO/IEC 23003-1: 2007 - Information Technology - MPEG Audio Technologies, Part 1: MPEG Surround。

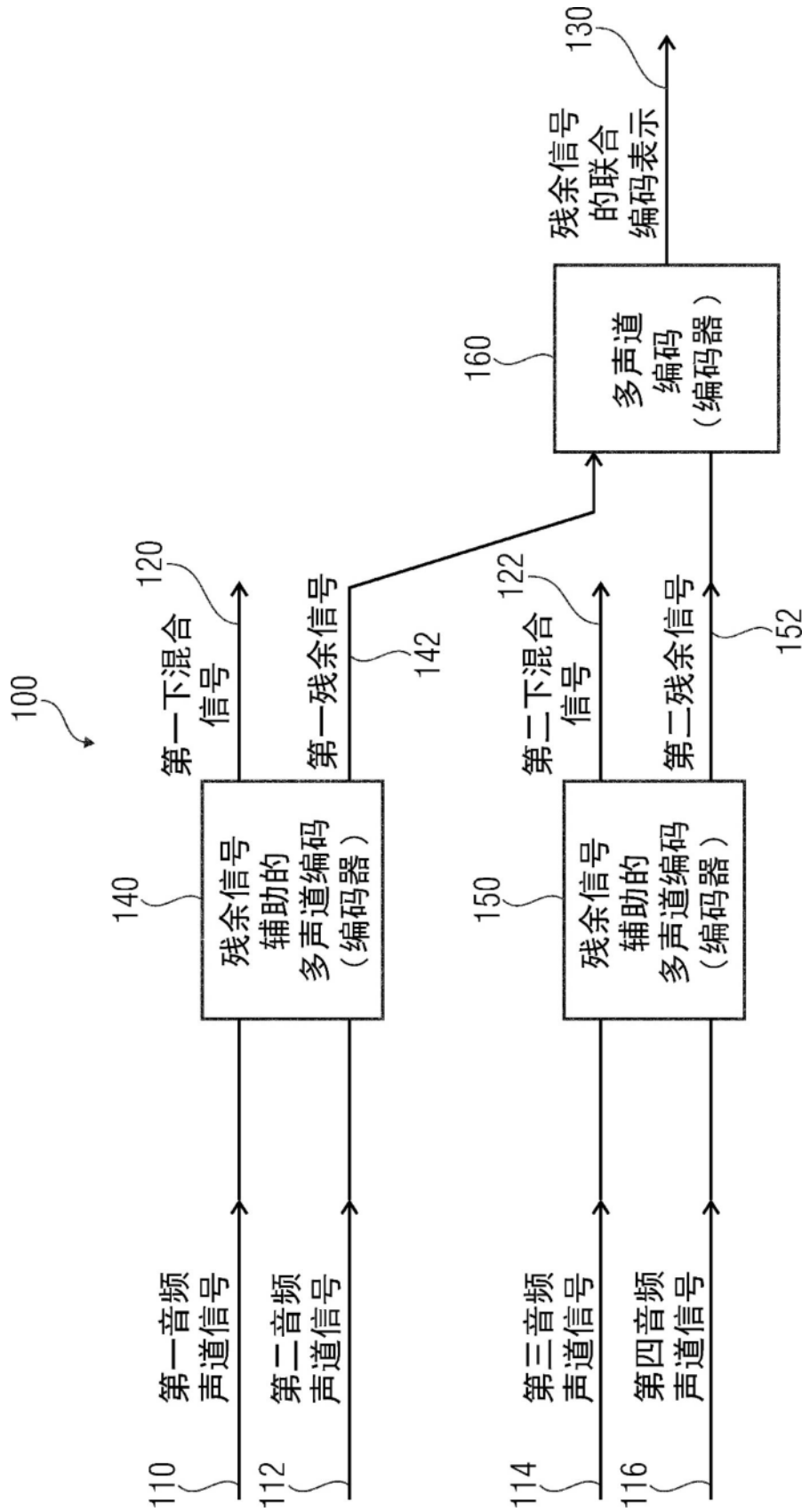


图1

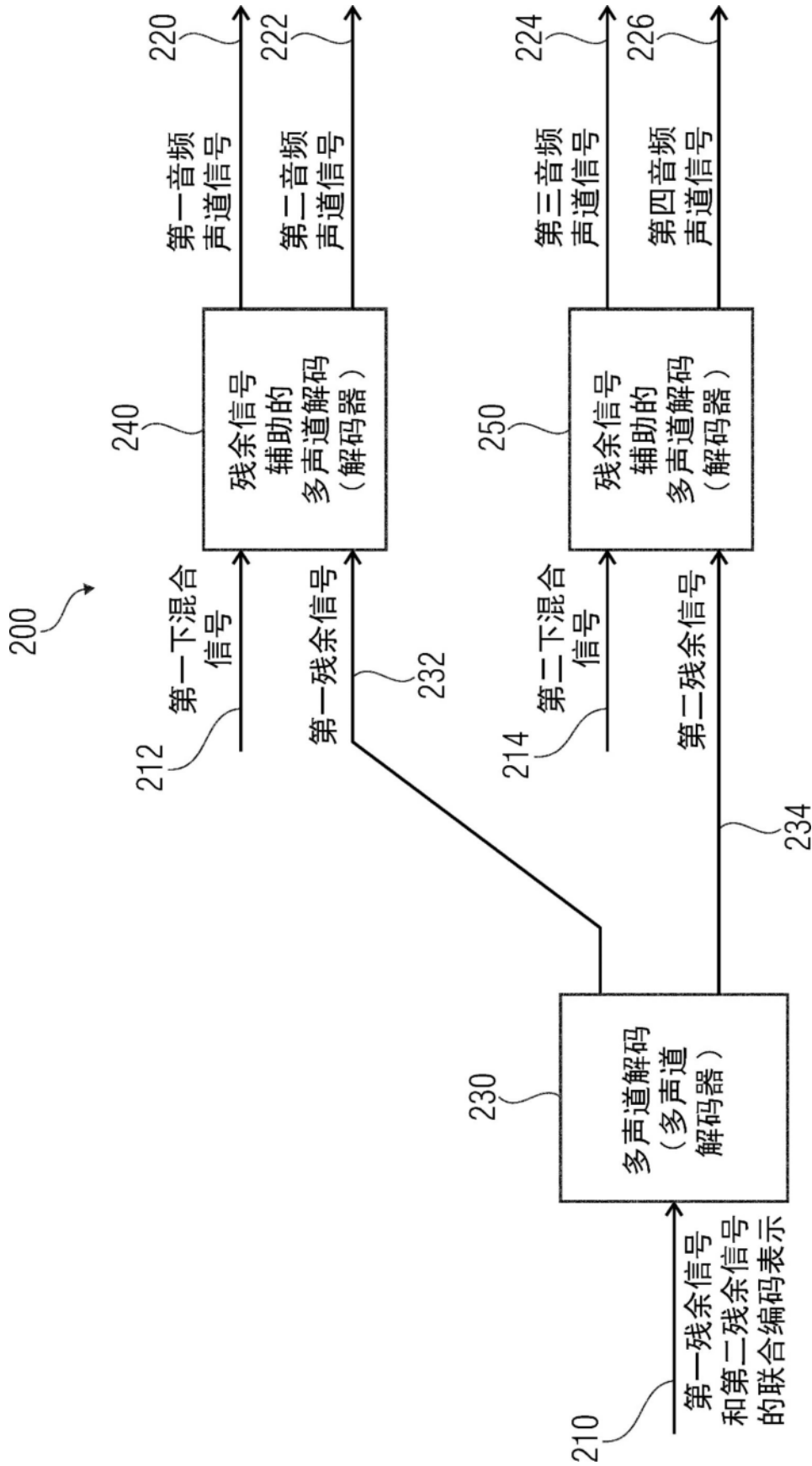


图2

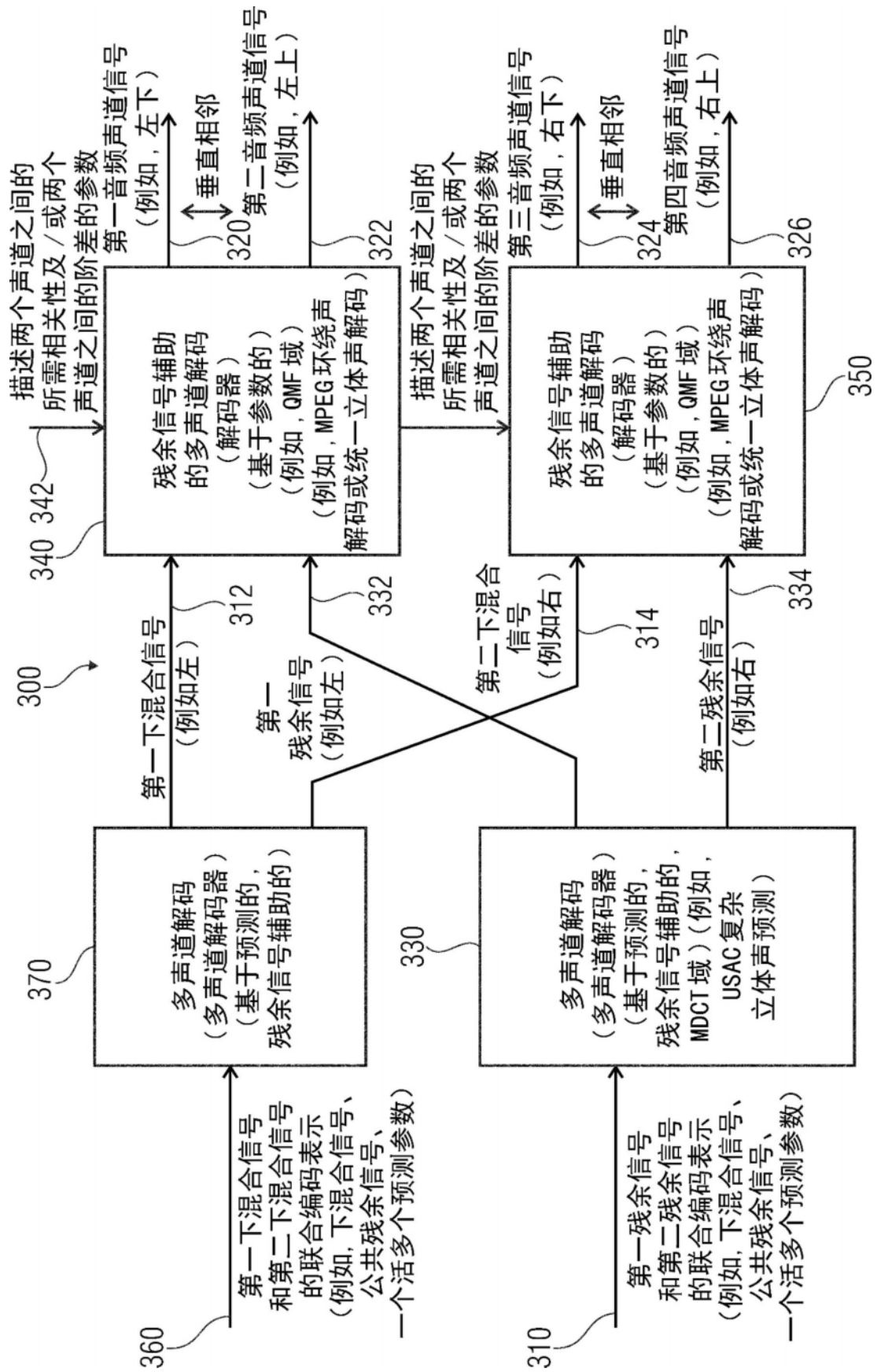


图3

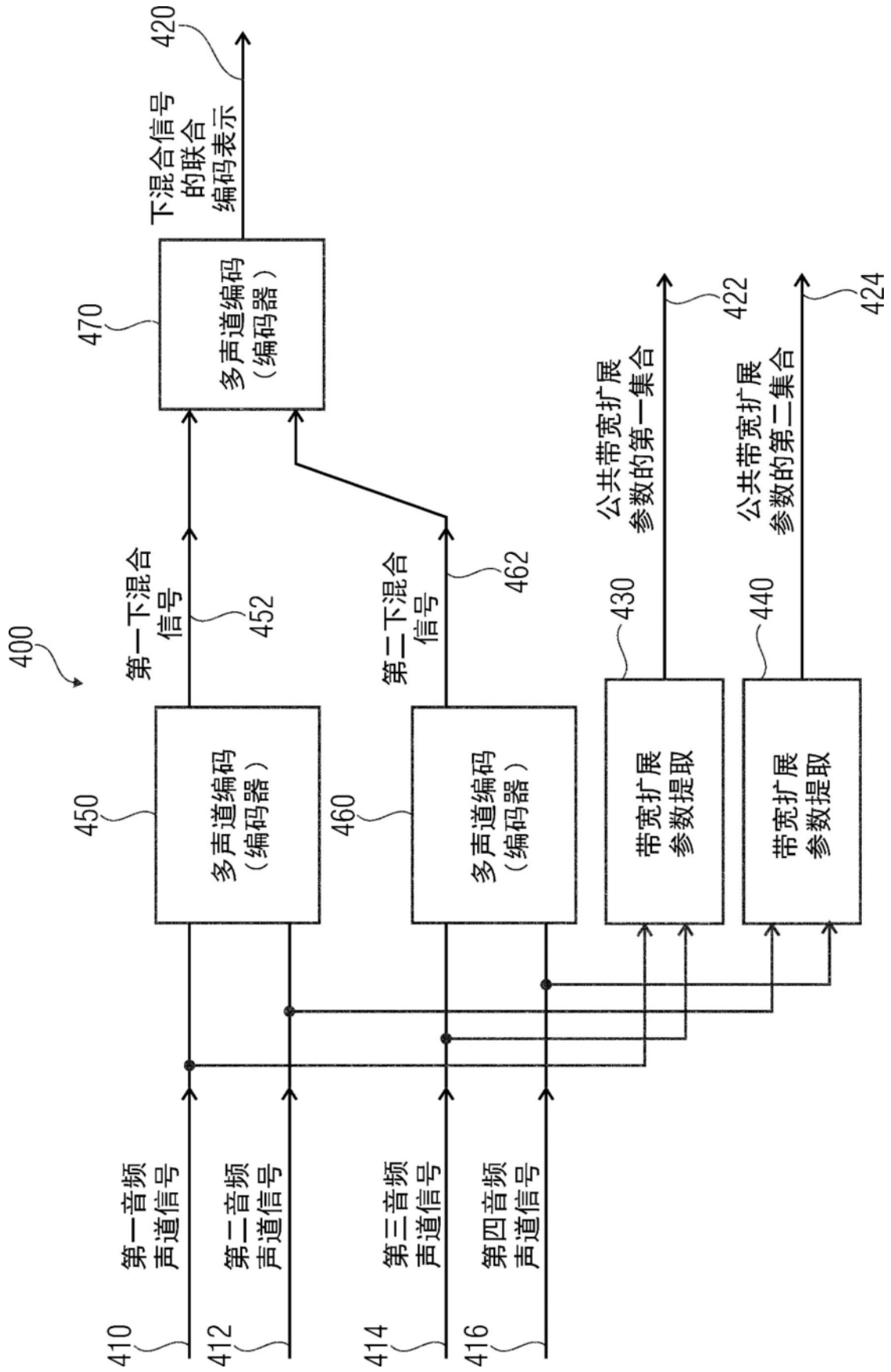


图4

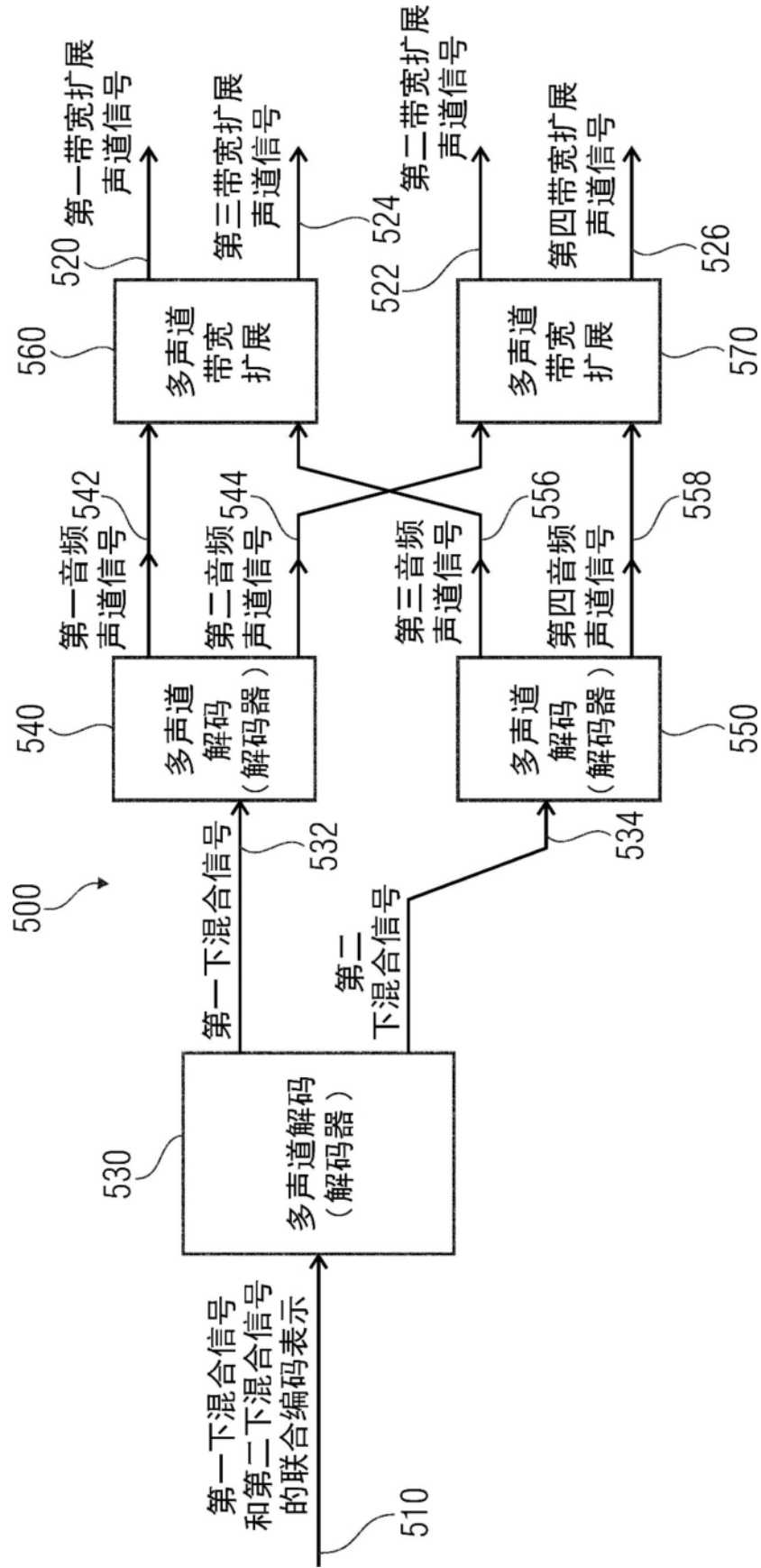


图5

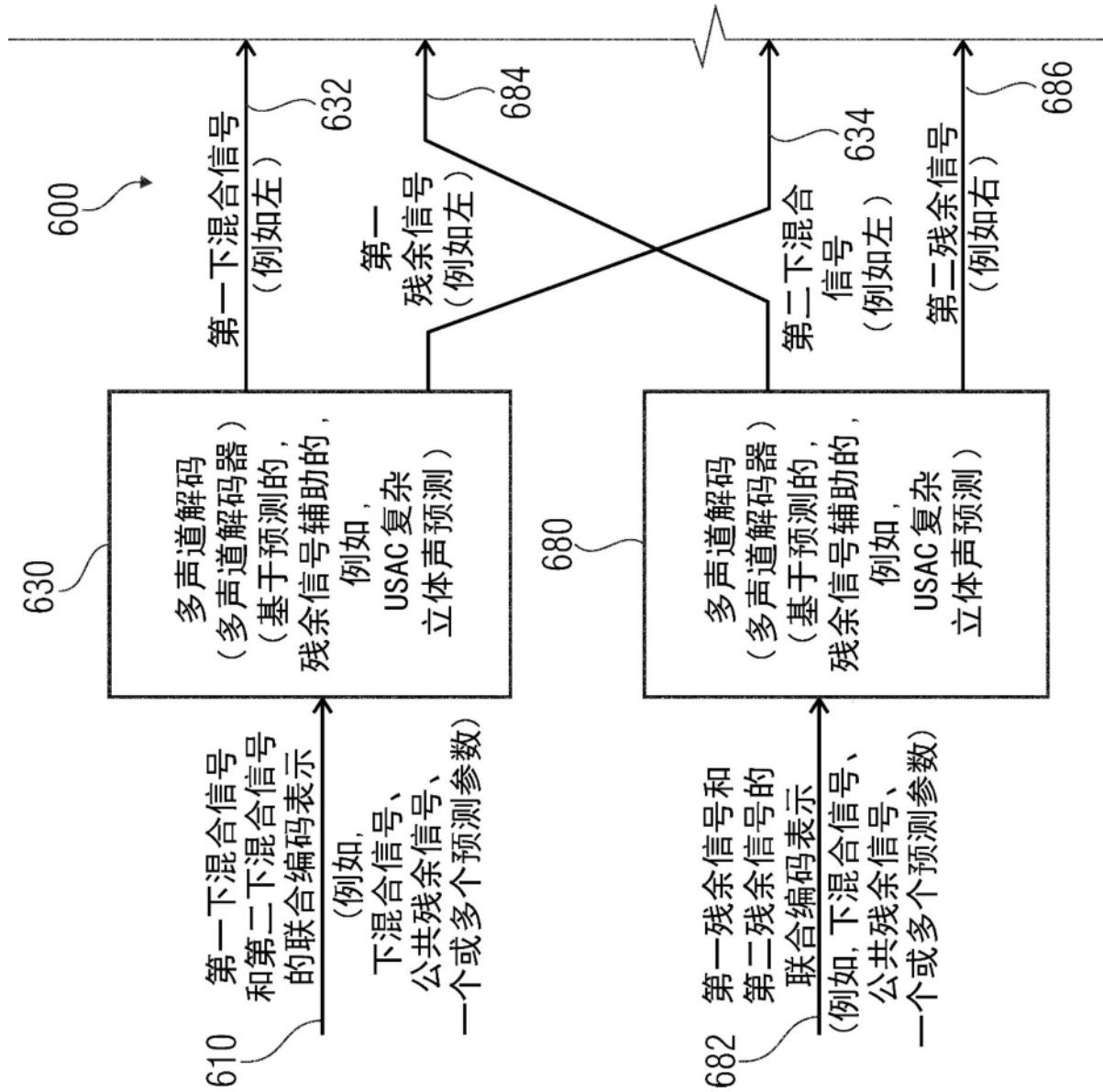


图6A

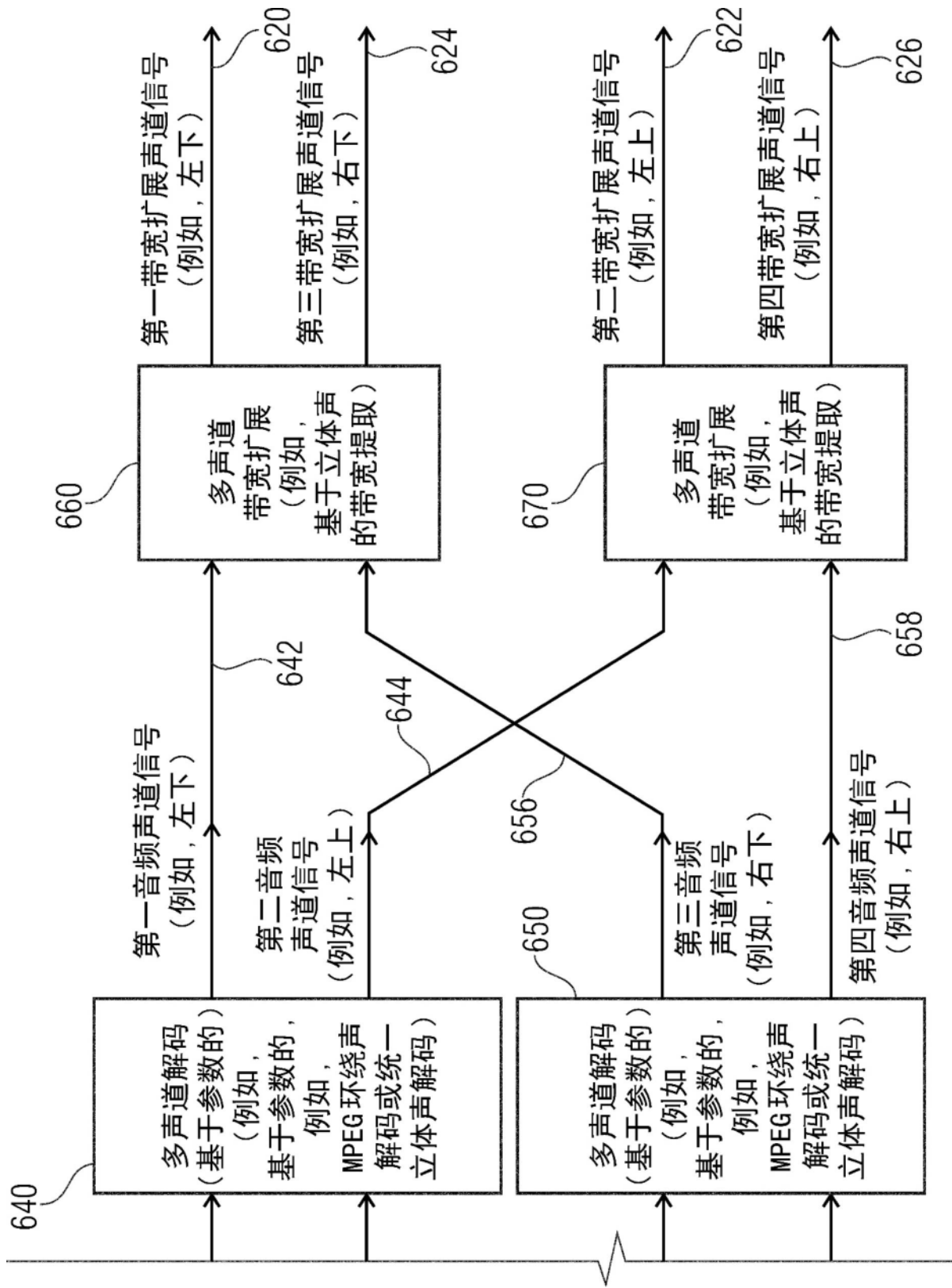


图6B

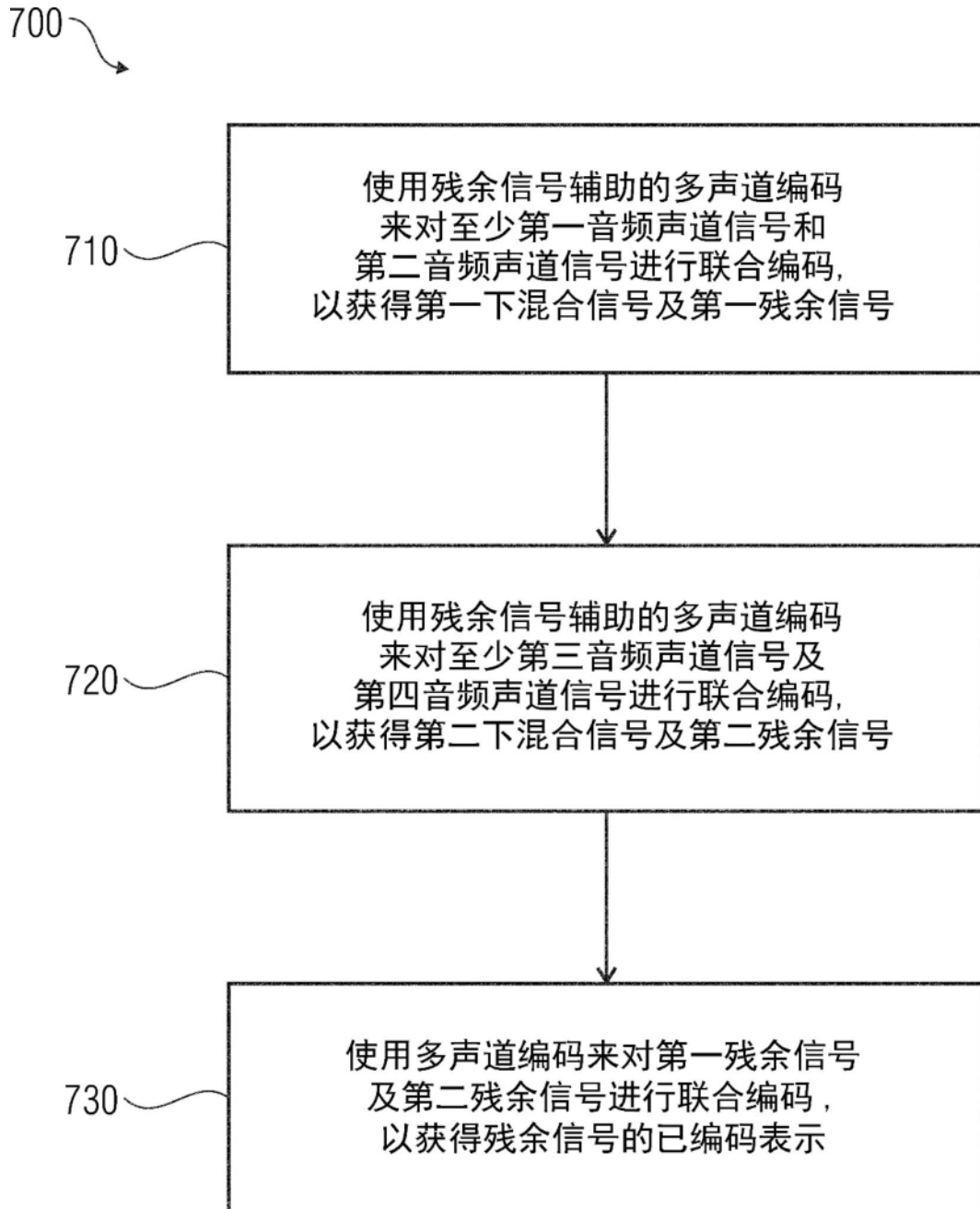


图7

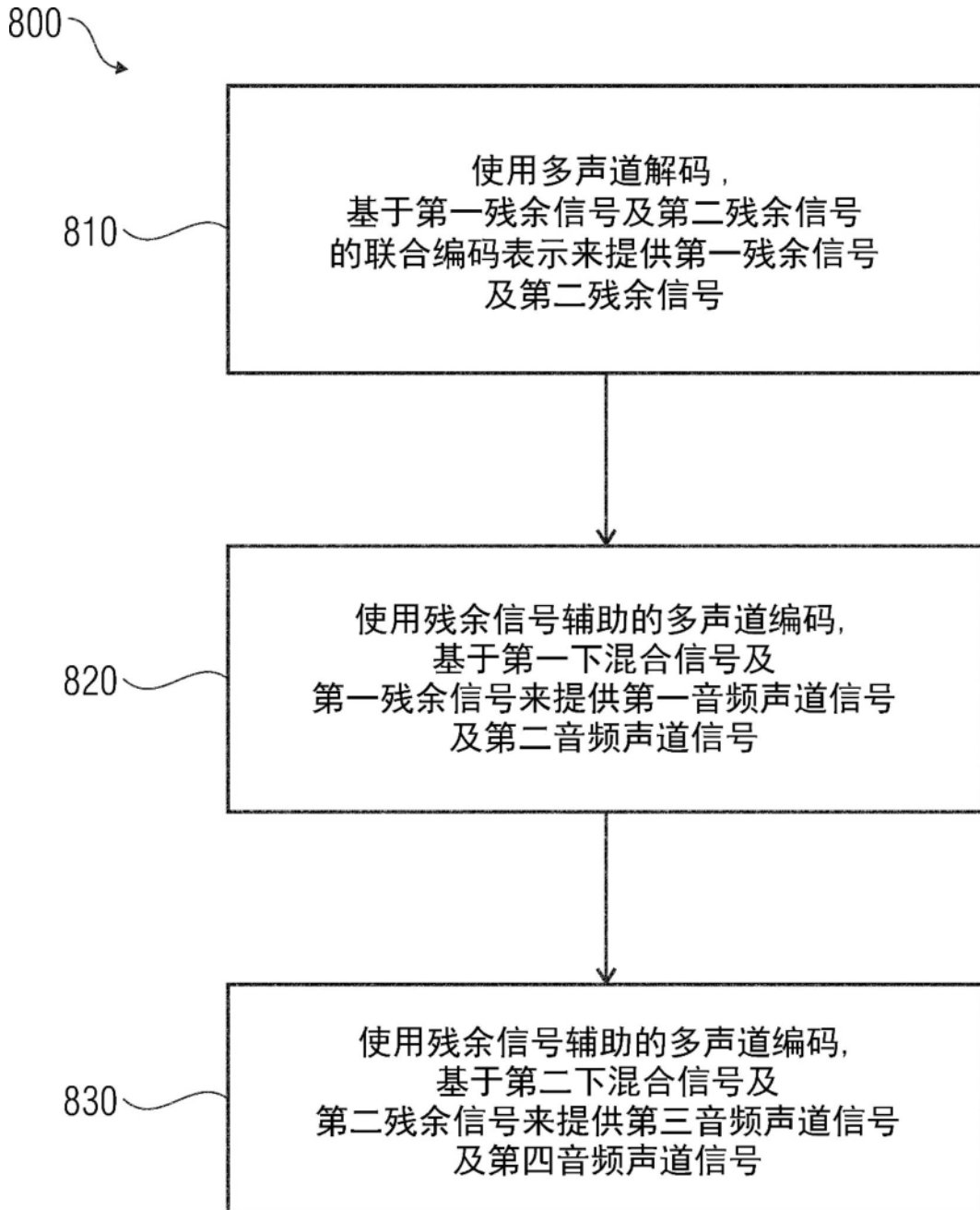


图8

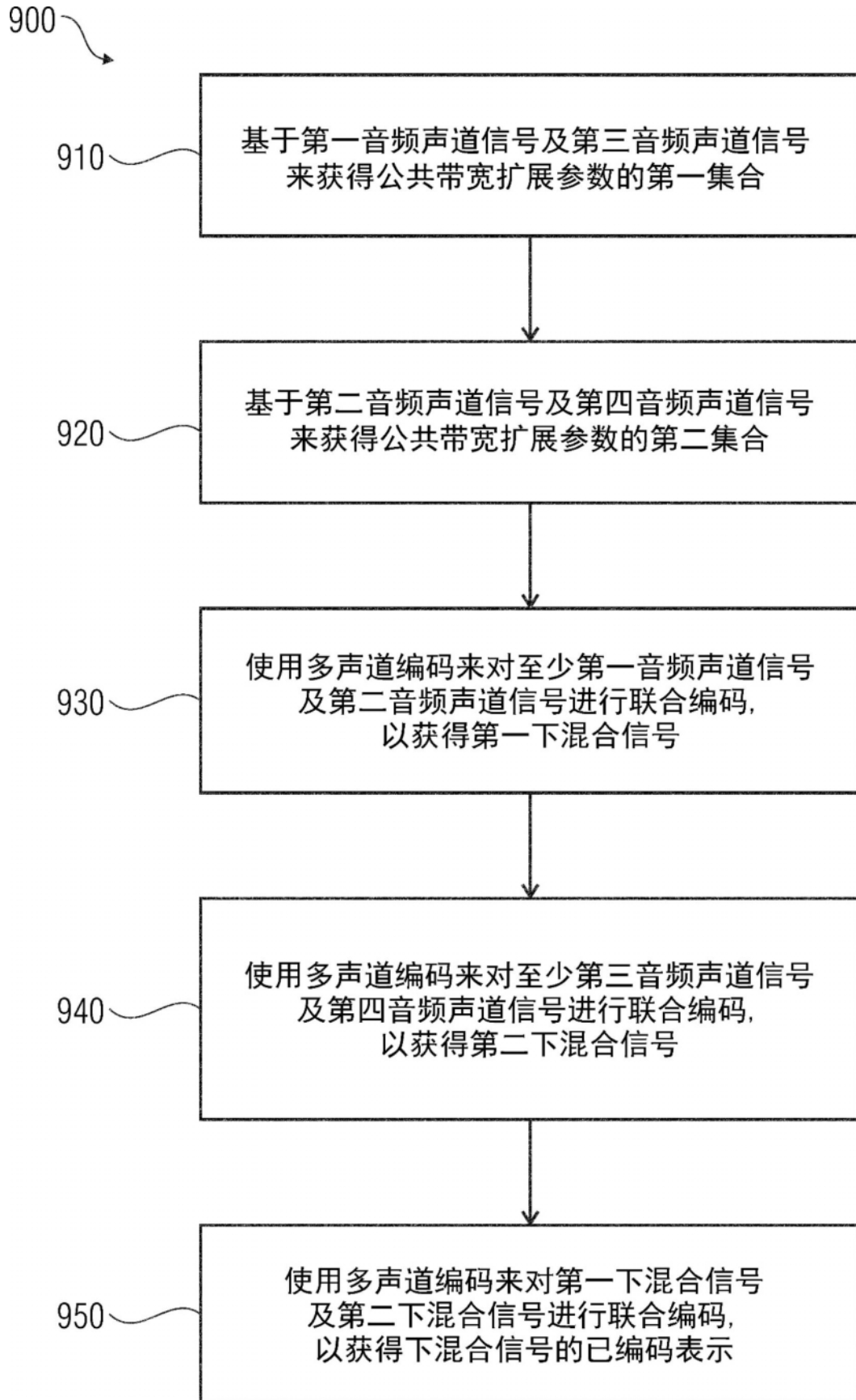


图9

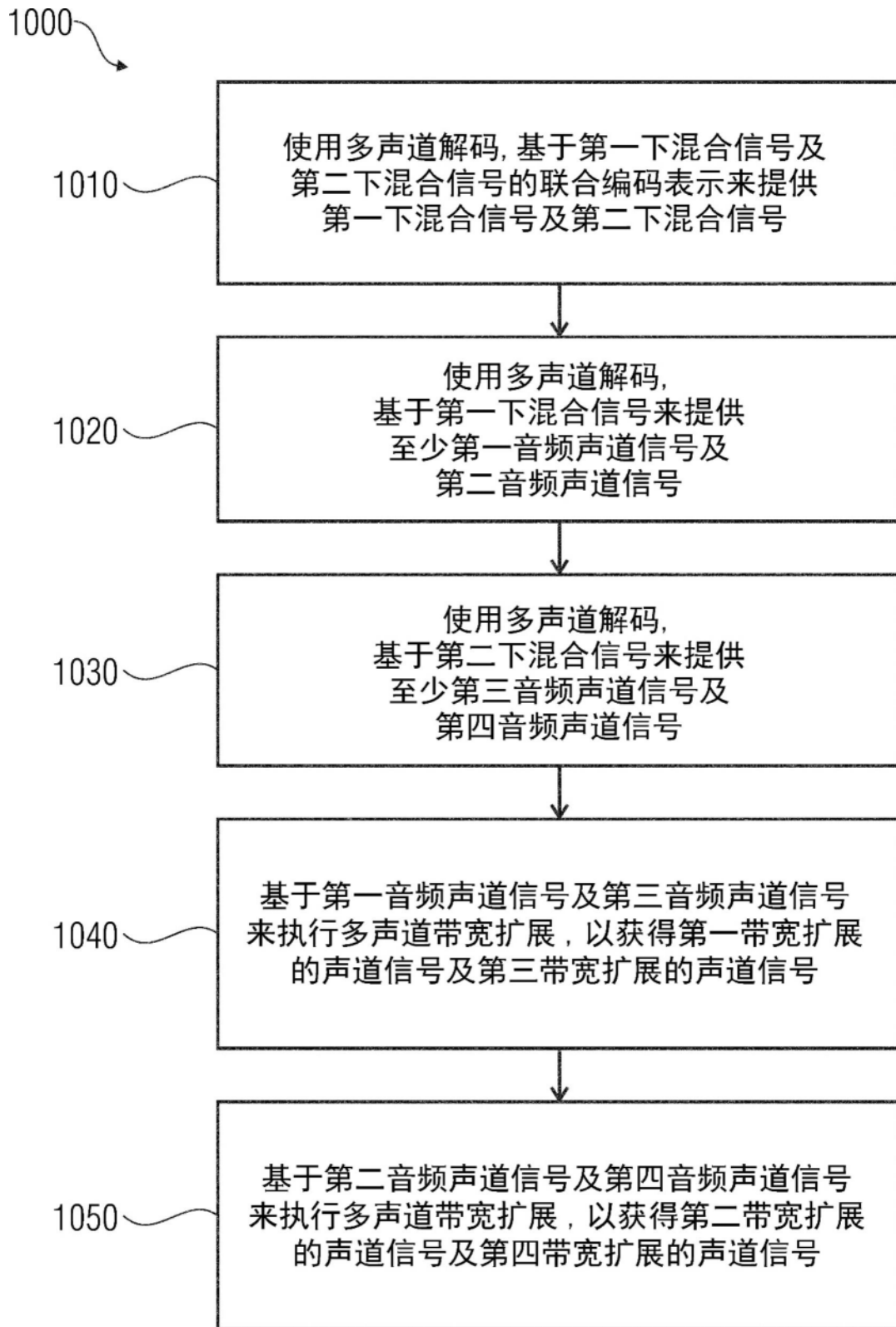


图10

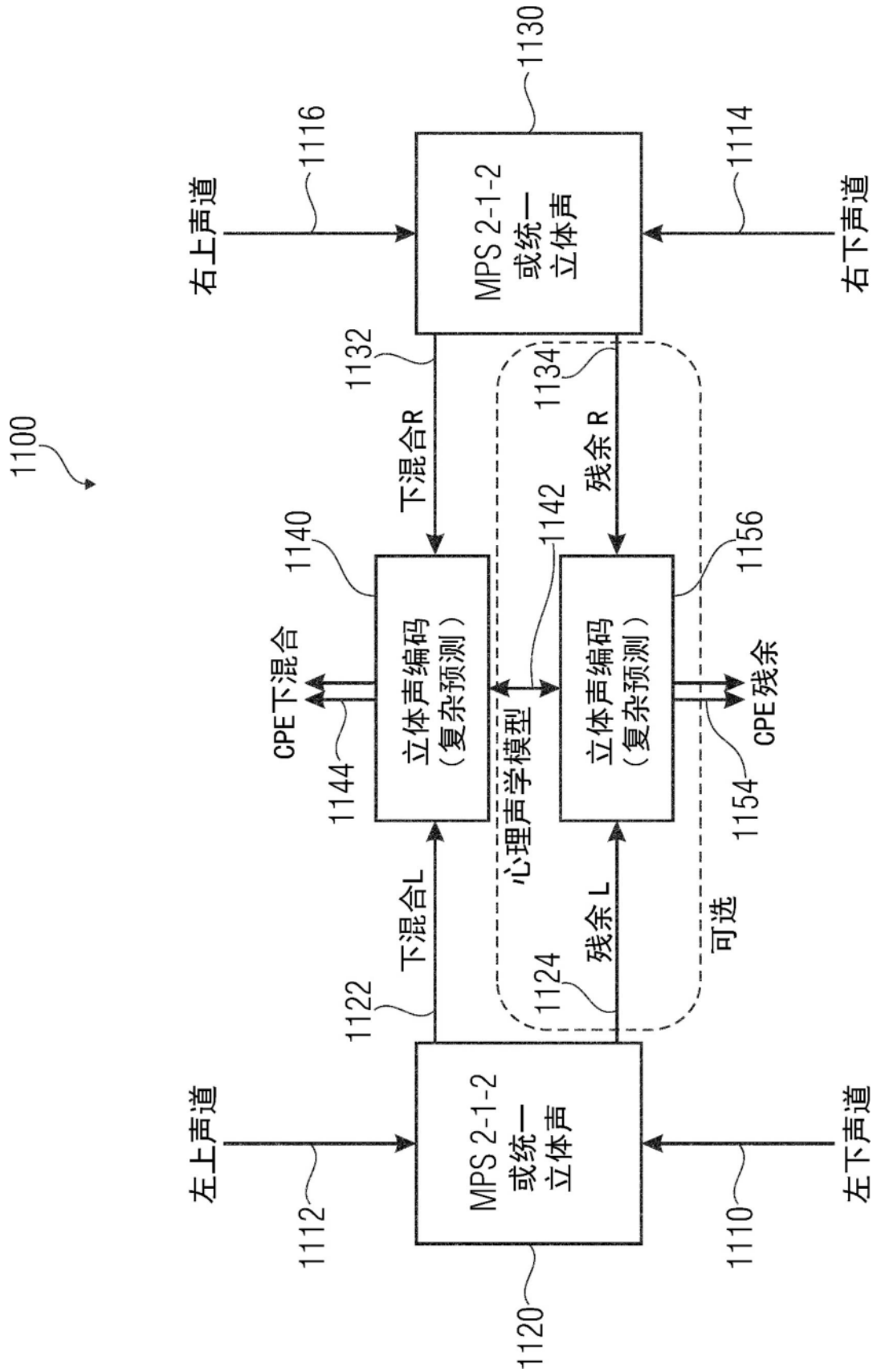


图11

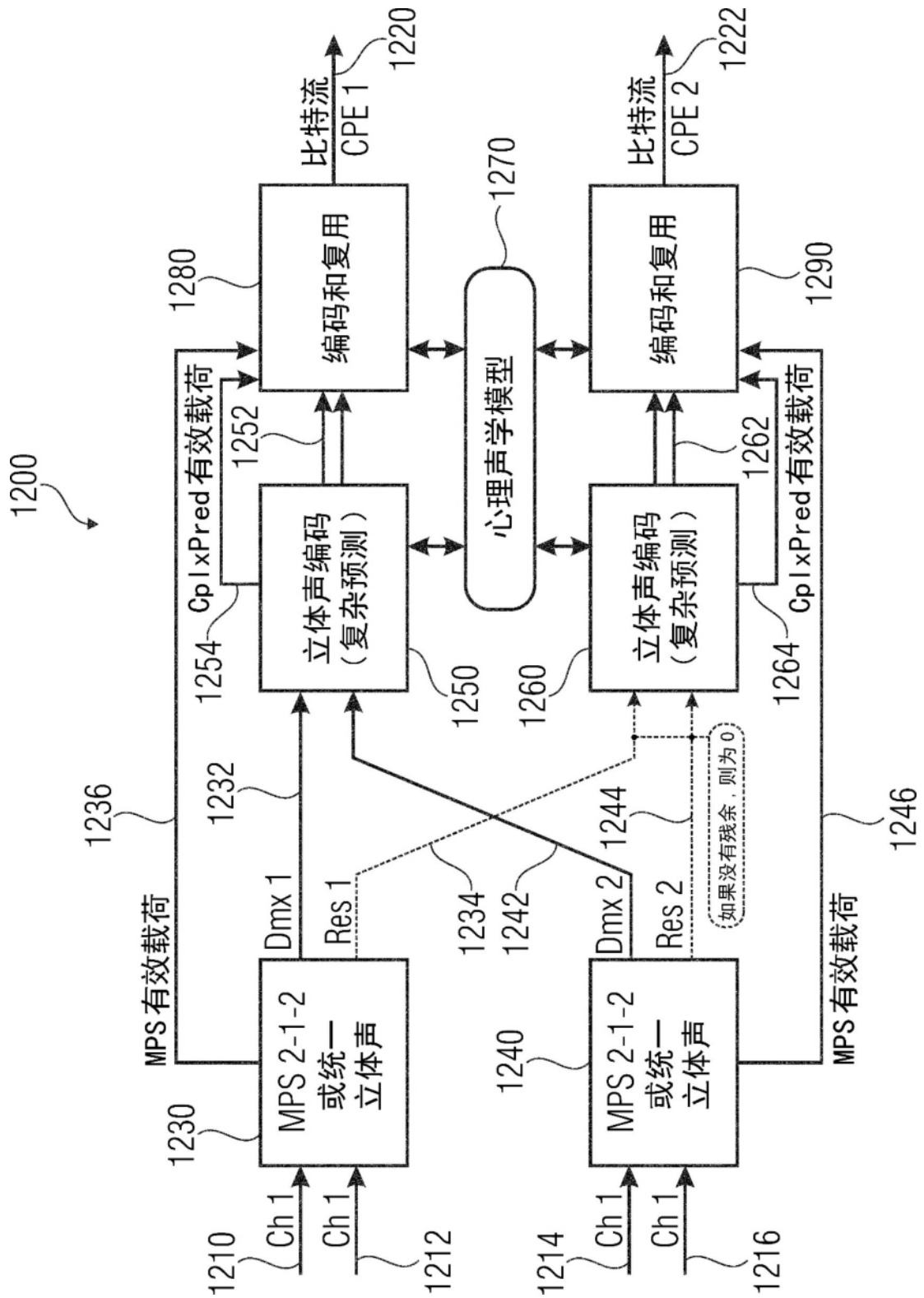


图12

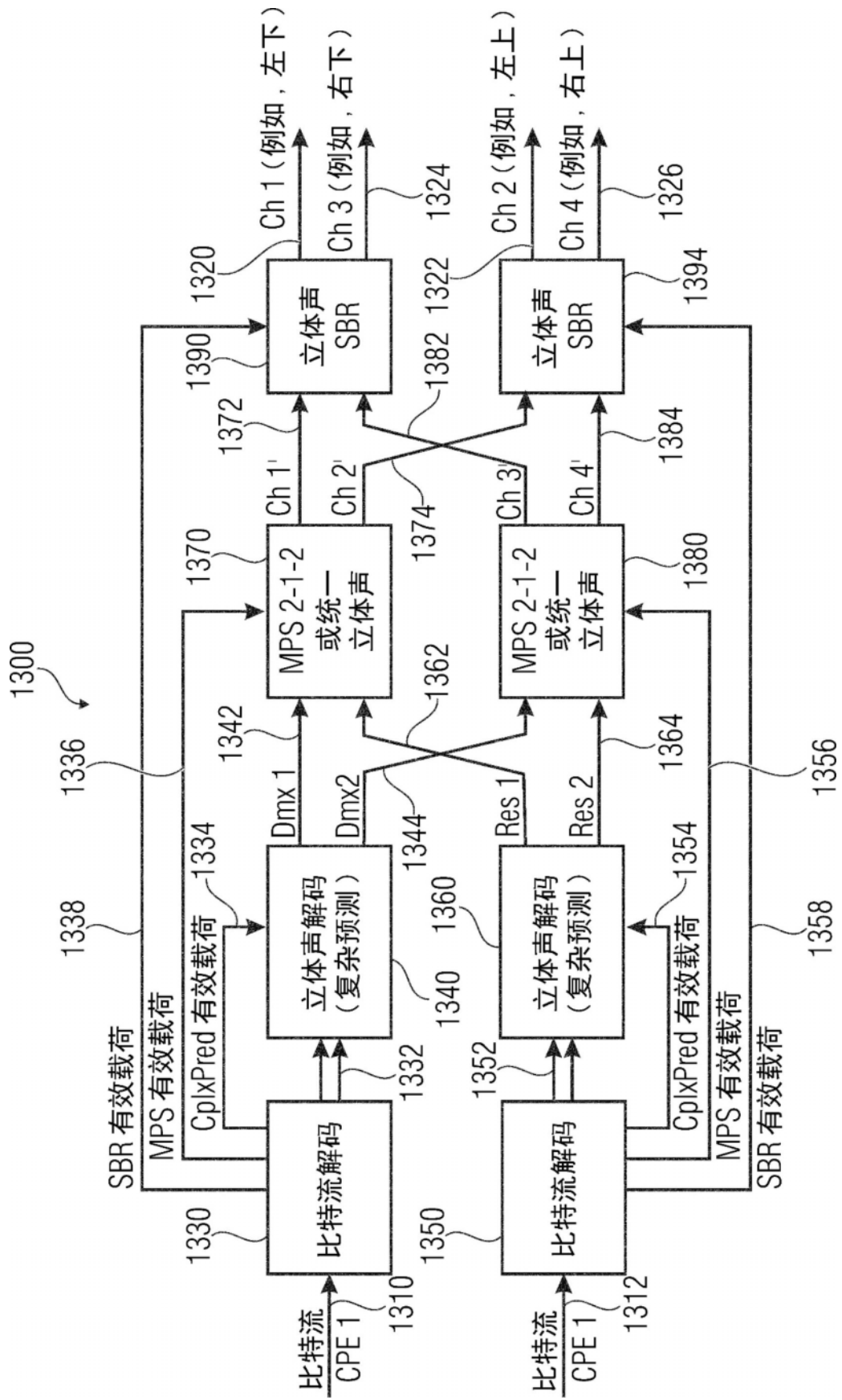


图13

```

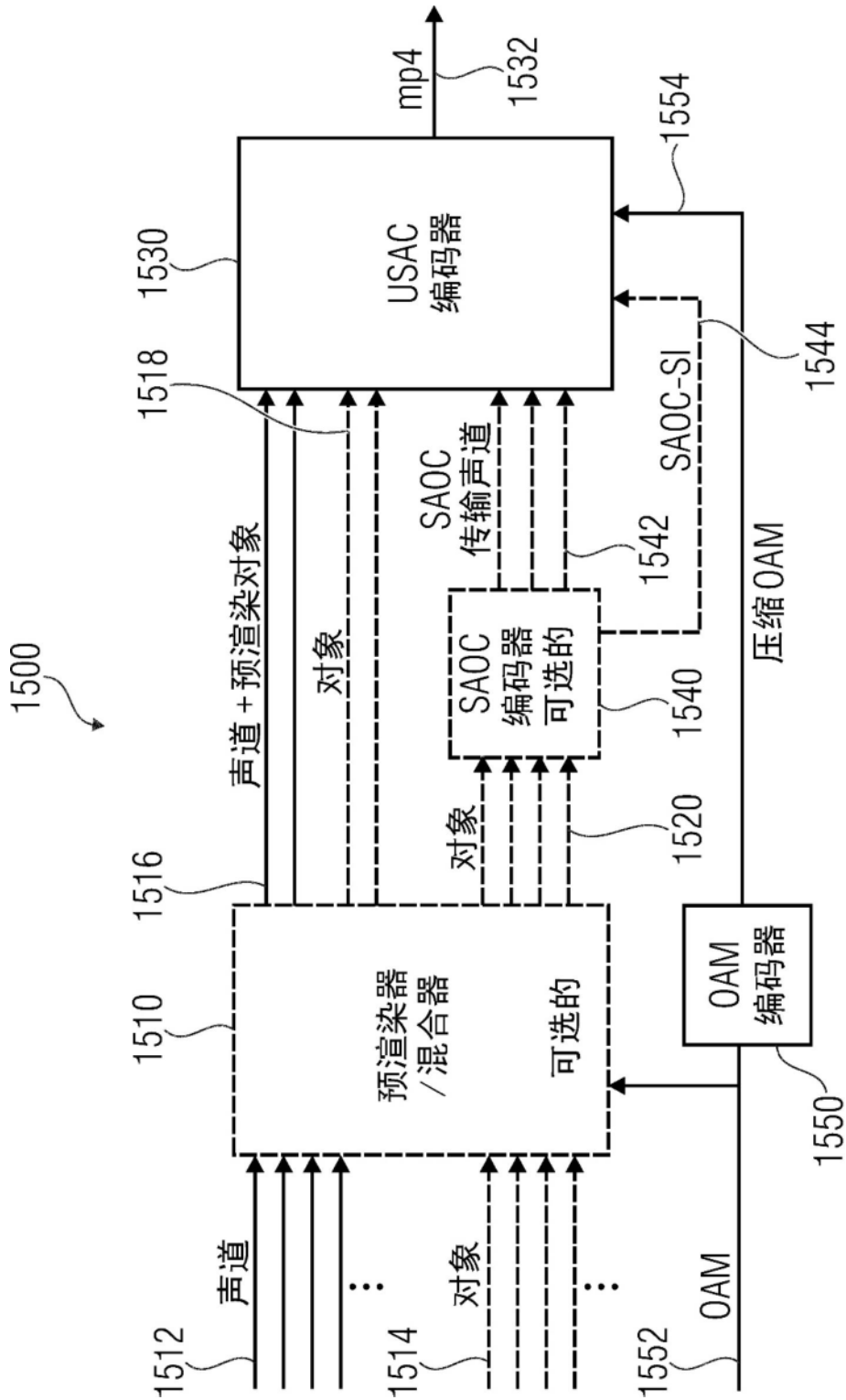
UsacChannelPairElementConfig (sbrRatiolIndex)
{
    UsacCoreConfig ();
    if (sbrRatiolIndex > 0) {
        SbrConfig ();
        stereoConfigIndex;                2         uimsbf
    } else {
        stereoConfigIndex = 0;
    }
    if (stereoConfigIndex > 0) {
        Mps212Config(stereoConfigIndex);
    }
+   qcelIndex                            2         uimsbf
}

```

图14A

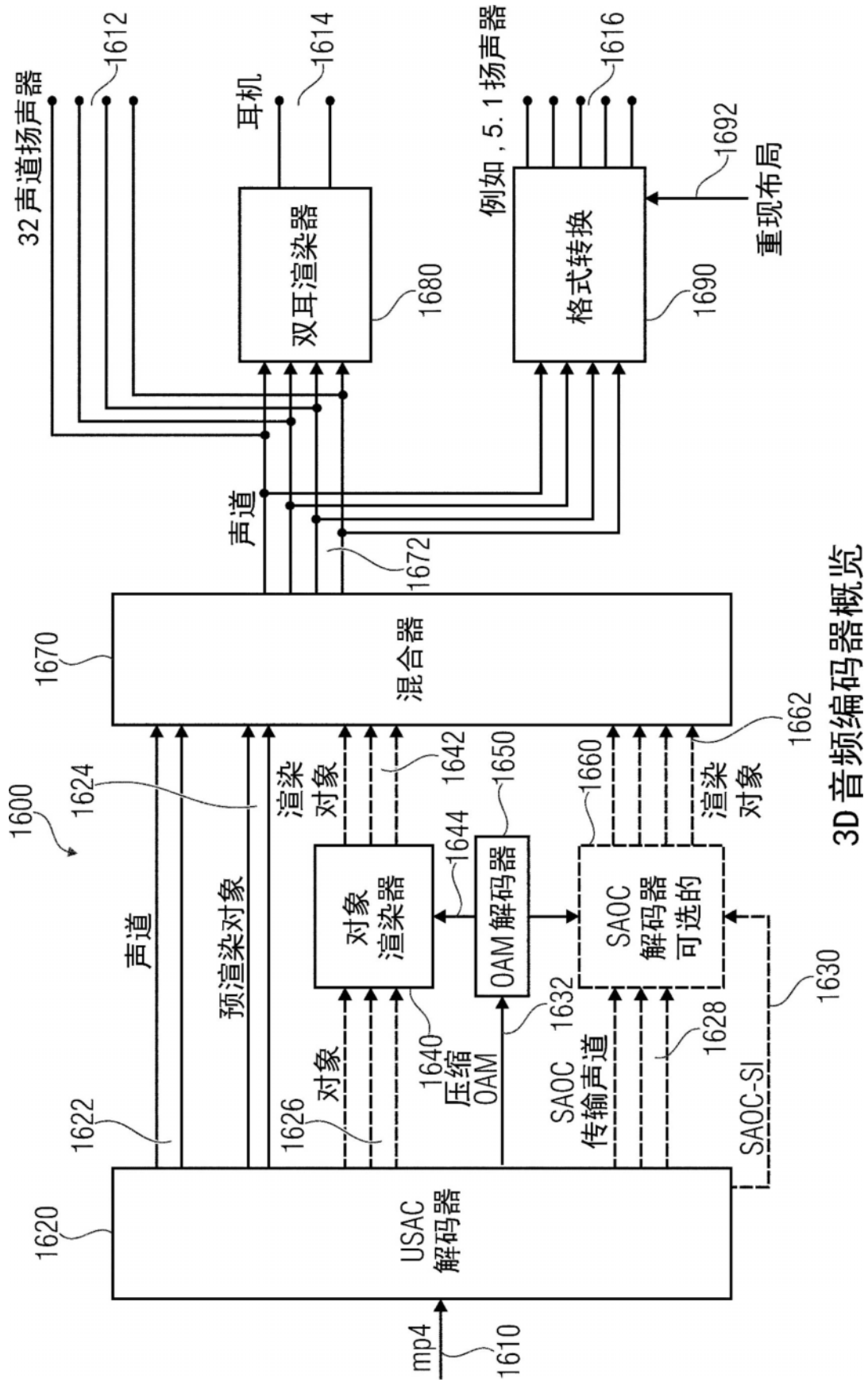
<b>qcelIndex</b>	含义
0	立体声 CPE
1	没有残余的 QCE
2	有残余的 QCE
3	- 预留 -

图14B



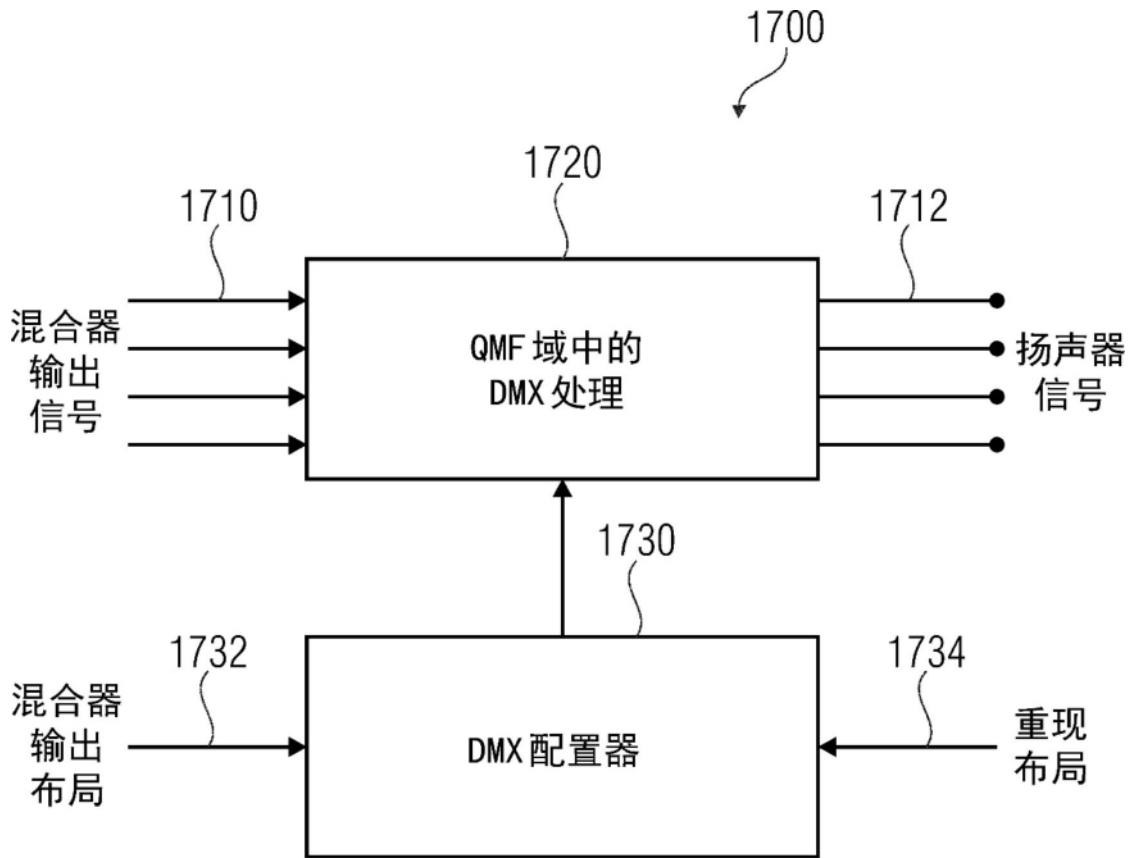
3D 音频编码器的概览

图15



3D 音频编码器概览

图16



格式转换器的结构

图17

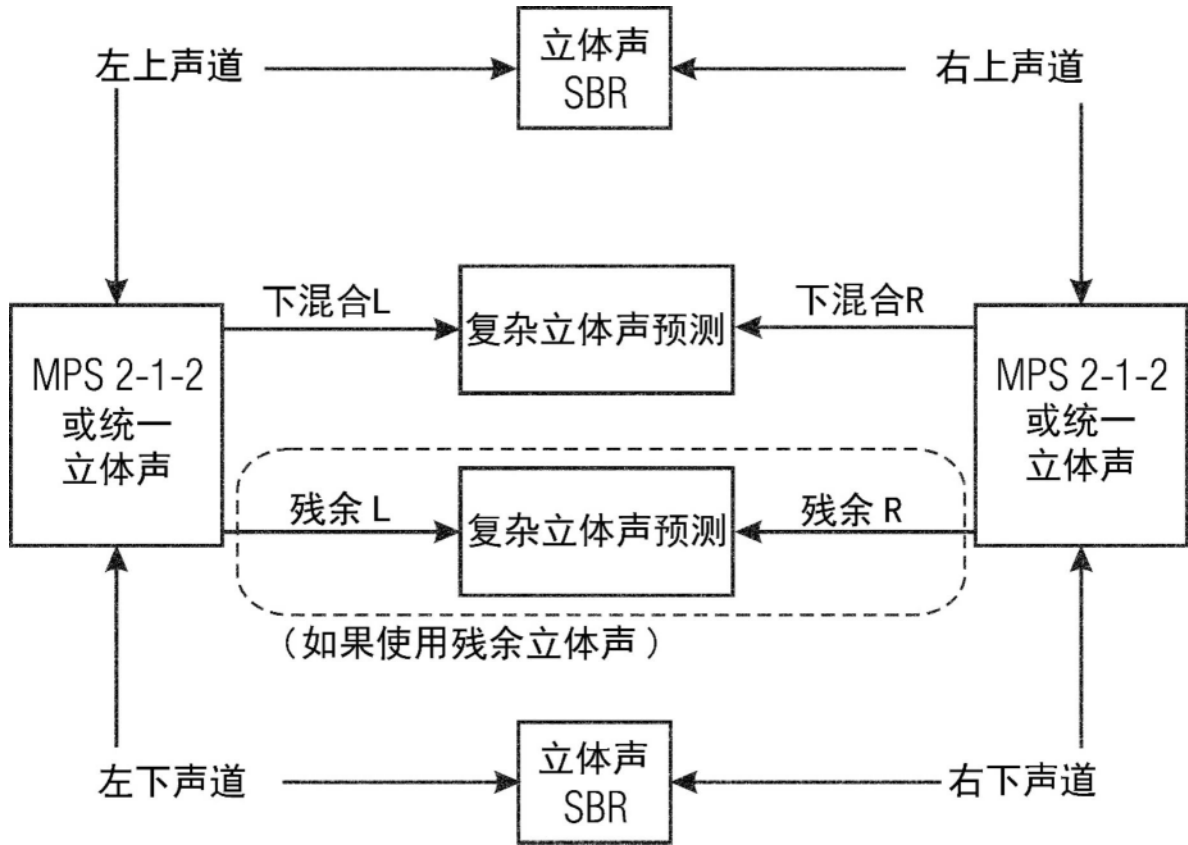


图18

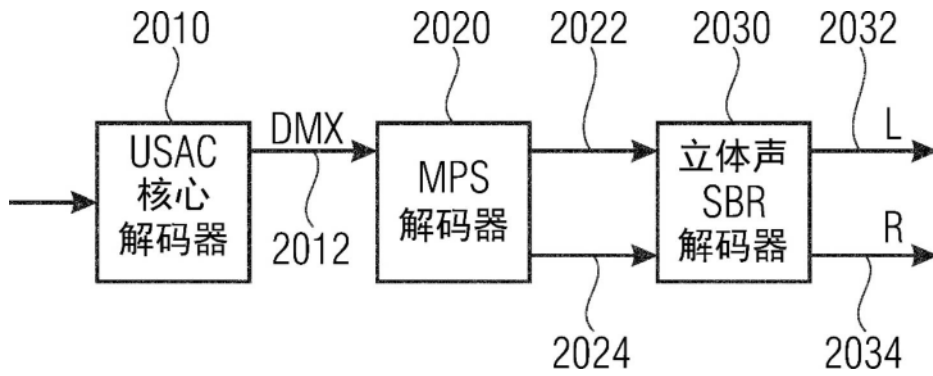


图19

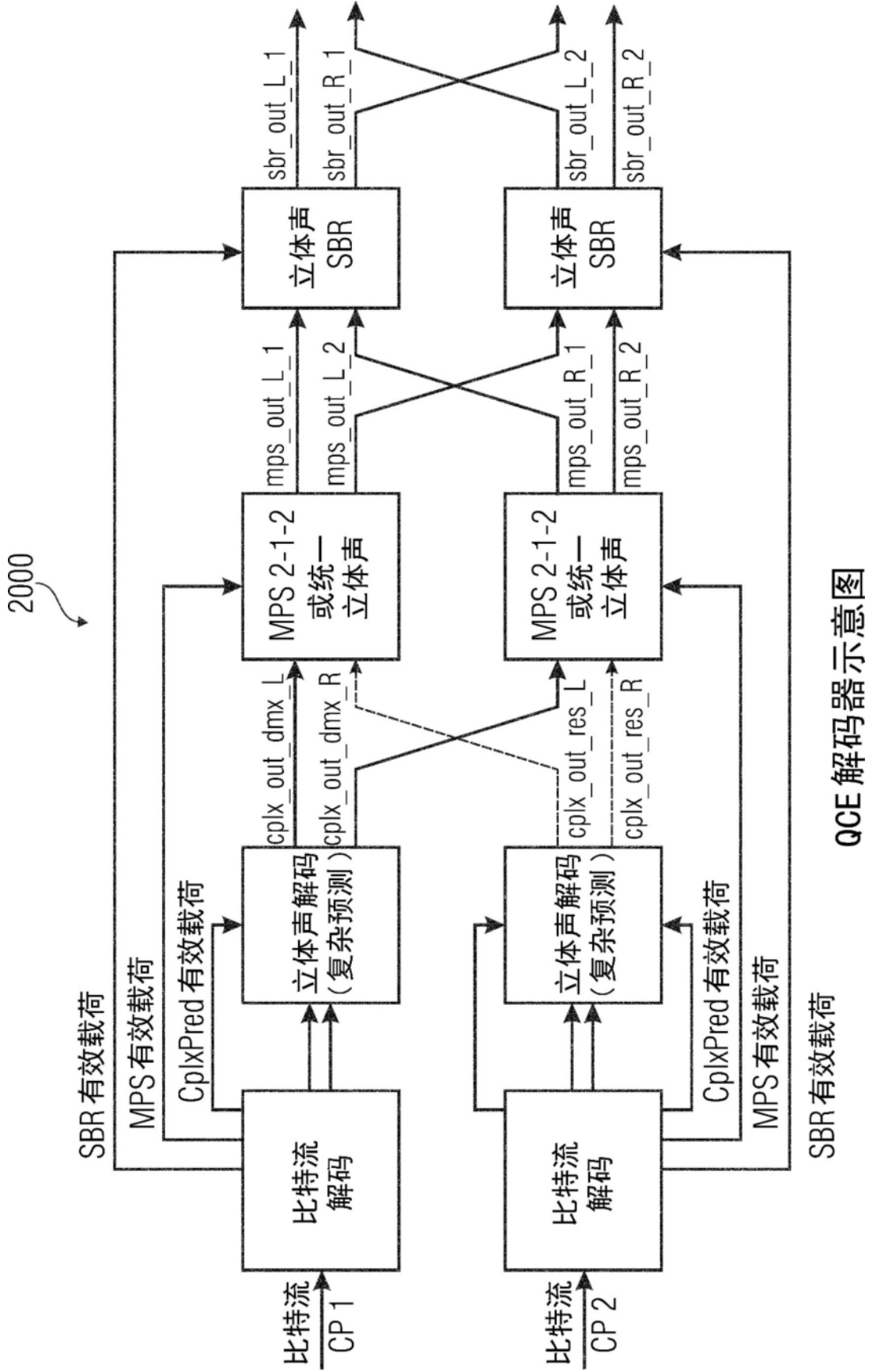


图20

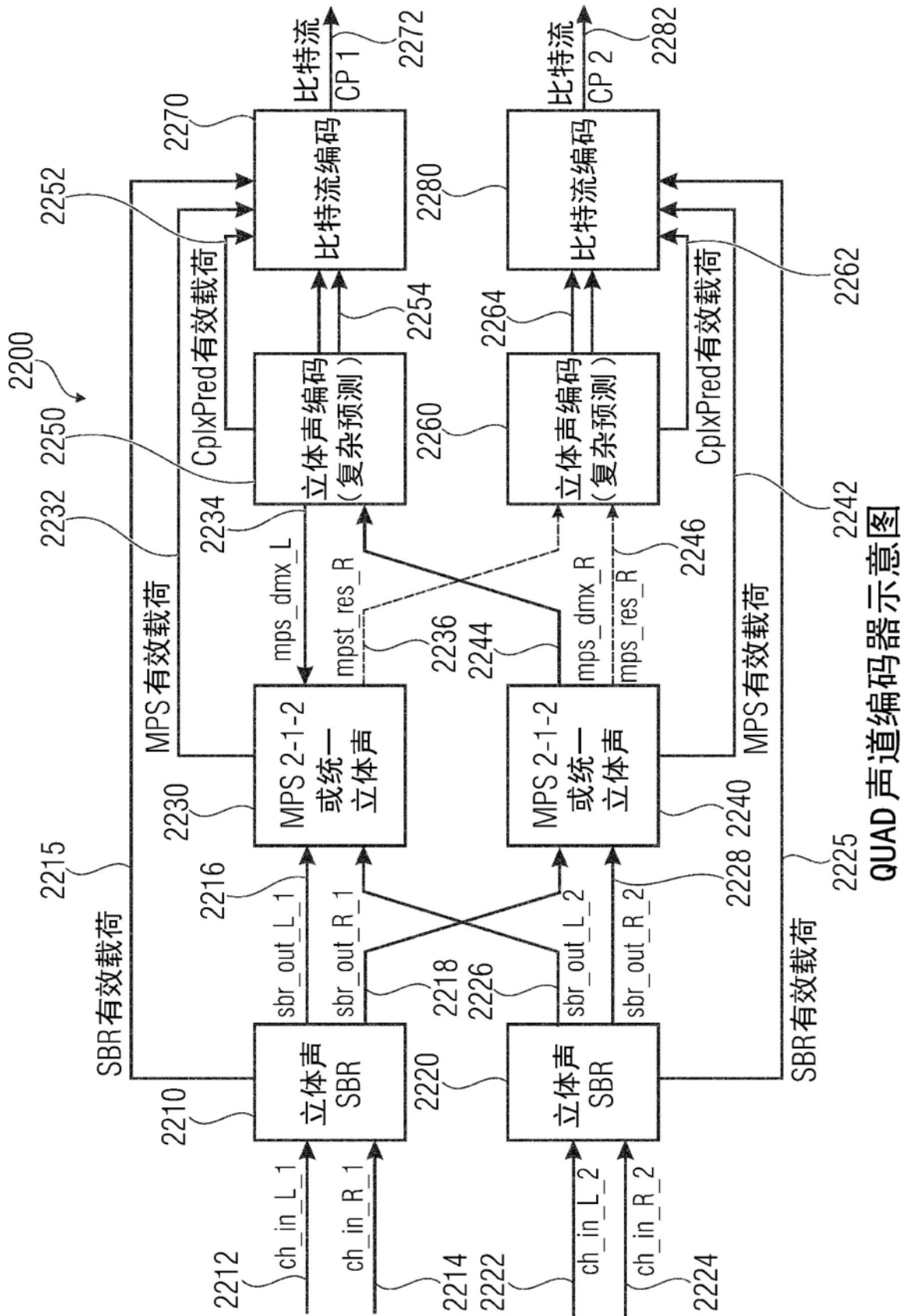


图21