



(12) 发明专利

(10) 授权公告号 CN 108206021 B

(45) 授权公告日 2020.12.18

(21) 申请号 201611171108.2

(22) 申请日 2016.12.16

(65) 同一申请的已公布的文献号
申请公布号 CN 108206021 A

(43) 申请公布日 2018.06.26

(73) 专利权人 南京青衿信息科技有限公司
地址 210000 江苏省南京市石头城69号302室

(72) 发明人 潘兴德 吴超刚 陈笑天

(74) 专利代理机构 南京先科专利代理事务所
(普通合伙) 32285

代理人 孙甫臣

(51) Int. Cl.

G10L 19/008 (2013.01)

G10L 19/16 (2013.01)

(56) 对比文件

CN 106205626 A, 2016.12.07

CN 105103543 A, 2015.11.25

CN 1278996 A, 2001.01.03

CN 103366750 A, 2013.10.23

CN 104064194 A, 2014.09.24

CN 101485094 A, 2009.07.15

CN 101632118 A, 2010.01.20

KR 20160016881 A, 2016.02.15

任鹏飞. AVS_3D实时解码器的设计与实现. 《中国优秀硕士学位论文全文数据库 信息科技辑》. 2016, (第2期), I136-344.

审查员 傅磊

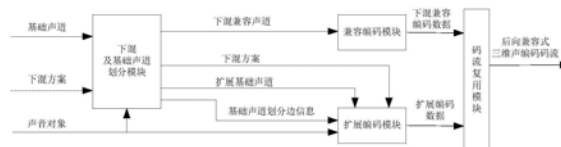
权利要求书4页 说明书21页 附图2页

(54) 发明名称

一种后向兼容式三维声编码器、解码器及其编解码方法

(57) 摘要

本发明公开一种后向兼容式三维声编码器, 包括: 下混及基础声道划分模块, 用于接收基础声道和/或声音对象, 按照下混方案进行下混和基础声道划分操作, 输出下混兼容声道、扩展基础声道、基础声道划分边信息; 兼容编码模块, 用于接收下混兼容声道, 输出下混兼容编码数据; 扩展编码模块, 用于接收声音对象、下混方案、扩展基础声道、基础声道划分边信息, 输出扩展编码数据; 码流复用模块, 复用所述下混兼容编码数据和扩展编码数据, 输出后向兼容式三维声编码码流。本发明在实现了兼容已有立体声、环绕声节目或系统的同时具备三维声编解码的能力。



1. 一种后向兼容式三维声编码器,其特征在于,包括:

下混及基础声道划分模块,用于接收基础声道和/或声音对象,按照下混方案进行下混和基础声道划分操作,输出下混兼容声道、扩展基础声道、基础声道划分边信息;

扩展编码模块,用于接收声音对象、下混方案、扩展基础声道、基础声道划分边信息,输出扩展编码数据;

码流复用模块,复用所述下混兼容声道和扩展编码数据,输出后向兼容式三维声编码码流。

2. 根据权利要求1所述的后向兼容式三维声编码器,其特征在于:所述下混及基础声道划分模块包括下混模块和基础声道划分模块,下混模块用于接收基础声道和/或声音对象,输出下混兼容声道和内部生成的下混方案,基础声道划分模块用于接收基础声道、内部生成的下混方案,输出扩展基础声道、基础声道划分边信息;当所述下混模块接收不包括基础声道时,下混及基础声道划分模块不进行基础声道划分操作,扩展编码模块不进行扩展基础声道、基础声道划分边信息的扩展编码操作;当所述下混模块接收不包括声音对象时,扩展编码模块不进行声音对象的扩展编码操作。

3. 根据权利要求1所述的后向兼容式三维声编码器,其特征在于:所述下混及基础声道划分模块包括下混模块和基础声道划分模块,下混模块用于接收基础声道和/或声音对象、外部输入的下混方案,输出下混兼容声道;基础声道划分模块用于接收基础声道、外部输入的下混方案,输出扩展基础声道、基础声道划分边信息;当所述下混模块接收不包括基础声道时,下混及基础声道划分模块不进行基础声道划分操作,扩展编码模块不进行扩展基础声道、基础声道划分边信息的扩展编码操作;当所述下混模块接收不包括声音对象时,扩展编码模块不进行声音对象的扩展编码操作。

4. 根据权利要求1所述的后向兼容式三维声编码器,其特征在于:所述下混及基础声道划分模块包括下混模块和基础声道划分模块,基础声道划分模块用于接收基础声道和外部输入的下混方案,输出所述扩展基础声道和基础声道划分边信息;所述扩展编码数据经扩展解码模块解码生成解码的基础声道、解码的基础声道划分边信息、解码的下混方案、解码的声音对象与所述基础声道输入至下混模块,下混模块输出所述下混兼容声道;当所述三维声不包括声音对象时,扩展编码模块不进行声音对象的扩展编码操作。

5. 根据权利要求2至4任一所述的后向兼容式三维声编码器,其特征在于:所述下混模块按照下混方案将基础声道、声音对象下混为下混兼容声道,下混兼容声道信号分为基础声道下混分量、声音对象下混分量,基础声道下混分量分为扩展基础声道下混分量、兼容基础声道下混分量。

6. 根据权利要求2至4任一所述的后向兼容式三维声编码器,其特征在于:所述基础声道划分模块将基础声道划分为兼容基础声道、扩展基础声道;所述基础声道划分模块采用的基础声道划分方案根据基础声道的声道配置、所要兼容的多声道系统及下混映射函数确定。

7. 根据权利要求6所述的后向兼容式三维声编码器,其特征在于:根据对应下混声道的划分方法确定的基础声道划分方案包括:

S11: 令集合 $S_{bedt} = S_{bed}$, $S_c = \emptyset$, $S_{srt} = S_{sr}$, 基础声道信号集合 $S_{bed} =$

$\{\overline{bedchannel}[i], i = 1 \dots Nb\}$, 下混兼容声道信号集合 $Ssr = \{\overline{srchannel}[k], k = 1 \dots Ns\}$, $fb(k, i)$ 为下混映射函数;

S12: 遍历集合 $Ssrt$, 找出一个满足如下关系的下混兼容声道 k :

$$\overline{srchannel_bed}[k] == \sum_{i=1}^{Nb} fb(k, i) (\overline{bedchannel}[i]),$$

对于所有属于 Sc 中的声道 n , 有 $fb(k, n) = 0$;

若没有找到, 执行步骤 S15;

S13: 对于步骤 S12 中的下混兼容声道 k , 遍历集合 $Sbedt$, 找出 $fb(k, m)$ 不为 0 且 $fb(k, m)$ 可逆的基础声道 m , 若没有找到, 执行步骤 S15;

S14: 将步骤 S13 中找到的基础声道 m 加入到集合 Sc 中, 得到新的 Sc , 从 $Ssrt$ 中去除下混兼容声道 k , 得到新的 $Ssrt$, 从 $Sbedt$ 中去除所有满足 $fb(k, i)$ 不为 0 的基础声道 i , 得到新的 $Sbedt$, 若新的 $Ssrt$ 、新的 $Sbedt$ 都不为空, 跳转到步骤 S12, 若新的 $Ssrt$ 、新的 $Sbedt$ 为空, 则执行步骤 S15;

S15: Sc 或 Sc 的子集作为基础声道集合 $Sbed$ 的兼容基础声道集合。

8. 根据权利要求 6 所述的后向兼容式三维声编码器, 其特征在于: 根据基于 QR 分解的划分方法确定的基础声道划分方案包括:

S21: 令 $Sbedc = Sbed$, $Sbed$ 为基础声道信号集合;

S22: 采用矩阵运算形式表达 $Sbedc$ 的下混: $Hs_bedcop = HAc * Hbc$, Hs_bedcop 为 $Sbedc$ 下混形成的下混分量组成的矩阵, Hbc 为 $Sbedc$ 中基础声道信号组成的矩阵, HAc 为 $Sbedc$ 下混系数矩阵;

S23: 对 HAc 进行 QR 分解, 得 $HAc = Q * HR$, 其中 Q 为一个 $Ns \times Ns$ 的正交矩阵, HR 是一个 $Ns \times Nbc$ 的上三角矩阵;

S24: 设 $M = \min(Ns, Nbc)$, Ns 为基础声道下混的声道数、 Nbc 为 $Sbedc$ 的声道数, 如果对于 HR 中每一个 $n = 1 \dots M$, 都满足 $r(n, n) > 0$, 执行步骤 S25; 否则, 对于每一个 $n = 1 \dots M$, 如果 HR 中 $r(n, n) = 0$, 将 $Sbedc$ 中第 n 个声道从 $Sbedc$ 中剔除形成新的集合 $Sbedc'$, 令 $Sbedc = Sbedc'$ 并执行步骤 S22;

S25: 保留 $Sbedc$ 中 $n = 1 \dots M$ 的声道组成的集合, 该集合或集合的子集作为基础声道集合 $Sbed$ 的兼容基础声道集合。

9. 根据权利要求 1 至 4 任一所述的后向兼容式三维声编码器, 其特征在于: 还包括兼容编码模块, 兼容编码模块用于接收所述下混兼容声道, 输出下混兼容编码数据, 所述码流复用模块复用所述下混兼容编码数据和扩展编码数据, 输出后向兼容式三维声编码码流。

10. 根据权利要求 9 所述的后向兼容式三维声编码器, 其特征在于: 所述兼容编码模块与所述扩展编码模块采用相同的编码格式进行编码或采用不同的编码格式进行编码。

11. 一种后向兼容式三维声编码方法, 其特征在于, 包括如下步骤:

按照下混方案将基础声道和/或声音对象下混为所兼容的下混兼容声道,

将基础声道划分为扩展基础声道、兼容基础声道并确定基础声道划分边信息;

对下混兼容声道进行编码生成下混兼容编码数据;

对声音对象、下混方案、扩展基础声道、基础声道划分边信息进行编码得到扩展编码数据;

复用下混兼容编码数据和扩展编码数据得到后向兼容式三维声编码码流。

12. 一种用于兼容系统的三维声解码器,用于对权利要求1至4任一所述后向兼容式三维声编码器输出的后向兼容式三维声编码码流进行解码,其特征在于,包括:

码流解复用模块,用于获取后向兼容式三维声编码码流进行解复用得到下混兼容编码数据和扩展编码数据;

兼容解码模块,用于接收下混兼容编码数据,输出解码的下混兼容声道。

13. 一种用于三维声系统的三维声解码器,用于对权利要求1至4任一所述后向兼容式三维声编码器输出的后向兼容式三维声编码码流进行解码,其特征在于:

码流解复用模块,用于获取后向兼容式三维声编码码流进行解复用得到下混兼容编码数据和扩展编码数据;

兼容解码模块,用于接收下混兼容编码数据,输出解码的下混兼容声道;

扩展解码模块,用于接收扩展编码数据,输出解码的下混方案、解码的扩展基础声道、解码的基础声道划分边信息、解码的声音对象;

去下混模块,用于接收解码的下混兼容声道、解码的下混方案、解码的扩展基础声道、解码的基础声道划分边信息、解码的声音对象,输出解码的兼容基础声道;

基础声道组合模块,用于接收解码的兼容基础声道、解码的扩展基础声道、解码的基础声道划分边信息,输出解码的基础声道。

14. 根据权利要求13所述的用于三维声系统的三维声解码器,其特征在于,所述去下混模块包括:按照解码的下混方案,从解码的下混兼容声道信号中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分,得到解码的兼容基础声道信号。

15. 根据权利要求13所述的用于三维声系统的三维声解码器,其特征在于,所述去下混模块包括如下步骤:

1) 计算兼容基础声道下混分量:按照解码的下混方案,从解码的下混兼容声道信号中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分,得到解码的兼容基础声道下混分量;

2) 逆映射:对解码的兼容基础声道下混分量进行逆映射得到解码的兼容基础声道信号。

16. 根据权利要求15所述的用于三维声系统的三维声解码器,其特征在于,针对根据对应下混声道的划分方法确定的基础声道划分方案进行的下混和基础声道划分操作,所述去下混模块包括如下步骤:

1) 计算兼容基础声道下混分量:对于每一个兼容基础声道 $n=1 \dots Nbc$, $k = \text{compat}(n)$ 为其对应的下混兼容声道,从解码的下混兼容声道信号 $\overrightarrow{srchannel}[k]$ 中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分得到兼容基础声道下混分量 $\overrightarrow{srchannel_bed_{cop}}[k]$;

2) 逆映射:对于每一个兼容基础声道 $n=1 \dots Nbc$, $k = \text{compat}(n)$ 为其对应的下混兼容声道,按照解码的下混方案中映射函数 $fb(k, bctob(n))$ 的逆函数 $fb^{-1}(k, bctob(n))$ 对解码的兼容基础声道下混分量 $\overrightarrow{srchannel_bed_{cop}}[k]$ 进行逆映射,得到解码的兼容基础声道 $\overrightarrow{bedchannel_{cop}}[n]$:

$$\overrightarrow{bedchannel_{cop}}[n] = fb^{-1}(k, bctob(n))(\overrightarrow{srchannel_bed_{cop}}[k])。$$

17. 根据权利要求15所述的用于三维声系统的三维声解码器, 其特征在于, 针对根据基于QR分解的划分方法确定的基础声道划分方案进行的下混和基础声道划分操作, 所述去下混模块包括如下步骤:

1) 计算兼容基础声道下混分量: 对于每一个下混兼容声道 k , 从解码的下混兼容声道信号 $\overrightarrow{srchannel[k]}$ 中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分得到兼容基础声道下混分量 $\overrightarrow{srchannel_bed_cop[k]}$, 所有的 $\overrightarrow{srchannel_bed_cop[k]}$, $k = 1 \dots N_s$ 组成矩阵 Hs_bedcop

$$Hs_bedcop = \begin{bmatrix} \overrightarrow{srchannel_bedcop[1]} \\ \vdots \\ \overrightarrow{srchannel_bedcop[N_s]} \end{bmatrix}$$

2) 逆映射: 根据解码的兼容基础声道映射系数逆矩阵 $invHRQ$, 对 Hs_bedcop 进行逆映射求得 Hbc

$$Hbc = invHRQ * Hs_bedcop$$

Hbc 中的第 n 行即为解码的兼容基础声道信号 $\overrightarrow{bedchannel_cop[n]}$, $n = 1 \dots N_{bc}$ 。

18. 一种后向兼容式三维声解码方法, 用于对权利要求11所述后向兼容式三维声编码方法得到的后向兼容式三维声编码码流进行解码, 其特征在于, 包括如下步骤:

获取后向兼容式三维声编码码流进行解复用得到下混兼容编码数据和扩展编码数据;

解码下混兼容编码数据得到解码的下混兼容声道, 解码扩展编码数据得到解码的下混方案、解码的扩展基础声道、解码的基础声道划分边信息、解码的声音对象;

通过从解码的下混兼容声道中减去扩展基础声道、声音对象的下混成分的方式, 解码得到解码的兼容基础声道;

对解码的兼容基础声道和扩展基础声道进行声道组合得到解码的基础声道。

一种后向兼容式三维声编码器、解码器及其编解码方法

技术领域

[0001] 本发明涉及一种三维声处理技术领域,具体涉及一种后向兼容式三维声编码器、解码器及其编解码方法。

背景技术

[0002] 经过多年发展,立体声、5.1、7.1环绕声等系统已经获得了广泛的应用,这些系统因缺乏声音的高度信息,最多只能呈现二维的声音。在真实的世界中,三维声是声音最真实的呈现和表达方式,无论自然界、艺术领域或视听娱乐领域,三维声都是未来的发展趋势。

[0003] 但是,因立体声、环绕声系统的普遍存在,三维声系统只有最大限度的兼容已经普遍存在的立体声或环绕声系统,才能获得市场的接受,并逐步成为主流。事实上,在近些年发展起来的一些多层多声道三维声技术中,如9.1、11.1、13.1和22.2等,都兼容了立体声、5.1或7.1环绕声的扬声器布局。但是,这只是还音系统的兼容,三维声音源格式与已有立体声、环绕声格式的兼容性还没有合理的解决方案。

发明内容

[0004] 发明目的:本发明目的在于针对现有技术的不足,提供一种后向兼容式三维声编码器、解码器及其编解码方法,实现了兼容已有立体声、环绕声节目或系统的同时具备三维声编解码的能力。

[0005] 技术方案:本发明所述的后向兼容式三维声编码器,包括:下混及基础声道划分模块,用于接收基础声道和/或声音对象,按照下混方案进行下混和基础声道划分操作,输出下混兼容声道、扩展基础声道、基础声道划分边信息;扩展编码模块,用于接收声音对象、下混方案、扩展基础声道、基础声道划分边信息,输出扩展编码数据;码流复用模块,复用所述下混兼容声道和扩展编码数据,输出后向兼容式三维声编码码流。

[0006] 进一步完善上述技术方案,针对下混方案为内部生成的方式,所述下混及基础声道划分模块包括下混模块和基础声道划分模块,下混模块用于接收基础声道和声音对象,输出下混兼容声道和内部生成的下混方案,基础声道划分模块用于接收基础声道、内部生成的下混方案,输出扩展基础声道、基础声道划分边信息。

[0007] 进一步地,针对下混方案采用外部输入的方式,所述下混及基础声道划分模块包括下混模块和基础声道划分模块,下混模块用于接收基础声道、声音对象、外部输入的下混方案,输出下混兼容声道;基础声道划分模块用于接收基础声道、外部输入的下混方案,输出扩展基础声道、基础声道划分边信息。

[0008] 进一步地,所述下混及基础声道划分模块包括下混模块和基础声道划分模块,基础声道划分模块用于接收基础声道和外部输入的下混方案,输出所述扩展基础声道和基础声道划分边信息;所述扩展编码数据经扩展解码模块解码生成解码的基础声道、解码的基础声道划分边信息、解码的下混方案、解码的声音对象与所述基础声道输入至下混模块,下混模块输出所述下混兼容声道。

[0009] 进一步地,所述下混模块按照下混方案将基础声道、声音对象下混为下混兼容声道,下混兼容声道信号分为基础声道下混分量、声音对象下混分量,基础声道下混分量可分为扩展基础声道下混分量、兼容基础声道下混分量;所述基础声道划分模块将基础声道划分为兼容基础声道、扩展基础声道,采用的基础声道划分方案根据基础声道的声道配置、所要兼容的多声道系统及下混映射函数确定,如根据对应下混声道的划分方法确定或者根据基于QR分解的划分方法确定。

[0010] 根据对应下混声道的划分方法确定基础声道划分方案,包括:

[0011] S11:令集合 $S_{bedt} = S_{bed}$, $S_c = \emptyset$, $S_{srt} = S_{sr}$, 基础声道信号集合 $S_{bed} = \{\overline{bedchannel}[i], i = 1 \dots N_b\}$, 下混兼容声道信号集合 $S_{sr} = \{\overline{srchannel}[k], k = 1 \dots N_s\}$, $fb(k, i)$ 为下混映射函数;

[0012] S12:遍历集合 S_{srt} , 找出一个满足如下关系的下混兼容声道 k :

[0013] $\overline{srchannel_bed}[k] == \sum_{i=1}^{N_b} fb(k, i) (\overline{bedchannel}[i])$, 对于所有属于 S_c 中的声道 n , 有 $fb(k, n) = 0$ 。

[0014] 若没有找到,执行步骤S15;

[0015] S13:对于步骤S12中的下混兼容声道 k , 遍历集合 S_{bedt} , 找出 $fb(k, m)$ 不为0且 $fb(k, m)$ 可逆的基础声道 m , 若没有找到,执行步骤S15;

[0016] S14:将步骤S13中找到的基础声道 m 加入到集合 S_c 中, 得到新的 S_c , 从 S_{srt} 中去除下混兼容声道 k , 得到新的 S_{srt} , 从 S_{bedt} 中去除所有满足 $fb(k, i)$ 不为0的基础声道 i , 得到新的 S_{bedt} , 若新的 S_{srt} 、新的 S_{bedt} 都不为空, 跳转到步骤S12, 若新的 S_{srt} 、新的 S_{bedt} 为空, 则执行步骤S15;

[0017] S15: S_c 或 S_c 的子集作为基础声道集合 S_{bed} 的兼容基础声道集合。

[0018] 根据基于QR分解的划分方法确定基础声道划分方案, 包括:

[0019] S21:令 $S_{bedc} = S_{bed}$, S_{bed} 为基础声道信号集合;

[0020] S22:采用矩阵运算形式表达 S_{bedc} 下混: $H_{s_bedcop} = H_{Ac} * H_{Bc}$, H_{s_bedcop} 为 S_{bedc} 下混形成的下混分量组成的矩阵, H_{Bc} 为 S_{bedc} 中基础声道信号组成的矩阵, H_{Ac} 为 S_{bedc} 下混系数矩阵;

[0021] S23:对 H_{Ac} 进行QR分解, 得 $H_{Ac} = Q * H_R$, 其中 Q 为一个 $N_s \times N_s$ 的正交矩阵, H_R 是一个 $N_s \times N_{bc}$ 的上三角矩阵;

[0022] S24:设 $M = \min(N_s, N_{bc})$, N_s 为基础声道下混的声道数、 N_{bc} 为 S_{bedc} 的声道数, 如果对于 H_R 中每一个 $n = 1 \dots M$, 都满足 $r(n, n) > 0$, 执行步骤S25; 否则, 对于每一个 $n = 1 \dots M$, 如果 H_R 中 $r(n, n) = 0$, 将 S_{bedc} 中第 n 个声道从 S_{bedc} 中剔除形成新的集合 S_{bedc}' , 令 $S_{bedc} = S_{bedc}'$ 并执行步骤S22;

[0023] S25:保留 S_{bedc} 中 $n = 1 \dots M$ 的声道组成的集合, 该集合或集合的子集作为基础声道集合 S_{bed} 的兼容基础声道集合。

[0024] 进一步地, 还包括兼容编码模块, 兼容编码模块用于接收所述下混兼容声道, 输出下混兼容编码数据, 所述码流复用模块复用所述下混兼容编码数据和扩展编码数据, 输出后向兼容式三维声编码码流。所述兼容编码模块与所述扩展编码模块采用相同的编码格式进行编码或采用不同的编码格式进行编码。具体而言, 所述扩展编码模块采用压缩编码或

非压缩编码方式进行编码,所述兼容编码模块采用有损或无损方式进行编码。具体而言对扩展基础声道的音频数据的编码和声音对象的编码可以采用非压缩的编码也可以采用压缩的编码方法,可以是矢量编码也可以是标量编码,如杜比AC3、MPEG-1layer3、MPEG-2/4AAC、MPEG H、杜比Atmos、AVS2-P3三维音频编码等编码格式。对声音对象的编码包括对对象渲染描述信息obj_info[j]和对对象信号obj_data[j]的编码。对下混方案进行编码时,可以对下混映射函数fb(k,i)、fo(k,j)的参数a(k,i)、 $\Delta(k,i)$ 等进行有损或无损的编码;也可以将对a(k,i)、 $\Delta(k,i)$ 进行矢量编码。对基础声道划分边信息(比如betob(m)($m=1\cdots Nbe$)、bctob(n)($n=1\cdots Nbc$))进行编码时,可以进行有损或无损的编码;也可以采用矢量编码的方式。

[0025] 采用上述后向兼容式三维编码器进行三维声编码方法,包括如下步骤:

[0026] 按照下混方案将基础声道和/或声音对象下混为所兼容的下混兼容声道,

[0027] 将基础声道划分为扩展基础声道、兼容基础声道并确定基础声道划分边信息;

[0028] 对下混兼容声道进行编码生成下混兼容编码数据;

[0029] 对声音对象、下混方案、扩展基础声道、基础声道划分边信息进行编码得到扩展编码数据;

[0030] 复用下混兼容编码数据和扩展编码数据得到后向兼容式三维声编码码流。

[0031] 用于上述三维编码器输出码流解码的后向兼容式三维声解码器,根据所用还音系统或后端系统的不同,自适应地确定是按照兼容还是三维声的方式进行解码,按照兼容的方式解码时,只包括码流解复用模块和兼容解码模块,得到解码的下混兼容声道信号;按照三维声的方式解码时,还包括扩展解码模块、去下混模块、基础声道组合模块,得到解码的基础声道信号和解码的声音对象,组成解码的三维声节目。码流解复用模块,用于获取后向兼容式三维声编码码流进行解复用得到下混兼容编码数据和扩展编码数据;兼容解码模块,用于接收下混兼容编码数据,输出解码的下混兼容声道;扩展解码模块,用于接收扩展编码数据,输出解码的下混方案、解码的扩展基础声道、解码的基础声道划分边信息、解码的声音对象;去下混模块,用于接收解码的下混兼容声道、解码的下混方案、解码的扩展基础声道、解码的基础声道划分边信息、解码的声音对象,按照解码的下混方案,从解码的下混兼容声道中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分,输出解码的兼容基础声道;基础声道组合模块,用于接收解码的兼容基础声道、解码的扩展基础声道、解码的基础声道划分边信息,输出解码的基础声道。

[0032] 所述去下混模块包括如下步骤:按照解码的下混方案,从解码的下混兼容声道信号中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分,得到解码的兼容基础声道信号。

[0033] 所述去下混模块包括如下步骤:

[0034] 1) 计算兼容基础声道下混分量:按照解码的下混方案,从解码的下混兼容声道信号中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分,得到解码的兼容基础声道下混分量;

[0035] 2) 逆映射:对解码的兼容基础声道下混分量进行逆映射得到解码的兼容基础声道信号。

[0036] 针对根据对应下混声道的划分方法确定的基础声道划分方案进行的下混和基础

声道划分操作,所述去下混模块包括如下步骤:

[0037] 1) 计算兼容基础声道下混分量:对于每一个兼容基础声道 $n=1\cdots Nbc$, $k=\text{compat}(n)$ 为其对应的下混兼容声道,从解码的下混兼容声道信号 $\overrightarrow{\text{srchannel}[k]}$ 中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分得到兼容基础声道下混分量 $\overrightarrow{\text{srchannel_bed}_{cop}[k]}$;

[0038] 2) 逆映射:对于每一个兼容基础声道 $n=1\cdots Nbc$, $k=\text{compat}(n)$ 为其对应的下混兼容声道,按照解码的下混方案中映射函数 $f_b(k, \text{bctob}(n))$ 的逆函数 $f_b^{-1}(k, \text{bctob}(n))$ 对解码的兼容基础声道下混分量 $\overrightarrow{\text{srchannel_bed}_{cop}[k]}$ 进行逆映射,得到解码的兼容基础声道 $\overrightarrow{\text{bedchannel}_{cop}[n]}$;

[0039] $\overrightarrow{\text{bedchannel}_{cop}[n]} = f_b^{-1}(k, \text{bctob}(n))(\overrightarrow{\text{srchannel_bed}_{cop}[k]})$

[0040] 进一步地,针对根据基于QR分解的划分方法确定的基础声道划分方案进行的下混和基础声道划分操作,所述去下混模块包括如下步骤:

[0041] 1) 计算兼容基础声道下混分量:对于每一个下混兼容声道 k ,从解码的下混兼容声道信号 $\overrightarrow{\text{srchannel}[k]}$ 中去除解码的扩展基础声道信号和/或解码的声音对象的下混成分得到兼容基础声道下混分量 $\overrightarrow{\text{srchannel_bed}_{cop}[k]}$,所有的 $\overrightarrow{\text{srchannel_bed}_{cop}[k]}$, $k=1\cdots Ns$ 组成矩阵 Hs_bedcop

[0042]
$$Hs_bedcop = \begin{bmatrix} \overrightarrow{\text{srchannel_bed}_{cop}[1]} \\ \vdots \\ \overrightarrow{\text{srchannel_bed}_{cop}[Ns]} \end{bmatrix}$$

[0043] 2) 逆映射:根据解码的兼容基础声道映射系数逆矩阵 invHRQ ,对 Hs_bedcop 进行逆映射求得 Hbc

[0044] $Hbc = \text{invHRQ} * Hs_bedcop$

[0045] Hbc 中的第 n 行即为解码的兼容基础声道信号 $\overrightarrow{\text{bedchannel}_{cop}[n]}$, $n=1\cdots Nbc$ 。

[0046] 采用上述三维声系统对三维声编码器输出码流进行解码的方法,包括如下步骤:

[0047] 获取后向兼容式三维声编码码流进行解复用得到下混兼容编码数据和扩展编码数据;

[0048] 解码下混兼容编码数据得到解码的下混兼容声道,解码扩展编码数据得到解码的下混方案、解码的扩展基础声道、解码的基础声道划分边信息、解码的声音对象;

[0049] 通过从解码的下混兼容声道中减去扩展基础声道、声音对象的下混成分的方式,解码得到解码的兼容基础声道;

[0050] 对解码的兼容基础声道和扩展基础声道进行声道组合得到解码的基础声道。

[0051] 有益效果:本发明提出的后向兼容式三维声编码器、解码器及其编解码方法,能够更好地兼容已有立体声、环绕声节目或系统;且方案灵活,适用于基础声道或声音对象或基础声道与声音对象的组合等多种情况的三维声信号,可以采用外部输入确定的下混方案也可以根据系统自适应选择的下混方案;在兼容的立体声、环绕声系统中,只需解码得到下混

兼容声道,在三维声系统进行播放时,解码出基础声道信号,能够获得三维声播放效果,从而实现了兼容已有立体声、环绕声节目或系统的同时具备三维声播放的能力;在侧重于提升三维声的质量的应用场合,本发明还提供了改进的三维声编码方法,提升三维声的编码质量,获得更好地播放效果。

附图说明

- [0052] 图1为实施例1的方法流程图;
- [0053] 图2为实施例2的方法流程图;
- [0054] 图3为下混算法原理图;
- [0055] 图4为下混算法原理图;
- [0056] 图5为去下混原理图;
- [0057] 图6为实施例3的方法流程图;
- [0058] 图7为实施例4解码方法流程图;
- [0059] 图8为实施例5的方法流程图。

具体实施方式

[0060] 下面通过附图对本发明技术方案进行详细说明,但是本发明的保护范围不局限于所述实施例。

[0061] 实施例1:

[0062] 三维声信号由多声道信号(即基础声道)和/或声音对象信号(包括对象渲染描述信息和对象音频数据)组成。为满足后向兼容多声道(立体声、环绕声)节目或系统的需要,本申请所提出的三维声编码方法按照下混方案将三维声信号下混为所兼容的下混兼容声道信号,将基础声道划分为扩展基础声道和兼容基础声道;对下混兼容声道信号进行编码生成下混兼容编码数据;对下混方案、扩展基础声道信号、基础声道划分边信息、声音对象信号进行编码得到扩展编码数据;复用下混兼容编码数据和扩展编码数据得到后向兼容式三维声编码码流。

[0063] 本申请所提的三维声解码方法对后向兼容式三维声编码码流进行解复用得到下混兼容编码数据和扩展编码数据;解码下混兼容编码数据得到解码的下混兼容声道信号;解码扩展编码数据得到解码的下混方案、扩展基础声道信号、基础声道划分边信息、声音对象信号;通过从解码的下混兼容声道信号中减去扩展基础声道信号、声音对象信号的下混成分的方式,解码得到解码的兼容基础声道信号;对解码的兼容基础声道信号和扩展基础声道信号进行声道组合得到解码的基础声道信号。在兼容的环绕声系统下,则只对下混兼容编码数据进行解复用和解码,并按照兼容的环绕声系统进行还音,这就实现了对立体声、环绕声系统的兼容。

[0064] 基础声道可以是立体声、5.1、7.1等多声道信号,也可以是9.1、11.1、13.1和22.2等多层多声道三维声信号,基础声道中每一个声道的数据表示为 $\overline{bedchannel}[i], i = 1 \dots Nb$, Nb 为基础声道的声道数量, Nb 等于0时该三维声信号中不含多声道信号而只有声音对象信号,所有基础声道信号组成集合 $S_{bed} = \{\overline{bedchannel}[i], i = 1 \dots Nb\}$;声音对象信号

obj_signal[j]包括对象渲染描述信息obj_info[j]和对象信号 $\overrightarrow{obj_data[j]}$, $j=1 \dots M$, M 为声音对象的数量, M 等于0时该三维声信号中不包含声音对象。声音对象信号可以是单声道、立体声或多声道信号。所兼容的下混兼容声道信号表示为 $\overrightarrow{srchannel[k]}$, $k=1 \dots N_s$, N_s 为所兼容的立体声、环绕声系统的声道数, 所有下混兼容声道信号组成集合 $S_{sr} = \{\overrightarrow{srchannel[k]}, k=1 \dots N_s\}$ 。每一个声道或声音对象的数据 $\overrightarrow{bedchannel[i]}$ 、 $\overrightarrow{obj_data[j]}$ 、 $\overrightarrow{srchannel[k]}$ 是时间序列信号, 即PCM(脉冲编码调制采样)数据, 当进行分帧处理时, 则可以表示一帧内的时间序列信号。

[0065] 按照某一个下混方案, 将三维声信号下混为下混兼容声道信号时, 每一个下混兼容声道信号包括基础声道下混分量和声音对象下混分量:

$$[0066] \quad \overrightarrow{srchannel[k]} = \overrightarrow{srchannel_bed[k]} + \overrightarrow{srchannel_obj[k]}, k = 1 \dots N_s$$

[0067] 下混方案可以表达为一组映射函数 $f_o(k, j)$ 、 $f_b(k, i)$ 。所有下混兼容声道信号的基础声道下混分量组成集合 $S_{sr_bed} = \{\overrightarrow{srchannel_bed[k]}, k=1 \dots N_s\}$, 所有下混兼容声道信号的声音对象下混分量组成集合 $S_{sr_obj} = \{\overrightarrow{srchannel_obj[k]}, k=1 \dots N_s\}$ 。基础声道下混分量由基础声道信号所生成, 可以表示为:

$$[0068] \quad \overrightarrow{srchannel_bed[k]} = \sum_{i=1}^{N_b} f_b(k, i) (\overrightarrow{bedchannel[i]}), k = 1 \dots N_s$$

[0069] $f_b(k, i)$ 是第 i 个基础声道信号向下混兼容声道信号的第 k 个声道进行下混时的下混映射函数。声音对象下混分量是根据所要兼容的下混兼容声道系统对每一个声音对象的信号进行下混渲染所生成, 可以表示为:

$$[0070] \quad \overrightarrow{srchannel_obj[k]} = \sum_{j=1}^M f_o(k, j) (\overrightarrow{obj_data[j]}), k = 1 \dots N_s$$

[0071] $f_o(k, j)$ 是第 j 个对象向下混兼容声道的第 k 个声道进行下混时的下混映射函数, 与对象的位置坐标等信息相关。函数 $f_o(k, j)$ 、 $f_b(k, i)$ 表达的可以是增益、延时等操作(此类操作后面简称PAN方法), 例如 $f_b(k, i)(x(t)) = a(k, i) * x(t - \Delta(k, i))$;也可以是更复杂的映射函数, 如WFS和HOA驱动函数等。WFS(Wave field Synthesis)是一种声音渲染方法, 通过在求解波动方程的时候采用Kirchhoff-Helmholtz积分, 实现用还音扬声器组还原原始声场。HOA(High Order Ambisonic)也是一种声音渲染方法, 通过在求解波动方程的时候采用球面波级数, 实现用扬声器组还原原始声场。WFS和HOA方法具体可参考文献“Comparison of Higher order Ambisonics and Wave Field Synthesis With Respect to Spatial Discretization Artifacts in Time Domain”(Sascha Spors and Jens Ahrens, 19th International Congress on Acoustics Madrid, 2-7 Sept. 2007)。

[0072] 可以将基础声道信号的集合 $S_{bed} = \{\overrightarrow{bedchannel[i]}, i=1 \dots N_b\}$ 进行声道划分, 分为两个集合 S_{bede} 和 S_{bedc} , 满足:

$$[0073] \quad S_{bede} = \{\overrightarrow{bedchannel_{ext}[m]}, m=1 \dots N_{be}\}$$

[0074] $S_{bedc} = \{\overline{bedchannel_{cop}[n]}, n = 1 \dots Nbc\}$

[0075] $Nbe + Nbc = Nb$

[0076] $S_{bede} \cup S_{bedc} = S_{bed}$

[0077] $S_{bede} \cap S_{bedc} = \emptyset$

[0078] 据此,基础声道下混分量 $\overline{srchannel_bed[k]}$ 又可以分为 $\overline{srchannel_bed_{ext}[k]}$ 和 $\overline{srchannel_bed_{cop}[k]}$ 之和,可以表示为:

[0079] $\overline{srchannel_bed[k]} = \overline{srchannel_bed_{cop}[k]} + \overline{srchannel_bed_{ext}[k]}, k = 1 \dots Ns$

[0080] $\overline{srchannel_bed_{cop}[k]} = \sum_{n=1}^{Nbc} fb(k, bctob(n)) (\overline{bedchannel[bctob(n)]}), k = 1 \dots Ns$

[0081] $\overline{srchannel_bed_{ext}[k]} = \sum_{m=1}^{Nbe} fb(k, betob(m)) (\overline{bedchannel[betob(m)]}), k = 1 \dots Ns$

[0082] 其中, $betob(m)$ ($m=1 \dots Nbe$) 是 S_{bede} 中第 m 个声道在基础声道中的序号,

[0083] $\overline{bedchannel_{ext}[m]} = \overline{bedchannel[betob(m)]}$

[0084] $bctob(n)$ ($n=1 \dots Nbc$) 是 S_{bedc} 中第 n 个声道在基础声道中的序号,

[0085] $\overline{bedchannel_{cop}[n]} = \overline{bedchannel[bctob(n)]}$ 。

[0086] $\overline{srchannel_bed_{cop}[k]}$ 由 S_{bed} 集合中的基础声道下混生成,所有 $\overline{srchannel_bed_{cop}[k]}$ 组成集合 $S_{srbed_{cop}} = \{\overline{srchannel_bed_{cop}[k]}, k = 1 \dots Ns\}$ 。

[0087] 如果能根据集合 $S_{srbed_{cop}}$ 和下混映射函数 $fb(k, i)$ 、基础声道划分边信息 ($betob(m)$ 、 $bctob(n)$) 等运算得到 S_{bedc} 中的所有信号,则称 S_{bedc} 为兼容基础声道集合,其中的声道信号 $\overline{bedchannel_{cop}[n]}$ 称为兼容基础声道; S_{bede} 称为与 S_{bedc} 对应的扩展基础声道集合, $\overline{bedchannel_{ext}[m]}$ 称为扩展基础声道。 $\overline{srchannel_bed_{ext}[k]}$ 称为扩展基础声道下混分量; $S_{srbed_{cop}}$ 称为兼容基础声道下混分量集合, $\overline{srchannel_bed_{cop}[k]}$ 称为兼容基础声道下混分量。

[0088] 对于基础声道集合 S_{bed} 来说,其兼容基础声道集合 S_{bedc} 与扩展基础声道集合 S_{bede} 的划分方式和划分准则可能有多种,显然, $S_{bede} = S_{bed}$, $S_{bedc} = \emptyset$ 的划分符合上述定义;如果 S_{bedc1} 是 S_{bed} 的兼容基础声道集合,则 S_{bedc1} 的任意子集 S_{bedc2} 也是 S_{bed} 的兼容基础声道集合。

[0089] 如图1所示,后向兼容式三维声编码器包括下混及基础声道划分模块、兼容编码模块、扩展编码模块、码流复用模块,其编码方法流程如下:

[0090] 步骤1.1) 下混及基础声道划分模块

[0091] 根据外界输入或系统自适应选择的下混方案(比如可以以一组映射函数 $fb(k, i)$ 、 $fo(k, j)$ 来表达),将三维声节目下混为所兼容的立体声/多声道信号,得到下混兼容声道信

号 $\overline{srchannel[k]}$, $k = 1 \dots N_s$;将基础声道划分为兼容基础声道和扩展基础声道两部分,并确定基础声道划分边信息(如betob(m)、bctob(n));输出所用下混方案。

[0092] “外界输入”一般指录音师在三维声节目缩混时人工选择的下混方案,这种方式允许录音师通过反复监听,比较和选择下混方案;“系统自适应”是指编码系统智能的选择一种下混的方案,如对于基础声道的下混,可以自适应地根据不同的基础声道系统与下混兼容多声道系统之间的扬声器布局的关系,将多层多声道系统中表达高度信息的声道层的信号按照位置关系下混到中间层的环绕声声道上(比如顶前左声道直接混入左声道,顶前右声道直接混入右声道),形成一种下混方案(可以以一组映射函数fb(k,i)来表达);如对于声音对象的下混,可以自适应地根据对象渲染描述信息(对象的位置坐标等),按照WFS/HOA/PAN等渲染方法,形成一种下混方案(可以以一组映射函数fo(k,j)来表达)。

[0093] 如果该三维声节目中不包括基础声道,或兼容基础声道、扩展基础声道与基础声道之间是简单明确的一一对应关系,则可以不需要进行该步骤1.1。

[0094] 步骤1.2) 扩展编码模块

[0095] 对扩展基础声道、基础声道划分边信息、下混方案和声音对象进行扩展编码,得到扩展编码数据。如果该三维节目没有基础声道,则可不包括对扩展基础声道、基础声道划分边信息进行编码的部分;如果该三维节目没有声音对象,则可不包括对声音对象进行编码的部分。如果在编码、解码中都遵循同样的可确定的基础声道划分方式,则可对基础声道划分边信息进行编码。如果在编码、解码中都遵循同样的可确定的下混方案,则可对下混方案进行编码。

[0096] 对扩展基础声道的音频数据的编码和声音对象的编码可以采用非压缩的编码也可以采用压缩的编码方法,可以是矢量编码也可以是标量编码,如杜比AC3、MPEG-1layer3、MPEG-2/4AAC、MPEG H、杜比Atmos、AVS2-P3三维音频编码等编码格式。对声音对象的编码包括对对象渲染描述信息obj_info[j]和对象信号obj_data[j]的编码。

[0097] 对下混方案进行编码时,可以对下混映射函数fb(k,i)、fo(k,j)的参数a(k,i)、 $\Delta(k,i)$ 等进行有损或无损的编码;也可以对将a(k,i)、 $\Delta(k,i)$ 进行矢量编码。

[0098] 对基础声道划分边信息(比如betob(m) ($m=1 \dots N_{be}$)、bctob(n) ($n=1 \dots N_{bc}$))进行编码时,可以进行有损或无损的编码;也可以采用矢量编码的方式。

[0099] 步骤1.3) 兼容编码模块

[0100] 对下混兼容声道信号进行编码,得到下混兼容编码数据。

[0101] 对下混兼容声道信号的编码可以采用非压缩的编码也可以采用压缩的编码方法,可以是矢量编码也可以是标量编码,如杜比AC3、MPEG-1layer3、MPEG-2/4AAC、AVS等编码格式,为满足兼容性需要,所用的编码方法应符合所兼容的多声道系统的要求。如果不需要对下混兼容声道信号进行编码,可以采用直通的方式,此时下混兼容编码数据与下混兼容声道信号一致,可以不需要步骤1.3) 兼容编码模块。

[0102] 需要说明的是,兼容编码模块的具体编码方法和扩展编码模块的具体编码方法,可以采用相同的编码格式,也可以采用不同的编码格式。

[0103] 步骤1.4) 码流复用模块

[0104] 复用下混兼容编码数据和扩展编码数据,得到后向兼容的三维声编码码流。后向兼容的三维声编码码流应符合所兼容的多声道系统的码流格式。

[0105] 实施例2:

[0106] 如图2所示的后向兼容式三维声编码器,包括下混模块、基础声道划分模块、兼容编码模块、扩展编码模块、码流复用模块。

[0107] 步骤2.1) 下混模块

[0108] 根据外界输入或系统自适应选择的下混方案,将三维声节目下混为所兼容的立体声/多声道信号,得到下混兼容声道信号 $\overrightarrow{srchannel[k]}$;输出所用下混方案。

[0109] 如前所述,下混兼容声道信号包括基础声道下混分量和声音对象下混分量:

$$[0110] \quad \overrightarrow{srchannel[k]} = \overrightarrow{srchannel_bed[k]} + \overrightarrow{srchannel_obj[k]}, k = 1 \dots Ns$$

[0111] 基础声道下混分量由每一个基础声道的信号所生成,可以按照如下公式计算每一个下混兼容声道k的基础声道下混分量:

$$[0112] \quad \overrightarrow{srchannel_bed[k]} = \sum_{i=1}^{Nb} fb(k, i) (\overrightarrow{bedchannel[i]}), k = 1 \dots Ns$$

[0113] $fb(k, i)$ 是第*i*个基础声道信号向下混多声道信号的第*k*个声道进行下混时的下混映射函数。

[0114] 声音对象下混分量是根据所要兼容的多声道系统对每一个声音对象的信号进行下混渲染所生成,可以根据如下公式计算每一个下混兼容声道k的声音对象下混分量:

$$[0115] \quad \overrightarrow{srchannel_obj[k]} = \sum_{j=1}^M fo(k, j) (\overrightarrow{obj_data[j]}), k = 1 \dots Ns$$

[0116] $fo(k, j)$ 是第*j*个对象向下混多声道信号的第*k*个声道进行下混时的下混映射函数。这种情况下下混运算的示意图见图3。

[0117] 基础声道下混分量 $\overrightarrow{srchannel_bed[k]}$ 又可以分为扩展基础声道下混分量 $\overrightarrow{srchannel_bed_cop[k]}$ 和兼容基础声道下混分量 $\overrightarrow{srchannel_bed_ext[k]}$:

$$[0118] \quad \overrightarrow{srchannel_bed[k]} = \overrightarrow{srchannel_bed_cop[k]} + \overrightarrow{srchannel_bed_ext[k]}, k = 1 \dots Ns$$

[0119] 可以分别按如下公式计算扩展基础声道下混分量 $\overrightarrow{srchannel_bed_ext[k]}$ 和兼容基础声道下混分量 $\overrightarrow{srchannel_bed_cop[k]}$:

$$[0120] \quad \overrightarrow{srchannel_bed_cop[k]} = \sum_{n=1}^{Nbc} fb(k, bctob(n)) (\overrightarrow{bedchannel[bctob(n)]}), k = 1 \dots Ns$$

$$[0121] \quad \overrightarrow{srchannel_bed_ext[k]} = \sum_{m=1}^{Nbe} fb(k, betob(m)) (\overrightarrow{bedchannel[betob(m)]}), k = 1 \dots Ns$$

[0122] 这种情况下下混运算的示意图见图4。

[0123] 步骤2.2) 基础声道划分模块

[0124] 根据确定或选择的基础声道划分方案,将基础声道划分为兼容基础声道和扩展基础声道两部分,并确定基础声道划分边信息(如betob(*m*)、bctob(*n*)等)。基础声道划分方案

可以由外界输入,也可以根据下混模块所用的下混方案(比如可以以一组映射函数 $fb(k, i)$ 、 $fo(k, j)$ 来表达)自适应地确定。一般地,对于某一种三维声系统与其所要兼容的多声道系统,可以确定采用固定的一种基础声道划分方案,不需要每次都进行确定基础声道划分方案的运算。

[0125] 如果该三维声节目中不包括基础声道,或兼容基础声道、扩展基础声道与基础声道之间是简单明确的一一对应关系,则可以不需要进行该步骤2.2。

[0126] 确定基础声道划分方案的方法可以有多种,下面以两种情况为例来对基础声道划分方法进行详细说明。

[0127] 1) 对应下混声道的划分方法:

[0128] 当三维声节目的基础声道的声道配置、所要兼容的多声道系统及下混映射函数 $fb(k, i)$ 都已确定时,可以根据如下划分规则将基础声道 $Sbed = \{\overline{bedchannel[i]}, i = 1 \dots N\}$ 分为 $Sbede = \{\overline{bedchannel_{ext}[m]}, m = 1 \dots N\}$ 、 $Sbedc = \{\overline{bedchannel_{cop}[n]}, n = 1 \dots Nbc\}$ 两部分:

[0129] 对于每一个 $n = 1 \dots Nbc$,存在 $k = compat(n)$,及 $n = invcompat(k)$,其中 $compat(n)$ 是第 n 个兼容基础声道所对应的下混兼容声道的序号,有

$$\begin{aligned} \overline{srchannel_{bed}[k]} &== \sum_{i=1}^{Nb} fb(k, i) (\overline{bedchannel[i]}) \\ [0130] \quad &== fb(k, bctob(n)) (\overline{bedchannel_{cop}[n]}) + \sum_{m=1}^{Nbe} fb(k, betob(m)) (\overline{bedchannel_{ext}[m]}) \end{aligned}$$

[0131] 且 $fb(k, bctob(n))$ 存在逆函数。

[0132] 在满足上述规则的情况下,可以按照如下公式来计算得到解码的兼容基础声道信号 $\overline{bedchannel_{cop}[n]}$,如图5所示,

[0133] 对于每一个 n ,存在 $k = compat(n)$

$$[0134] \quad \overline{srchannel_{bed_{cop}}[k]} = \overline{srchannel_{bed}[k]} - \sum_{m=1}^{Nbe} fb(k, betob(m)) (\overline{bedchannel_{ext}[m]})$$

$$[0135] \quad \overline{bedchannel_{cop}[n]} = fb^{-1}(k, bctob(n)) (\overline{srchannel_{bed_{cop}}[k]})$$

[0136] 即,集合 $Sbedc = \{\overline{bedchannel_{cop}[n]}, n = 1 \dots Nbc\}$ 可以根据集合 $Ssrbed_{cop} = \{\overline{srchannel_{bed_{cop}}[k]}, k = 1 \dots Nbc\}$ 和下混映射函数 $fb(k, i)$ 、基础声道划分边信息($betob(m)$ 、 $bctob(n)$)等运算得到,满足前述对兼容基础声道集合的定义,所以 $Sbedc$ 为 $Sbed$ 的兼容基础声道集合。这种情况下的特征是:兼容基础声道 n 可以通过对其所对应的下混兼容声道 k 的兼容基础声道下混分量 $\overline{srchannel_{bed_{cop}}[k]}$ 进行逆映射得到,逆映射的函数为下混映射函数的逆函数。

[0137] 这种情况下,基础声道划分边信息包括 $betob(m)$ ($m = 1 \dots Nbe$)、 $bctob(n)$ ($n = 1 \dots Nbc$)、 $compat(n)$ ($n = 1 \dots Nbc$)、 $fb^{-1}(k, bctob(n))$ 。其中, $compat(n)$ 是第 n 个兼容基础声道所

对应的下混兼容声道的序号, $fb^{-1}(k, bctob(n))$ 为下混映射函数逆函数, 可以通过求解下混映射函数的逆函数的方法求得。 $fb(k, bctob(n))$ 存在逆函数 $fb^{-1}(k, bctob(n))$, 例如

$$[0138] \quad fb(k, i)(x(t)) = a(k, i) * x(t - \Delta(k, i))$$

$$[0139] \quad fb^{-1}(k, i)(x(t)) = x(t + \Delta(k, i)) / a(k, i)$$

[0140] 特殊情况下,

$$[0141] \quad fb(k, i)(x(t)) = x(t)$$

$$[0142] \quad fb^{-1}(k, i)(x(t)) = x(t)$$

[0143] 例如基础声道为5.1.4 (5.1+4H, 5.1环绕声系统加上4个顶部音箱), 所兼容的多声道系统为5.1时, 假定5.1.4的声道顺序为左声道(1)、右声道(2)、中央声道(3)、重低音声道(4)、左环绕声道(5)、右环绕声道(6)、顶前左声道(7)、顶前右声道(8)、顶后左声道(9)、顶后右声道(10), 5.1的声道顺序为左声道(1)、右声道(2)、中央声道(3)、重低音声道(4)、左环绕声道(5)、右环绕声道(6), 下混算法按照顶前左声道(7)、顶后左声道(9)直接混入左声道, 顶前右声道(8)、顶后右声道(10)直接混入右声道的方式进行, 声道映射函数简化为:

$$[0144] \quad betob(m) = m + 6$$

$$[0145] \quad bctob(n) = n$$

[0146] 兼容基础声道与下混兼容声道的对应关系 $compat(n)$ 简化为:

$$[0147] \quad compat(n) = n$$

[0148] 基础声道的下混映射函数简化为

$$[0149] \quad fb(k, i)(x(t)) = x(t)$$

[0150] 此时可以将基础声道分为扩展基础声道{顶前左声道(7)、顶前右声道(8)、顶后左声道(9)、顶后右声道(10)}和兼容基础声道{左声道(1)、右声道(2)、中央声道(3)、重低音声道(4)、左环绕声道(5)、右环绕声道(6)}两部分。

[0151] 比如, 可以按照如下步骤进行基础声道划分:

[0152] 步骤2.2a.1): 令集合 $Sbedt = Sbed$, $S_c = \emptyset$, $Ssrt = Ssr$ 。

[0153] 步骤2.2a.2): 遍历集合 $Ssrt$, 直到找到一个下混兼容声道 k 满足如下关系或遍历结束:

$$[0154] \quad \overrightarrow{srchannel_bed[k]} = \sum_{i=1}^{Nb} fb(k, i) \left(\overrightarrow{bedchannel[i]} \right)$$

[0155] 对于所有属于 S_c 中的声道 n , 有 $fb(k, n) = 0$ 。

[0156] 如果找不到这样的下混兼容声道 k , 则跳转到步骤2.2a.5。

[0157] 步骤2.2a.3): 对于步骤2.2a.2中找到的下混兼容声道 k , 从 $Sbedt$ 中选择一个基础声道 m , 使得 $fb(k, m)$ 不为0, 且 $fb(k, m)$ 可逆, 且基础声道 m 不在 S_e 中。如果找不到这样的基础声道 m , 则跳转到步骤2.2a.5。

[0158] 步骤2.2a.4): 从 $Sbedt$ 中去除所有满足 $fb(k, i)$ 不为0的基础声道 i , 得到新的 $Sbedt$; 将步骤2.2a.3)中找到的基础声道 m 加入到 S_c 中, 得到新的 S_c ; 从 $Ssrt$ 中去除下混兼容声道 k , 得到新的 $Ssrt$ 。如果 $Ssrt$ 、 $Sbedt$ 都不为空, 则跳转到步骤2.2a.2。

[0159] 步骤2.2a.5): 选择 S_c 或 S_c 的子集为兼容基础声道集合 $Sbedc$, $Sbede = Sbed - Sbedc$ 为扩展基础声道集合。

[0160] 2) 基于QR分解的划分方法

[0161] 如果下混函数可以表达为 $f_b(k, i)(x(t)) = a(k, i) * (x(t))$ ($a(k, i)$ 为实数), 基础声道的下混可以表达为矩阵运算的形式:

[0162] $\overrightarrow{srchannel_bed[k]}$ 组成下混多声道信号矩阵 Hs_bed

$$[0163] \quad Hs_bed = \begin{bmatrix} \overrightarrow{srchannel_bed[1]} \\ \vdots \\ \overrightarrow{srchannel_bed[Ns]} \end{bmatrix}$$

[0164] $\overrightarrow{bedchannel[k]}$ 组成基础声道信号矩阵 Hb

$$[0165] \quad Hb = \begin{bmatrix} \overrightarrow{bedchannel[1]} \\ \vdots \\ \overrightarrow{bedchannel[Nb]} \end{bmatrix}$$

[0166] $a(k, i)$ 组成基础声道下混系数矩阵 HA , HA 为 $Ns * Nb$ 矩阵

$$[0167] \quad HA = \begin{bmatrix} a(1,1) & \dots & a(1, Nb) \\ \vdots & \ddots & \vdots \\ a(Ns, 1) & \dots & a(Ns, Nb) \end{bmatrix}$$

[0168] $Hs_bed = HA * Hb$

[0169] 此时, 可以按照如下步骤进行基础声道的划分:

[0170] 步骤2.2b.1): 令 $Sbedc = Sbed$;

[0171] 步骤2.2b.2): 由 $Sbedc$ 中所有的声道信号 $\overrightarrow{bedchannel_cop[n]}$, $n = 1 \dots Nbc$ 组成矩阵

$$[0172] \quad Hbc = \begin{bmatrix} \overrightarrow{bedchannel[1]} \\ \vdots \\ \overrightarrow{bedchannel[Nbc]} \end{bmatrix}$$

[0173] 他们所对应的 $a(k, i)$ 组成兼容基础声道下混系数矩阵 HAc , HAc 为 $Ns * Nbc$ 矩阵

$$[0174] \quad HAc = \begin{bmatrix} a(1,1) & \dots & a(1, Nbc) \\ \vdots & \ddots & \vdots \\ a(Ns, 1) & \dots & a(Ns, Nbc) \end{bmatrix}$$

[0175] 他们下混形成的兼容基础声道下混分量集合 $Ssrbed_cop$ 的所有 $\overrightarrow{srchannel_bed_cop[k]}$ 组成矩阵

$$[0176] \quad Hs_bedcop = \begin{bmatrix} \overrightarrow{srchannel_bedcop[1]} \\ \vdots \\ \overrightarrow{srchannel_bedcop[Ns]} \end{bmatrix}$$

[0177] 满足: $Hs_bedcop = HAc * Hbc$

[0178] 步骤2.2b.3): 对 HAc 进行QR分解, 可得 $HAc = Q * HR$, 其中 Q 为一个 $Ns \times Ns$ 的正交矩阵, HR 是一个 $Ns \times Nbc$ 的上三角矩阵。

$$[0179] \quad HR = \begin{bmatrix} r(1,1) & \dots & r(1,Nbc) \\ \vdots & \ddots & \vdots \\ 0 & \dots & r(Ns,Nbc) \end{bmatrix}$$

[0180] 步骤2.2b.4): 设M为Ns和Nbc中的最小值, $M = \min(Ns, Nbc)$, 如果对于每一个 $n = 1 \dots M$ 都满足 $r(n, n) > 0$, 则执行2.2b.5; 否则进行如下操作: 对于每一个 $n = 1 \dots M$, 如果 $r(n, n) = 0$ 则将Sbedc中第n个声道 $\overrightarrow{bedchannel_{cop}[n]}$ 从Sbedc中剔除, 如此, 形成新的集合Sbedc', 令Sbedc = Sbedc' 并跳转到步骤2.2b.2。

[0181] 步骤2.2b.5): 只保留Sbedc中 $n = 1 \dots M$ 的声道, 组成新的Sbedc, $Nbc = M \leq Ns$; 也可以Sbedc的任意子集作为新的Sbedc。Sbedc为兼容基础声道集合, Sbede = Sbed - Sbedc为扩展基础声道集合。

[0182] 通过以上步骤运算得到Sbedc, 由Sbedc中所有的声道信号 $\overrightarrow{bedchannel_{cop}[n]}$, $n = 1 \dots Nbc$ 组成矩阵

$$[0183] \quad Hbc = \begin{bmatrix} \overrightarrow{bedchannel[1]} \\ \vdots \\ \overrightarrow{bedchannel[Nbc]} \end{bmatrix}$$

[0184] 他们所对应的 $a(k, i)$ 组成下混系数矩阵HAc, HAc为Ns*Nbc矩阵

$$[0185] \quad HAC = \begin{bmatrix} a(1,1) & \dots & a(1,Nbc) \\ \vdots & \ddots & \vdots \\ a(Ns, 1) & \dots & a(Ns, Nbc) \end{bmatrix}$$

[0186] 他们下混形成的兼容基础声道下混分量集合Ssrbed_{cop}中的信号 $\overrightarrow{srchannel_bed_{cop}[k]}$ 组成矩阵

$$[0187] \quad Hs_bedcop = \begin{bmatrix} \overrightarrow{srchannel_bedcop[1]} \\ \vdots \\ \overrightarrow{srchannel_bedcop[Ns]} \end{bmatrix}$$

[0188] 满足: $Hs_bedcop = HAC * Hbc$ 。对HAc进行QR分解的结果是 $HAC = Q * HR$, 其中Q为一个 $Ns \times Ns$ 的酉矩阵, HR是一个 $Ns \times Nbc$ 的上三角矩阵:

$$[0189] \quad HR = \begin{bmatrix} r(1,1) & \dots & r(1,Nbc) \\ \vdots & \ddots & \vdots \\ 0 & \dots & r(Ns,Nbc) \end{bmatrix}$$

[0190] $M = Nbc \leq Ns$, 对于每一个 $n = 1 \dots M$ 都满足 $r(n, n) > 0$ 。

[0191] 可以证明通过以上步骤运算得到的Sbedc为Sbed的兼容基础声道集合:

$$[0192] \quad Hs_bedcop = HAC * Hbc = Q * HR * Hbc$$

[0193] 因为Q为一个 $Ns \times Ns$ 的酉矩阵, 所以其存在逆矩阵 Q^{-1} , 可得

$$[0194] \quad HR * Hbc = Q^{-1} * Hs_bedcop = QHs_bedcop$$

[0195] 因为 $M = \min(Ns, Nbc) = Nbc$, 且对每一个 $n = 1 \dots M$ 都有 $r(n, n) > 0$, 所以可以取HR的前Nbc行构造方阵HRm

$$[0196] \quad HR_m = \begin{bmatrix} r(1,1) & \dots & r(1,Nbc) \\ \vdots & \ddots & \vdots \\ 0 & \dots & r(M,Nbc) \end{bmatrix}$$

[0197] 方阵 HR_m 是上三角矩阵且每一个 $n=1 \cdots M$ 都有 $r(n,n) > 0$,所以 HR_m 存在逆矩阵 HR_m^{-1} 。取 Q^{-1} 的前 Nbc 行构造矩阵 $invQ_m$,满足

$$[0198] \quad HR_m * Hbc = invQ_m * Hs_bedcop$$

[0199] 所以有:

$$[0200] \quad Hbc = HR_m^{-1} * HR_m * Hbc = HR_m^{-1} * invQ_m * Hs_bedcop$$

[0201] 令 $invHRQ = HR_m^{-1} * invQ_m$,有

$$[0202] \quad Hbc = invHRQ * Hs_bedcop$$

[0203] 即 $Sbedc$ 中所有的声道信号 $\overline{bedchannel_{cop}[n]}$, $n = 1 \dots Nbc$ 都可以根据集合 $Ssrbecop$ 和 $fb(k, i)$ 、 $betob(m)$ 、 $bctob(n)$ 运算得到,满足前述对兼容基础声道集合的定义,所以 $Sbedc$ 为 $Sbed$ 的兼容基础声道集合。这种情况下的特征是:兼容基础声道组成的矩阵可以通过对所有兼容基础声道下混分量所组成的矩阵进行逆映射求得,逆映射的过程可以通过矩阵 $invHRQ$ 来表达。这种情况下,基础声道划分边信息包括 $betob(m)$ ($m=1 \cdots Nbe$)、 $bctob(n)$ ($n=1 \cdots Nbc$)、 $invHRQ$,将 $invHRQ$ 称为兼容基础声道映射系数逆矩阵,它可以通过 $fb(k, i)$ 、 $betob(m)$ 、 $bctob(n)$ 等信息计算得到。

[0204] 更宽泛的范围内,如果 $fb(k, i)(x(t)) = a(k, i) * fb(k, 1)(x(t))$,且 $fb(k, 1)$ 存在逆函数 $fb^{-1}(k, 1)$,则也可以按照上述基于QR分解的划分方法进行基础声道划分,这种情况下在解码端通过矩阵运算得到 Hbc 后,还需要对每个解码的声道信号进行 $fb^{-1}(k, 1)$ 的操作,以得到最终的解码的兼容基础声道信号。

[0205] 步骤2.3) 扩展编码模块

[0206] 对扩展基础声道、基础声道划分边信息、下混方案和声音对象进行扩展编码,得到扩展编码数据。如果该三维节目没有基础声道,则可不包括对扩展基础声道、基础声道划分边信息进行编码的部分;如果该三维节目没有声音对象,则可不包括对声音对象进行编码的部分。如果在编码、解码中都遵循同样的可确定的基础声道划分方式,则可对基础声道划分边信息进行编码。如果在编码、解码中都遵循同样的可确定的下混方案,则可对下混方案进行编码。

[0207] 对扩展基础声道的音频数据的编码和声音对象的编码可以采用非压缩的编码也可以采用压缩的编码方法,可以是矢量编码也可以是标量编码,如杜比AC3、MPEG-1layer3、MPEG-2/4AAC、AVS、杜比Atmos、AVS三维音频编码等编码方式。对声音对象的编码包括对对象渲染描述信息 $obj_info[j]$ 和对象信号 $obj_data[j]$ 的编码。

[0208] 对下混方案进行编码时,可以对下混映射函数 $fb(k, i)$ 、 $fo(k, j)$ 的参数 $a(k, i)$ 、 $\Delta(k, i)$ 等进行有损或无损的编码;也可以对将 $a(k, i)$ 、 $\Delta(k, i)$ 进行矢量编码。

[0209] 对基础声道划分边信息进行编码时,分别对 $betob(m)$ ($m=1 \cdots Nbe$)、 $bctob(n)$ ($n=1 \cdots Nbc$)、 $compat(n)$ ($n=1 \cdots Nbc$)、下混映射函数逆函数 $fb^{-1}(k, bctob(n))$ 或兼容基础声道映射系数逆矩阵 $invHRQ$ 进行有损或无损的编码;也可以采用矢量编码的方式。下混映射函数逆函数 $fb^{-1}(k, bctob(n))$ 、兼容基础声道映射系数逆矩阵 $invHRQ$ 可以根据下混映射函数等信息计算得到,所以也可以不对其进行编码,在解码端可通过采用与前述相关过程中同

样的计算方式计算得到。如果编解码端都遵循同样的基础声道划分方案,则可不对基础声道划分边信息进行编码。

[0210] 步骤2.4) 兼容编码模块

[0211] 对下混兼容声道信号进行编码,得到下混兼容编码数据。

[0212] 对下混兼容声道信号的编码可以采用非压缩的编码也可以采用压缩的编码方法,可以是矢量编码也可以是标量编码,如杜比AC3、MPEG-1layer3、MPEG-2/4AAC、AVS等编码方式,为满足兼容性需要,所用的编码方法应符合所兼容的多声道系统的要求。

[0213] 步骤2.5) 码流复用模块

[0214] 复用下混兼容编码数据和扩展编码数据,得到后向兼容的三维声编码码流。后向兼容的三维声编码码流应符合所兼容的多声道系统的码流格式。

[0215] 步骤2.3) 和步骤2.4) 之间、步骤2.2) 和步骤2.3) 之间可以没有先后顺序的约束。

[0216] 实施例3:后向兼容式三维声编码方法1.2

[0217] 如图6所示,后向兼容式三维声编码器包括基础声道划分模块、下混模块、兼容编码模块、扩展编码模块、码流复用模块。

[0218] 步骤3.1) 基础声道划分模块

[0219] 根据确定或选择的基础声道划分方案,将基础声道划分为兼容基础声道和扩展基础声道两部分,并确定基础声道划分边信息(如betob(m)、bctob(n)等)。基础声道划分方案可以由外部输入,也可以根据下混模块所用的下混方案(比如可以以一组映射函数fb(k,i)、fo(k,j)来表达)自适应地生成。一般地,对于某一种三维声系统与其所要兼容的多声道系统,可以确定采用固定的一种基础声道划分方案,不需要每次都进行确定基础声道划分方案的运算。

[0220] 如果该三维声节目中不包括基础声道,或兼容基础声道、扩展基础声道与基础声道之间是简单明确的一一对应关系,则可以不需要进行该步骤3.1)。

[0221] 步骤3.2) 下混模块

[0222] 根据确定或选择所用的下混方案(比如可以以一组映射函数fb(k,i)、fo(k,j)来表达),将三维声节目下混为所兼容的立体声/多声道信号,得到下混兼容声道信号 $\overrightarrow{srchannel}[k]$ 。

[0223] 如前所述,下混兼容声道信号包括基础声道下混分量和声音对象下混分量:

$$[0224] \quad \overrightarrow{srchannel}[k] = \overrightarrow{srchannel_bed}[k] + \overrightarrow{srchannel_obj}[k], k = 1 \dots Ns$$

[0225] 基础声道下混分量由每一个基础声道的信号所生成,可以按照如下公式计算每一个下混兼容声道k的基础声道信号下混分量:

$$[0226] \quad \overrightarrow{srchannel_bed}[k] = \sum_{i=1}^{Nb} fb(k,i) (\overrightarrow{bedchannel}[i]), k = 1 \dots Ns$$

[0227] fb(k,i)是第i个基础声道信号向下混多声道信号的第k个声道进行下混时的下混映射函数。

[0228] 声音对象下混分量是根据所要兼容的多声道系统对每一个声音对象的信号进行渲染所生成,可以根据如下公式计算每一个下混兼容声道k的声音对象下混分量:

$$[0229] \quad \overrightarrow{srchannel_obj[k]} = \sum_{j=1}^M f_o(k, j) (\overrightarrow{obj_data[j]}), k = 1 \dots N_s$$

[0230] $f_o(k, j)$ 是第 j 个对象向下混多声道信号的第 k 个声道进行下混时的下混映射函数。这种情况下下混运算的示意图见图3。

[0231] 基础声道下混分量 $\overrightarrow{srchannel_bed[k]}$ 又可以分为扩展基础声道下混分量 $\overrightarrow{srchannel_bed_{cop}[k]}$ 和兼容基础声道下混分量 $\overrightarrow{srchannel_bed_{ext}[k]}$ ：

$$[0232] \quad \overrightarrow{srchannel_bed[k]} = \overrightarrow{srchannel_bed_{cop}[k]} + \overrightarrow{srchannel_bed_{ext}[k]}, k = 1 \dots N_s$$

[0233] 可以分别按如下公式计算扩展基础声道下混分量 $\overrightarrow{srchannel_bed_{ext}[k]}$ 和兼容基础声道下混分量 $\overrightarrow{srchannel_bed_{cop}[k]}$ ：

$$[0234] \quad \overrightarrow{srchannel_bed_{cop}[k]} = \sum_{n=1}^{Nbc} f_b(k, bctob(n)) (\overrightarrow{bedchannel[bctob(n)]}), k = 1 \dots N_s$$

$$[0235] \quad \overrightarrow{srchannel_bed_{ext}[k]} = \sum_{m=1}^{Nbe} f_b(k, betob(m)) (\overrightarrow{bedchannel[betob(m)]}), k = 1 \dots N_s$$

[0236] 这种情况下下混运算的示意图见图4。

[0237] 步骤3.3) 扩展编码模块

[0238] 对扩展基础声道、基础声道划分边信息、下混方案和声音对象进行扩展编码，得到扩展编码数据。如果该三维节目没有基础声道，则可不包括对扩展基础声道、基础声道划分边信息进行编码的部分；如果该三维节目没有声音对象，则可不包括对声音对象进行编码的部分。如果在编码、解码中都遵循同样的可确定的基础声道划分方式，则可对基础声道划分边信息进行编码。如果在编码、解码中都遵循同样的可确定的下混方案，则可对下混方案进行编码。

[0239] 步骤3.4) 兼容编码模块

[0240] 对下混兼容声道信号进行编码，得到下混兼容编码数据。

[0241] 步骤3.5) 码流复用模块

[0242] 复用下混兼容编码数据和扩展编码数据，得到后向兼容的三维声编码码流。后向兼容的三维声编码码流应符合所兼容的多声道系统的码流格式。

[0243] 步骤3.3) 和步骤3.4) 之间、步骤3.2) 和步骤3.3) 之间可以没有先后顺序的约束。

[0244] 实施例4: 后向兼容立体声/多声道的三维声解码方法

[0245] 如图7所示，后向兼容立体声/多声道的三维声解码器包括码流解复用模块、兼容解码模块、扩展解码模块、去下混模块、基础声道组合模块。

[0246] 步骤4.1) 码流解复用模块

[0247] 对后向兼容环绕声的三维声编码码流进行解复用，得到下混兼容编码数据和扩展编码数据。

[0248] 步骤4.2) 兼容解码模块

[0249] 解码下混兼容编码数据,得到解码的下混兼容声道信号。

[0250] 步骤4.3) 扩展解码模块

[0251] 解码扩展编码数据,得到解码的声音对象、解码的扩展基础声道信号、解码的下混方案和解码的基础声道划分边信息。如果在编码、解码中遵循确定的同样的下混方案、基础声道划分边信息,则可以不对基础声道划分边信息、下混方案进行解码,而是按照确定的规则生成。

[0252] 扩展解码的过程是前述扩展编码的逆过程。

[0253] 步骤4.4) 去下混模块

[0254] 按照解码的下混方案,从解码的下混兼容声道信号中去除解码的扩展基础声道信号、解码的声音对象的下混成分,并进行逆映射得到解码的兼容基础声道信号。分为两个步骤:

[0255] 步骤4.4.1): 计算兼容基础声道下混分量

[0256] 按照解码的下混方案,对解码的扩展基础声道信号、解码的声音对象进行下混,从解码的下混兼容声道信号中去除解码的扩展基础声道信号、解码的声音对象的下混成分,得到解码的兼容基础声道下混分量;

[0257] 步骤4.4.2): 逆映射

[0258] 对解码的兼容基础声道下混分量进行逆映射得到解码的兼容基础声道信号。

[0259] 如果逆映射的过程是简单的直通关系,即解码的兼容基础声道信号与解码的兼容基础声道信号相同,这时则不需要步骤4.4.2): 逆映射。

[0260] 下面以与编码端相对应的两种情况为例对去下混过程进行详细说明。

[0261] 1) 如果编码过程中使用的是对应下混声道的划分方法:

[0262] 此时,解码的基础声道划分边信息包括betob(m) ($m=1 \cdots N_{be}$)、bctob(n) ($n=1 \cdots N_{bc}$)、compat(n) ($n=1 \cdots N_{bc}$)、下混映射函数逆函数 $fb^{-1}(k, bctob(n))$,解码的下混方案包括下混映射函数 $fb(k, i)$ 、 $fo(k, j)$ 。

[0263] 步骤4.4a.1): 计算兼容基础声道下混分量

[0264] 对于每一个兼容基础声道 $n=1 \cdots N_{bc}$, $k=compat(n)$ 为其对应的下混兼容声道,从解码的下混兼容声道信号 $\overrightarrow{srchannel}[k]$ 中去除解码的扩展基础声道信号、解码的声音对象的下混成分得到兼容基础声道下混分量 $\overrightarrow{srchannel_bed}_{cop}[k]$:

$$[0265] \quad \overrightarrow{srchannel_bed}[k] = \overrightarrow{srchannel}[k] - \overrightarrow{srchannel_obj}[k], \quad k = 1 \dots N_s$$

$$[0266] \quad \overrightarrow{srchannel_bed}_{cop}[k] = \overrightarrow{srchannel_bed}[k] - \overrightarrow{srchannel_bed}_{ext}[k], \quad k = 1 \dots N_s$$

[0267] 其中,扩展基础声道下混分量 $\overrightarrow{srchannel_bed}_{ext}[k]$ 按照解码的下混方案对解码的扩展基础声道信号下混而来:

$$[0268] \quad \overrightarrow{srchannel_bed}_{ext}[k] = \sum_{m=1}^{N_{be}} fb(k, betob(m)) \left(\overrightarrow{bedchannel}[betob(m)] \right), \quad k = 1 \dots N_s$$

[0269] 声音对象下混分量 $\overrightarrow{srchannel_obj}[k]$ 按照解码的下混方案对解码的声音对象下混渲染而来:

$$[0270] \quad \overrightarrow{srchannel_obj[k]} = \sum_{j=1}^M fo(k, j) (\overrightarrow{obj_data[j]}), k = 1 \dots Ns$$

[0271] 步骤4.4a.2): 逆映射

[0272] 对于每一个兼容基础声道 $n=1 \dots Nbc$, $k=compat(n)$ 为其对应的下混兼容声道按照解码的下混方案中映射函数 $fb(k, bctob(n))$ 的逆函数 $fb^{-1}(k, bctob(n))$ 对兼容基础声道下混分量 $\overrightarrow{srchannel_bed_cop[k]}$ 进行逆映射, 得到解码的兼容基础声道 $\overrightarrow{bedchannel_cop[n]}$:

$$[0273] \quad \overrightarrow{bedchannel_cop[n]} = fb^{-1}(k, bctob(n))(\overrightarrow{srchannel_bed_cop[k]})$$

[0274] 逆函数 $fb^{-1}(k, bctob(n))$ 的运算就是逆映射的过程。

[0275] 如果编码端没有对 $fb^{-1}(k, bctob(n))$ 进行编码, 则可通过对映射函数 $fb(k, bctob(n))$ 求解逆函数的方法而得到。

[0276] 2) 如果编码过程中使用的是基于QR分解的划分方法:

[0277] 此时, 解码的基础声道划分边信息包括 $betob(m)$ ($m=1 \dots Nbe$)、 $bctob(n)$ ($n=1 \dots Nbc$)、兼容基础声道映射系数逆矩阵 $invHRQ$, 解码的下混方案包括下混映射函数 $fb(k, i)$ 、 $fo(k, j)$ 。

[0278] 步骤4.4b.1): 计算兼容基础声道下混分量

[0279] 对于每一个下混兼容声道 k , 从解码的下混兼容声道信号 $\overrightarrow{srchannel[k]}$ 中去除解码的扩展基础声道信号、解码的声音对象的下混成分得到兼容基础声道下混分量 $\overrightarrow{srchannel_bed_cop[k]}$:

$$[0280] \quad \overrightarrow{srchannel_bed[k]} = \overrightarrow{srchannel[k]} - \overrightarrow{srchannel_obj[k]}, k = 1 \dots Ns$$

$$[0281] \quad \overrightarrow{srchannel_bed_cop[k]} = \overrightarrow{srchannel_bed[k]} - \overrightarrow{srchannel_bed_ext[k]}, k = 1 \dots Ns$$

[0282] 其中, 扩展基础声道下混分量 $\overrightarrow{srchannel_bed_ext[k]}$ 按照解码的下混方案对解码的扩展基础声道信号下混而来:

$$[0283] \quad \overrightarrow{srchannel_bed_ext[k]} = \sum_{m=1}^{Nbe} fb(k, betob(m)) (\overrightarrow{bedchannel[betob(m)]}), k = 1 \dots Ns$$

[0284] 声音对象下混分量 $\overrightarrow{srchannel_obj[k]}$ 按照解码的下混方案对解码的声音对象下混渲染而来:

$$[0285] \quad \overrightarrow{srchannel_obj[k]} = \sum_{j=1}^M fo(k, j) (\overrightarrow{obj_data[j]}), k = 1 \dots Ns$$

[0286] 所有的 $\overrightarrow{srchannel_bed_cop[k]}, k = 1 \dots Ns$ 组成矩阵

$$[0287] \quad Hs_bedcop = \begin{bmatrix} \overrightarrow{srchannel_bedcop[1]} \\ \vdots \\ \overrightarrow{srchannel_bedcop[Ns]} \end{bmatrix}$$

[0288] 步骤4.4b.2):逆映射

[0289] 如前所述,此时的下混映射函数 $f_b(k, i)$ 满足:

[0290] $f_b(k, i)(x(t)) = a(k, i)$

[0291] 根据解码的兼容基础声道映射系数逆矩阵 invHRQ ,通过如下逆映射求得 H_{bc}

[0292] $H_{bc} = \text{invHRQ} * H_{s_bedcop}$

[0293] H_{bc} 中的第 n 行即为解码的兼容基础声道信号 $\overrightarrow{\text{bedchannel}_{cop}[n]}$, $n = 1 \dots N_{bc}$ 。

[0294] 如果编码端没有对兼容基础声道映射系数逆矩阵 invHRQ 进行编码,则解码端通过与步骤2.2)基础声道划分模块中所述同样的方法计算而得到 invHRQ 。

[0295] 需要指出的是,上述表达式是对逆映射的数学关系的表达,与其等效的实现方式和过程可以有多种,比如也可以通过如下过程求得 H_{bc} :

[0296] 由兼容基础声道的下混系数 $a(k, i)$ 构造兼容基础声道下混系数矩阵 H_{Ac} , H_{Ac} 为 $N_s * N_{bc}$ 矩阵:

$$[0297] \quad H_{Ac} = \begin{bmatrix} a(1,1) & \dots & a(1, N_{bc}) \\ \vdots & \ddots & \vdots \\ a(N_s, 1) & \dots & a(N_s, N_{bc}) \end{bmatrix}$$

[0298] 对 H_{Ac} 进行QR分解的结果是 $H_{Ac} = Q * HR$,其中 Q 为一个 $N_s \times N_s$ 的正交矩阵, HR 是一个 $N_s \times N_{bc}$ 的上三角矩阵:

$$[0299] \quad HR = \begin{bmatrix} r(1,1) & \dots & r(1, N_{bc}) \\ \vdots & \ddots & \vdots \\ 0 & \dots & r(N_s, N_{bc}) \end{bmatrix}$$

[0300] $M = N_{bc} \leq N_s$,对于每一个 $n = 1 \dots M$ 都满足 $r(n, n) > 0$ 。

[0301] 因为 Q 为一个 $N_s \times N_s$ 的正交矩阵,所以其存在逆矩阵 Q^{-1} ,构造矩阵

[0302] $QH_{s_bedcop} = Q^{-1} * H_{s_bedcop}$

[0303] 因为 $M = \min(N_s, N_{bc}) = N_{bc}$,且对每一个 $n = 1 \dots M$ 都有 $r(n, n) > 0$,所以可以取 HR 的前 N_{bc} 行构造方阵

$$[0304] \quad HR_m = \begin{bmatrix} r(1,1) & \dots & r(1, N_{bc}) \\ \vdots & \ddots & \vdots \\ 0 & \dots & r(M, N_{bc}) \end{bmatrix}$$

[0305] 方阵 HR_m 是上三角矩阵且每一个 $n = 1 \dots M$ 都有 $r(n, n) > 0$,所以 HR_m 存在逆矩阵 HR_m^{-1} 。

可以取 QH_{s_bedcop} 的前 N_{bc} 行构造方阵 $QH_{s_bedcopm}$,按照如下方式进行运算得到 H_{bc} :

[0306] $H_{bc} = HR_m^{-1} * QH_{s_bedcopm}$

[0307] H_{bc} 中的第 n 行即为解码的兼容基础声道信号 $\overrightarrow{\text{bedchannel}_{cop}[n]}$, $n = 1 \dots N_{bc}$ 。

[0308] 更宽泛的范围内,如果 $f_b(k, i)(x(t)) = a(k, i) * f_b(k, 1)(x(t))$,且 $f_b(k, 1)$ 存在逆函数则 $f_b^{-1}(k, 1)$,则按照上述运算得到 H_{bc} 后,还需要对每个解码的声道信号进行 $f_b^{-1}(k, 1)$ 的运算,以得到最终的解码的兼容基础声道信号。

[0309] 步骤4.5)基础声道组合模块

[0310] 根据前述解码的基础声道划分边信息,将解码的兼容基础声道信号与解码的扩展基础声道信号进行组合,得到解码的基础声道信号。

[0311] 如果该三维节目没有基础声道信号,或者兼容基础声道信号、扩展基础声道信号

与基础声道之间是简单的一一对应关系,或扩展基础声道已经包含了所有基础声道,则可不进行此步骤4.5)的操作。

[0312] 步骤4.2)与步骤4.3)之间没有先后顺序的约束关系。

[0313] 在兼容的立体声、环绕声系统中,则只进行步骤4.1)、步骤4.2),得到解码的下混兼容声道信号。也可以根据所用还音系统或后端系统的不同,自适应地确定是按照兼容还是三维声的方式进行解码,按照兼容的方式解码时,只执行步骤4.1)、步骤4.2),得到解码的下混兼容声道信号;按照三维声的方式解码时,执行步骤4.1)到4.5),解码的基础声道信号和解码的声音对象组成解码的三维声节目。

[0314] 实施例5后向兼容的三维声编码方法1.2a

[0315] 改进的三维声编码方法:实施例3中的步骤3.2中的下混模块中,也可以选择用解码的扩展基础声道信号、解码的声音对象来进行下混。具体步骤如下(图8):

[0316] 步骤3a.1)基础声道划分模块

[0317] 与前述步骤3.1相同。

[0318] 步骤3a.2)扩展编码模块

[0319] 与前述步骤3.3相同。

[0320] 步骤3a.3)扩展解码模块

[0321] 与前述步骤3.3相同。

[0322] 步骤3a.4)下混模块

[0323] 与前述步骤3.2类似,用“解码的扩展基础声道信号”替换“扩展基础声道信号”,“解码的声音对象”替换“声音对象”来进行下混。

[0324] 步骤3a.5)兼容编码模块

[0325] 与前述步骤3.4相同。

[0326] 步骤3a.6)码流复用模块

[0327] 与前述步骤3.5相同。

[0328] 因为在下混模块使用了解码后扩展基础声道和声音对象的数据,改进的三维声编码方法有如下特点:

[0329] 1、采用现有的立体声或环绕声系统播放所述改进的三维声编码方法获得的声音数据时,下混兼容声道的质量有一定下降。这是因为下混到“下混兼容声道”的扩展基础声道和声音对象数据作了两次编码,造成了这些成分声音质量的下降。

[0330] 2、采用三维声系统播放所述改进的三维声编码方法获得的声音数据时,在编码模块的编码失真小的情况下,改进的三维声编码方法可以提升三维声的编码质量。这是因为采用所述改进的三维声编码方法时,三维声解码器在去下混时引入较小的新误差,因而提升了三维声信号中兼容基础声道的质量。

[0331] 因此,所述改进的三维声编码方法适用于编码模块的编码失真足够小、以及侧重于提升三维声的质量的应用场合。

[0332] 特别的,实施例1中步骤1.3兼容编码模块中的编码方法、实施例2中步骤2.4兼容编码模块中的编码方法、实施例3中步骤3.4兼容编码模块、实施例5中步骤3a.5兼容编码模块中的编码方法,可以采用实施例1、2、3或5中所描述的后向兼容三维声编码方法,形成可支持多层次后向兼容的码流;与之对应,步骤4.2中的需要采用实施例4中所描述的后向兼

容三维声解码方法。

[0333] 如上所述,尽管参照特定的优选实施例已经表示和表述了本发明,但其不得解释为对本发明自身的限制。在不脱离所附权利要求定义的本发明的精神和范围前提下,可对其在形式上和细节上作出各种变化。

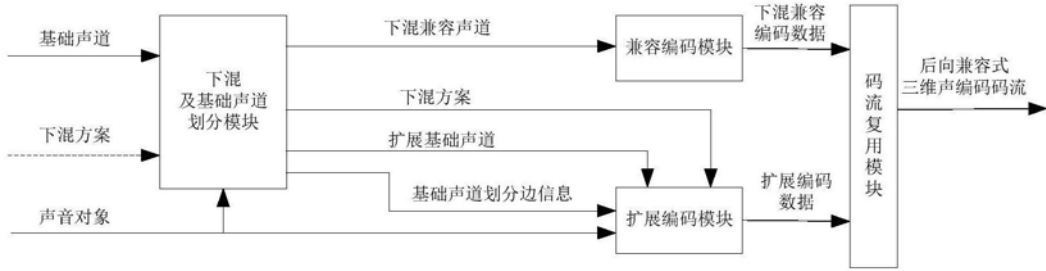


图1

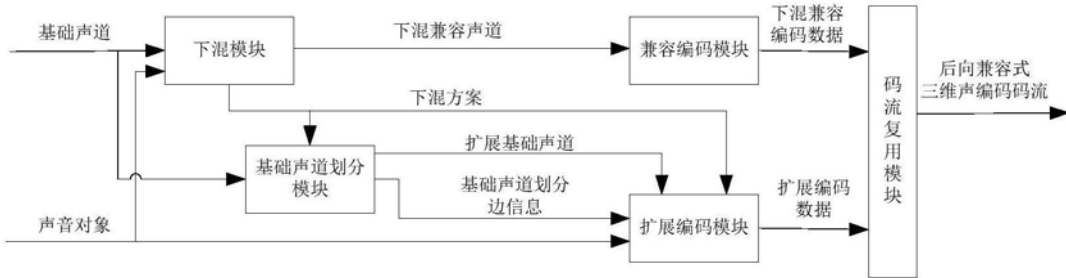


图2

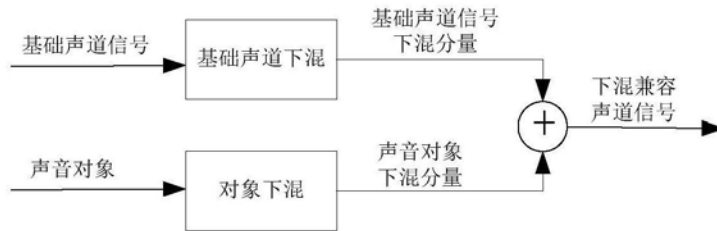


图3

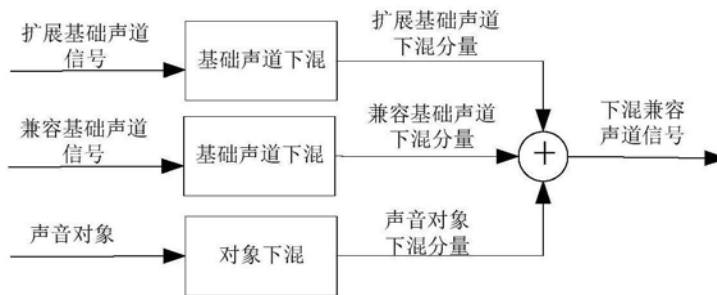


图4

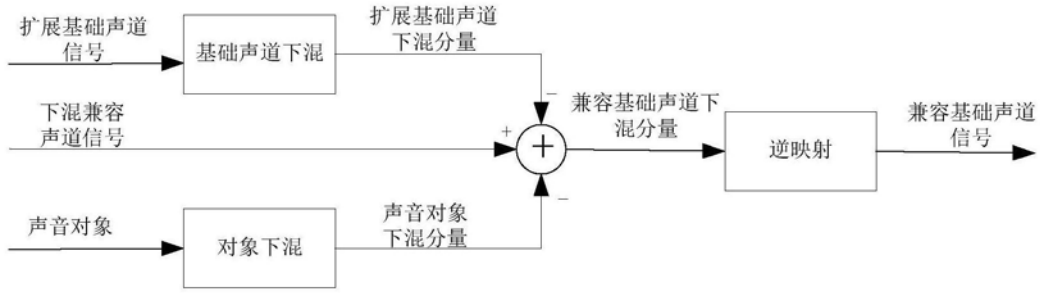


图5

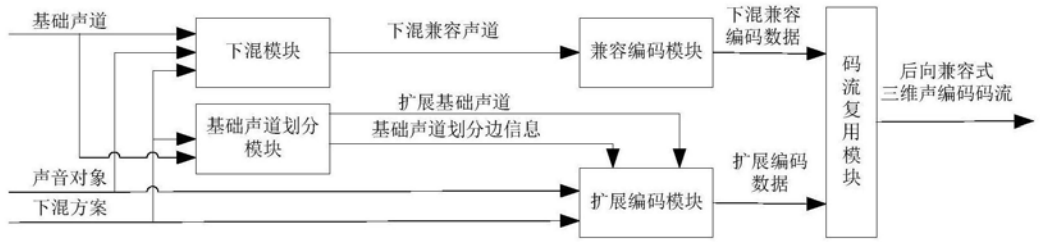


图6

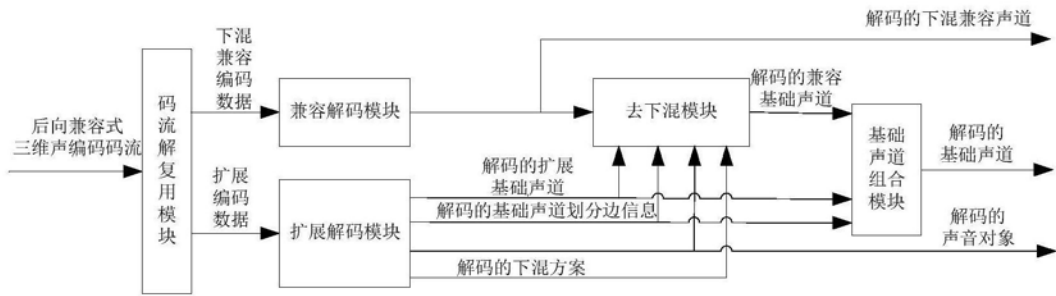


图7

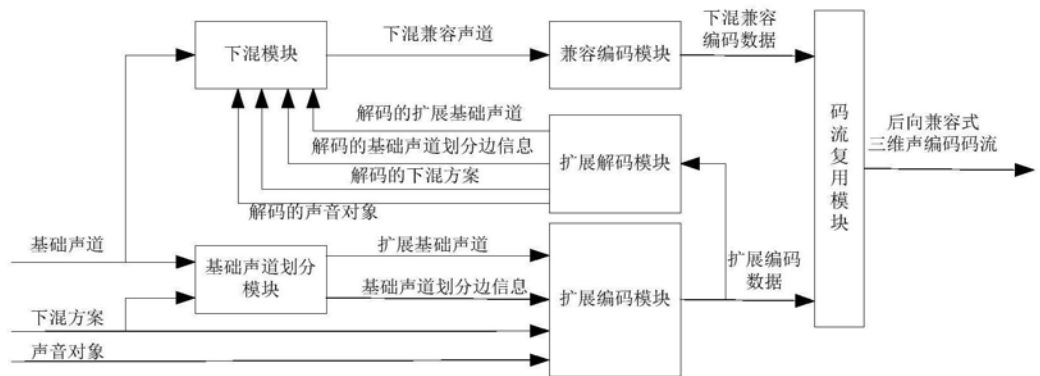


图8