



(12)发明专利

(10)授权公告号 CN 105981410 B

(45)授权公告日 2018.01.02

(21)申请号 201480074092.6

(22)申请日 2014.11.18

(65)同一申请的已公布的文献号

申请公布号 CN 105981410 A

(43)申请公布日 2016.09.28

(30)优先权数据

13306629.0 2013.11.28 EP

(85)PCT国际申请进入国家阶段日

2016.07.26

(86)PCT国际申请的申请数据

PCT/EP2014/074903 2014.11.18

(87)PCT国际申请的公布数据

W02015/078732 EN 2015.06.04

(73)专利权人 杜比国际公司

地址 荷兰阿姆斯特丹

(72)发明人 H·克罗普 S·埃伯林格

(74)专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

代理人 欧阳帆

(51)Int.Cl.

H04S 3/00(2006.01)

(56)对比文件

EP 2645748 A1,2013.10.02,

JOHANNES BOEHM ET AL.RMO-HOA Working draft text.《XP030059861》.2013,

FAZI FILIPPO M ET AL..The III-Conditioning problem in sound field reconstruction.《XP040508388》.2007,

FAZI FILIPPO ET AL.Surround system based on three-dimensional sound field reconstruction.《XP040508793》.2008,

审查员 陈嵘

权利要求书3页 说明书15页 附图6页

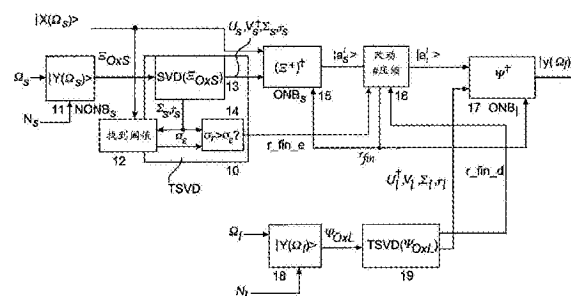
(54)发明名称

使用奇异值分解进行高阶高保真立体声编码和解码的方法和装置

(57)摘要

使用奇异值分解来对HOA信号进行编码和解码包括:基于声源方向值和高保真立体声阶次,形成(11)对应的球谐右矢量($|Y(\Omega_s)\rangle$)和编码器模式矩阵(Ξ_{0XS})。根据音频输入信号($|x(\Omega_s)\rangle$),确定奇异阈值(σ_e)。对编码器模式矩阵执行奇异值分解(13),以便得到与所述阈值比较的相关奇异值,导致最终的编码器模式矩阵秩(r_{fin_e})。基于扬声器的方向值(Ω_1)和解码器高保真立体声阶次(N_1),形成(18)对应的右矢量($|Y(\Omega_1)\rangle$)以及解码器模式矩阵(Ψ_{0XL})。对解码器模式矩阵执行奇异值分解(19),以提供最终的解码器模式矩阵秩(r_{fin_d})。根据最终的编码器模式矩阵秩和最终的解码器模式矩阵秩,确定最终的模式矩阵秩(r_{fin}),并且根据最终的模式矩阵秩以及编码器侧的奇异值分解,计算编码

器模式矩阵(Ξ_{0XS})的伴随伪逆(Ξ^+)[†]和高保真立体声右矢量($|a'_s\rangle$)。根据所述最终的模式矩阵秩(r_{fin})减少(16)高保真立体声右矢量($|a'_s\rangle$)的分量的数量,以便提供改动的高保真立体声右矢量($|a'_1\rangle$)。根据改动的高保真右矢量、解码器侧的奇异值分解的输出值和所述最终的模式矩阵秩计算(15)伴随解码器模式矩阵(Ψ^+),得到所有扬声器的输出信号的右矢量($|y(\Omega_1)\rangle$)。



1. 一种用于高阶高保真立体声 (HOA) 编码的方法, 包括:

- 接收音频输入信号 ($|x(\Omega_s)\rangle$);
- 基于声源的方向值 (Ω_s) 和所述音频输入信号 ($|x(\Omega_s)\rangle$) 的高保真立体声阶次 (N_s), 至少确定球谐右矢量 ($|Y(\Omega_s)\rangle$) 和编码器模式矩阵 (Ξ_{0xs});
- 基于对所述编码器模式矩阵 (Ξ_{0xs}) 的奇异值分解, 确定两个编码器酉矩阵 ($U_s, V_s^\dagger$)、以及包含奇异值和相关的编码器模式矩阵秩 (r_s) 的编码器对角矩阵 (Σ_s);
- 基于所述音频输入信号 ($|x(\Omega_s)\rangle$)、所述编码器对角矩阵 (Σ_s) 的所述奇异值以及所述编码器模式矩阵秩 (r_s), 确定阈值 (σ_ϵ);
- 基于所述奇异值中的至少一个 (σ_r) 与所述阈值 (σ_ϵ) 的比较, 确定最终的编码器模式矩阵秩 (r_{fine}).

2. 根据权利要求1所述的方法, 其中, 所述球谐右矢量 ($|Y(\Omega_s)\rangle$) 和所述编码器模式矩阵 (Ξ_{0xs}) 基于平移函数 (f_s), 所述平移函数 (f_s) 包括线性运算以及所述音频输入信号 ($|x(\Omega_s)\rangle$) 中的源位置到扬声器输出信号的右矢量 ($|y(\Omega_l)\rangle$) 中的扬声器位置的映射。

3. 根据权利要求1或2所述的方法, 其中, 基于在所述奇异值内从第一个奇异值 (σ_1) 开始检测的量值差距来确定所述阈值 (σ_ϵ), 并且如果后一奇异值 (σ_{i+1}) 的量值比当前奇异值 (σ_i) 的量值小, 则该当前奇异值的量值被取作所述阈值 (σ_ϵ)。

4. 根据权利要求1或2所述的方法, 其中, 基于所有源信号的采样块的信噪比 SNR 来确定所述阈值 (σ_ϵ), 并且将所述阈值 (σ_ϵ) 设置为 $\sigma_\epsilon = \frac{1}{\sqrt{SNR}}$ 。

5. 一种用于高阶高保真立体声 (HOA) 编码的装置, 包括:

- 接收器, 用于接收音频输入信号 ($|x(\Omega_s)\rangle$);
- 处理器, 被配置为基于声源的方向值 (Ω_s) 和所述音频输入信号 ($|x(\Omega_s)\rangle$) 的高保真立体声阶次 (N_s), 确定球谐右矢量 ($|Y(\Omega_s)\rangle$) 和编码器模式矩阵 (Ξ_{0xs}),
- 所述处理器还被配置为基于对所述编码器模式矩阵 (Ξ_{0xs}) 的奇异值分解, 确定两个编码器酉矩阵 ($U_s, V_s^\dagger$)、包含奇异值和相关的编码器模式矩阵秩 (r_s) 的编码器对角矩阵 (Σ_s);
- 其中所述处理器还被配置为基于所述音频输入信号 ($|x(\Omega_s)\rangle$)、所述编码器对角矩阵 (Σ_s) 的所述奇异值以及所述编码器模式矩阵秩 (r_s), 确定阈值 (σ_ϵ);
- 其中所述处理器还配置为基于所述奇异值中的至少一个 (σ_r) 与所述阈值 (σ_ϵ) 的比较, 确定最终的编码器模式矩阵秩 (r_{fine}).

6. 根据权利要求5所述的装置, 其中, 所述球谐右矢量 ($|Y(\Omega_s)\rangle$) 和所述编码器模式矩阵 (Ξ_{0xs}) 基于平移函数 (f_s), 所述平移函数 (f_s) 包括线性运算以及所述音频输入信号 ($|x(\Omega_s)\rangle$) 中的源位置到扬声器输出信号的右矢量 ($|y(\Omega_l)\rangle$) 中的扬声器位置的映射。

7. 根据权利要求5或6所述的装置, 其中, 基于在所述奇异值内从第一个奇异值 (σ_1) 开始检测的量值差距来确定所述阈值 (σ_ϵ), 并且如果后一奇异值 (σ_{i+1}) 的量值比当前奇异值 (σ_i) 的量值小, 则该当前奇异值的量值被取作所述阈值 (σ_ϵ)。

8. 根据权利要求5或6所述的装置, 其中, 基于所有源信号的采样块的信噪比 SNR 来确定所述阈值 (σ_ϵ), 并且将所述阈值 (σ_ϵ) 设置为 $\sigma_\epsilon = \frac{1}{\sqrt{SNR}}$ 。

9. 一种用于高阶高保真立体声 (HOA) 解码的方法, 包括:

- 接收关于扬声器的方向值 (Ω_1) 和解码器高保真立体声阶次 (N_1) 的信息;
- 基于扬声器的方向值 (Ω_1) 和解码器高保真立体声阶次 (N_1), 确定用于位于与所述方向值 (Ω_1) 对应的方向上的扬声器的球谐右矢量 ($|Y(\Omega_1)\rangle$) 以及解码器模式矩阵 (Ψ_{0xL});
- 基于对所述解码器模式矩阵 (Ψ_{0xL}) 的奇异值分解, 确定两个对应的解码器酉矩阵 ($U_L^\dagger, V_L$) 以及解码器对角矩阵 (Σ_L), 该解码器对角矩阵 (Σ_L) 包含奇异值和最终的解码器模式矩阵秩 (r_{fin_d});
- 基于最终的编码器模式矩阵秩 (r_{fin_e}) 和所述最终的解码器模式矩阵秩 (r_{fin_d}), 确定最终的模式矩阵秩 (r_{fin});
- 基于编码器酉矩阵 ($U_s, V_s^\dagger$)、编码器对角矩阵 (Σ_s) 和所述最终的模式矩阵秩 (r_{fin}), 确定编码器模式矩阵 (Ξ_{0xS}) 的伴随伪逆 (Ξ^+)[†], 得到高保真立体声右矢量 ($|a'_s\rangle$);
- 根据所述最终的模式矩阵秩 (r_{fin}) 基于所述高保真立体声右矢量 ($|a'_s\rangle$) 的分量数量的减少, 确定改动的高保真立体声右矢量 ($|a'_1\rangle$);
- 基于所述改动的高保真立体声右矢量 ($|a'_1\rangle$)、所述解码器酉矩阵 ($U_L^\dagger, V_L$)、所述解码器对角矩阵 (Σ_L) 和所述最终的模式矩阵秩, 确定伴随解码器模式矩阵 (Ψ^+), 得到所有扬声器的输出信号的右矢量 ($|y(\Omega_1)\rangle$).

10. 根据权利要求9所述的方法, 其中, 用于扬声器的所述球谐右矢量 ($|Y(\Omega_1)\rangle$) 和所述解码器模式矩阵 (Ψ_{0xL}) 基于对应的平移函数 (f_1), 所述平移函数 (f_1) 包括线性运算以及音频输入信号 ($|x(\Omega_s)\rangle$) 中的源位置到扬声器输出信号的所述右矢量 ($|y(\Omega_1)\rangle$) 中的扬声器位置的映射。

11. 根据权利要求9所述的方法, 其中, 在确定所述伴随解码器模式矩阵 (Ψ^+) 之后, 确定所有扬声器的时间相关的输出信号的初步改动的右矢量, 并且其中, 基于平移矩阵 (G) 来确定所有扬声器的时间相关的输出信号的所述初步改动的右矢量, 得到所有扬声器的输出信号的所述右矢量 ($|y(\Omega_1)\rangle$).

12. 一种用于高阶高保真立体声 (HOA) 解码的装置, 包括:

- 接收器, 用于接收关于扬声器的方向值 (Ω_1) 和解码器高保真立体声阶次 (N_1) 的信息;
- 处理器, 被配置为基于扬声器的方向值 (Ω_1) 和解码器高保真立体声阶次 (N_1), 确定用于位于与所述方向值 (Ω_1) 对应的方向上的扬声器的球谐右矢量 ($|Y(\Omega_1)\rangle$) 以及解码器模式矩阵 (Ψ_{0xL}), 并且基于对所述解码器模式矩阵 (Ψ_{0xL}) 的奇异值分解, 确定两个对应的解码器酉矩阵 ($U_L^\dagger, V_L$) 以及解码器对角矩阵 (Σ_L), 该解码器对角矩阵 (Σ_L) 包含奇异值和最终的解码器模式矩阵秩 (r_{fin_d});

其中所述处理器还被配置为基于最终的编码器模式矩阵秩 (r_{fin_e}) 和所述最终的解码器模式矩阵秩 (r_{fin_d}), 确定最终的模式矩阵秩 (r_{fin});

其中所述处理器还被配置为基于编码器酉矩阵 ($U_s, V_s^\dagger$)、编码器对角矩阵 (Σ_s) 和所述最终的模式矩阵秩 (r_{fin}), 确定编码器模式矩阵 (Ξ_{0xS}) 的伴随伪逆 (Ξ^+)[†], 得到高保真立体

声右矢量 ($|a'_s\rangle$)，

其中所述处理器还被配置为根据所述最终的模式矩阵秩 (r_{fin}) 基于所述高保真立体声右矢量 ($|a'_s\rangle$) 的分量数量的减少，确定改动的高保真立体声右矢量 ($|a'_1\rangle$)；

其中所述处理器还被配置为基于所述改动的高保真立体声右矢量 ($|a'_1\rangle$)、所述解码器酉矩阵 ($U_l^\dagger, V_l$)、所述解码器对角矩阵 (Σ_l) 和所述最终的模式矩阵秩，确定伴随解码器模式矩阵 (Ψ^+)，得到所有扬声器的输出信号的右矢量 ($|y(\Omega_l)\rangle$)。

13. 根据权利要求12所述的装置，其中，用于扬声器的所述球谐右矢量 ($|Y(\Omega_l)\rangle$) 和所述解码器模式矩阵 (Ψ_{oxL}) 基于对应的平移函数 (f_l)，所述平移函数 (f_l) 包括线性运算以及音频输入信号 ($|x(\Omega_s)\rangle$) 中的源位置到扬声器输出信号的所述右矢量 ($|y(\Omega_l)\rangle$) 中的扬声器位置的映射。

14. 根据权利要求12所述的装置，其中，在确定所述伴随解码器模式矩阵 (Ψ^+) 之后，确定所有扬声器的时间相关的输出信号的初步改动的右矢量，并且其中，基于平移矩阵 (G) 来确定所有扬声器的时间相关的输出信号的所述初步改动的右矢量，得到所有扬声器的输出信号的所述右矢量 ($|y(\Omega_l)\rangle$)。

15. 一种非暂时性计算机可读介质，其包括当在计算机上被执行时执行根据权利要求1或9所述的方法的指令。

使用奇异值分解进行高阶高保真立体声编码和解码的方法和装置

技术领域

[0001] 本发明涉及用于使用奇异值分解进行高阶高保真立体声 (Ambisonics) 编码和解码的方法和装置。

背景技术

[0002] 高阶高保真立体声 (Higher Order Ambisonics, HOA) 表示三维声音。其他技术是波场合成 (WFS) 或基于声道的方法 (比如, 22.2)。然而, 与基于声道的方法相反, HOA 表示提供独立于特定扬声器设置的优点。但是, 这种灵活性是以 HOA 表示在特定扬声器设置上的回放所需的解码处理为代价的。与所需扬声器的数量通常非常大的 WFS 方法相比, HOA 也可以被渲染到仅由几个扬声器组成的设置。HOA 的其他优点是, 同一表示也可以无任何修改地用于对耳机的双耳渲染。

[0003] HOA 基于复谐波平面波振幅的空间密度的通过截断球谐 (SH) 扩展的表示。每个扩展系数是角频率的函数, 其可以等效地由时域函数表示。因此, 不失一般性地, 整个 HOA 声场表示实际上可被假定为由 O 个时域函数组成, 其中 O 表示扩展系数的数量。这些时域函数在下面将被等效地称为 HOA 系数序列或 HOA 声道。HOA 表示可以被表达为包含 HOA 系数的 HOA 数据帧的时间序列。HOA 表示的空间分辨率随着扩展的最大阶次 N 增长而提高。对于 3D 情况, 扩展系数的数量 O 随着阶次 N 呈二次方地增长, 特别地为 $O = (N+1)^2$ 。

[0004] 复矢量空间

[0005] 高保真立体声必须处理复函数。因此引入了基于复矢量空间的符号。该符号随抽象的复矢量操作, 所述复矢量不表示从三维 “xyz” 坐标系知道的真实的几何矢量。相反, 每个复矢量描述物理系统的可能的状态, 并且在 d 维空间中由具有 d 个分量 x_i 的列矢量形成, 并且根据 Dirac, 这些面向列的矢量被称为右矢量 (ket vector), 被表示为 $|x\rangle$ 。在 d 维空间中, 任意的 $|x\rangle$ 由其分量 x_i 和 d 个正交基矢量 $|e_i\rangle$ 形成:

$$[0006] \quad |x\rangle = x_1 |e_1\rangle + x_2 |e_2\rangle + \dots + x_d |e_d\rangle = \sum_{i=1}^d x_i |e_i\rangle \quad (1)$$

[0007] 这里, d 维空间不是通常的 “xyz” 3D 空间。

[0008] 右矢量的共轭复数被称为左矢量 (bra vector) $|x\rangle^* = \langle x|$ 。左矢量表示基于行的描述, 并且形成原始右矢量空间的对偶空间——左矢量空间。

[0009] 该 Dirac 符号将被用在高阶高保真立体声相关音频系统的以下描述中。

[0010] 内积可以根据同一维度的左矢量和右矢量来构建, 得到复数标量值。如果随机矢量 $|x\rangle$ 由其在正交矢量基中的分量描述, 则对于特定基的特定分量, 即, $|x\rangle$ 到 $|e_i\rangle$ 上的投影, 由内积给出:

$$[0011] \quad x_i = \langle x | e_i \rangle = \langle x | e_i \rangle \quad (2)$$

[0012] 在左矢量和右矢量之间, 仅考虑一个条、而非两个条。

[0013] 对于同一基中的不同矢量 $|x\rangle$ 和 $|y\rangle$, 内积通过将左矢量 $\langle x|$ 乘以右矢量 $|y\rangle$ 而得

到,使得:

[0014]

$$\langle x | y \rangle = \sum_{i=1}^d \langle x_i e_i | \sum_{j=1}^d y_j | e_j \rangle = \sum_{i,j=1}^d x_i^* y_j \langle e_i | e_j \rangle = \sum_{i,j=1}^d x_i^* y_j = \sum_{i,j=1}^d y_j^* x_i \quad (3)$$

[0015] 如果维度 $m \times 1$ 的右矢量和维度 $1 \times n$ 的左矢量被乘以外积,则具有 m 行 n 列的矩阵 A 被导出:

$$[0016] \quad A = |x\rangle\langle y| \quad (4)$$

[0017] 高保真立体声矩阵

[0018] 基于高保真立体声的描述考虑了用于将整个声场映射到时变矩阵中所需的相关性(dependency)。在高阶高保真立体声(HOA)编码或解码矩阵中,行(列)的数量与从声源或声宿(sound sink)起的特定方向相关。在编码器侧,变化数量的 S 个声源被考虑,其中, $s=1, \dots, S$ 。每个声源 s 可以具有离原点的个体距离 r_s 、个体方向 $\Omega_s = (\Theta_s, \Phi_s)$,其中, Θ_s 描述从 z 轴开始的倾斜角, Φ_s 描述从 x 轴开始的方位角。对应的时间相关信号 $X_s = (t)$ 具有个体时间行为。

[0019] 为简单起见,仅考虑方向部分(径向相关性将由贝塞尔函数描述)。那么,特定方向 Ω_s 用列矢量 $|Y_n^m(\Omega_s)\rangle$ 描述,其中, n 表示高保真立体声度, m 是高保真立体声阶次 N 的索引。对应的值分别按 $m=1, \dots, N$ 和 $n=-m, \dots, 0, \dots, m$ 取值。

[0020] 一般来说,特定HOA描述根据 N 来对2D或3D情况下的每个右矢量 $|Y_n^m(\Omega_s)\rangle$ 限制分量的数量0:

$$[0021] \quad O = \begin{cases} 2N+1, & 2D \\ (N+1)^2, & 3D \end{cases} \quad (5)$$

[0022] 对于多于一个的声源,如果阶次 n 的 s 个个体矢量 $|Y_n^m(\Omega_s)\rangle$ 被组合,则所有方向都被包括。这导致包含 $O \times S$ 个模式分量的模式矩阵 Ξ ,即, Ξ 的每个列表示特定方向:

$$[0023] \quad \Xi = \begin{bmatrix} Y_0^0(\Omega_1) & \dots & Y_0^0(\Omega_S) \\ Y_1^{-1}(\Omega_1) & \dots & Y_1^{-1}(\Omega_S) \\ \vdots & \ddots & \vdots \\ Y_N^N(\Omega_1) & \dots & Y_N^N(\Omega_S) \end{bmatrix} \quad (6)$$

[0024] 所有信号值都被组合在信号矢量 $|x(kT)\rangle$ 中,但是被以共同的采样速率 $\frac{1}{T}$ 采样,信号矢量 $|x(kT)\rangle$ 考虑每个源信号 $x_s(kT)$ 的时间相关性:

$$[0025] \quad |x(kT)\rangle = \begin{bmatrix} x_1(kT) \\ x_2(kT) \\ \vdots \\ x_S(kT) \end{bmatrix} \quad (7)$$

[0026] 在下文中,为简单起见,在比如 $|x(kT)\rangle$ 的时变信号中,采样编号 k 不再被描述,即,它将被忽视。那么,如方程(8)中所示, $|x\rangle$ 被乘以模式矩阵 Ξ 。这确保所有信号分量都与对应

的同一方向 Ω_s 的列线性地组合, 根据方程 (5) 得到具有 0 个高保真立体声模式分量或系数的右矢量 $|a_s\rangle$:

$$[0027] \quad |a_s\rangle = \Xi |x\rangle \quad (8)$$

[0028] 解码器具有再现由专用数量的 1 个扬声器信号 $|y\rangle$ 所表示的声场 $|a_1\rangle$ 的任务。因此, 扬声器模式矩阵 Ψ 由基于球谐的单位矢量 $|Y_n^m(\Omega_l)\rangle$ 的 L 个单独的列组成 (类似于方程 (6)), 即, 每个扬声器方向一个右矢量:

$$[0029] \quad \Omega_1: |a_1\rangle = \Psi |y\rangle \quad (9)$$

[0030] 对于其中模式的数量等于扬声器的数量的二次矩阵, $|y\rangle$ 可以由模式矩阵 Ψ 的逆矩阵确定。在其中行和列的数量可以不同的任意矩阵的一般情况下, 扬声器信号 $|y\rangle$ 可以由伪逆确定, 参见 M.A. Poletti, “A Spherical Harmonic Approach to 3D Surround Sound Systems”, Acusticum 论坛, 布达佩斯, 2005。那么, 利用 Ψ 的伪逆 Ψ^+ :

$$[0031] \quad |y\rangle = \Psi^+ |a_1\rangle \quad (10)$$

[0032] 假定在编码器侧和解码器侧描述的声场几乎是相同的, 即, $|a_s\rangle \approx |a_1\rangle$ 。然而, 扬声器位置可以不同于源位置, 即, 对于有限的高保真立体声阶次, 由 $|x\rangle$ 描述的实值源信号和由 $|y\rangle$ 描述的扬声器信号是不同的。因此, 将映射 $|x\rangle$ 在 $|y\rangle$ 上的平移矩阵 G 可以被使用。那么, 根据方程 (8) 和 (10), 编码器和解码器的串运算是:

$$[0033] \quad |y\rangle = G \Psi^+ \Xi |x\rangle \quad (11)$$

[0034] 线性泛函

[0035] 为了保持以下方程更简单, 平移矩阵直到“发明内容”章节都将被忽视。如果所需的基矢量的数量变为无限, 则基矢量可以从离散基变为连续基。因此, 函数 f 可以被解释为具有无限数量的模式分量的矢量。这从数学意义上来讲被称为“泛函”, 因为它以确定性的方式执行从右矢量到特定的输出右矢量的映射。它可以用函数 f 和右矢量 $|x\rangle$ 之间的内积表示, 得到一般的复数 c :

$$[0036] \quad \langle f | (|x\rangle) = \sum_{i=1}^N f_i \cdot x_i = c \quad (12)$$

[0037] 如果泛函保持右矢量的线性组合, 则 f 被称为“线性泛函”。

[0038] 只要对 Hermitean 运算符存在约束, 就应考虑以下特性。Hermitean 运算符总是具有:

[0039] • 实的特征值;

[0040] • 对于不同特征值的完备的正交特征函数集。

[0041] 因此, 每个函数可以从这些特征函数建立, 参见 H. Vogel、C. Gerthsen、H.O. Kneser, “Physik”, Springer Verlag, 1982。任意函数可以被表示为球谐 $Y_n^m(\theta, \phi)$ 与复数常数 C_n^m 的线性组合:

$$[0042] \quad f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-N}^N C_n^m \cdot Y_n^m(\theta, \phi) \quad (13)$$

$$[0043] \quad \langle f(\theta, \phi) | Y_{n'}^{m'}(\theta, \phi) \rangle = \int_0^{2\pi} \int_0^\pi f(\theta, \phi)^* Y_{n'}^{m'}(\theta, \phi) \sin \theta d\theta d\phi \quad (14)$$

[0044] 索引n、m被以确定性的方式使用。它们被一维索引j取代,索引n'、m'被相同大小的索引i替代。由于每个子空间正交于具有不同i、j的子空间的事实,它们可被描述为无限维的空间中的线性独立的正交单位矢量:

$$[0045] \quad \langle f(\theta, \phi) | Y_i(\theta, \phi) \rangle = \int_0^{2\pi} \int_0^\pi \left(\sum_{j=0}^{\infty} C_j Y_j(\theta, \phi) \right)^* Y_i(\theta, \phi) \sin \theta d\theta d\phi \quad (15)$$

[0046] C_j 的常数值可以被设置在积分的前面:

$$[0047] \quad \langle f(\theta, \phi) | Y_i(\theta, \phi) \rangle = \sum_{j=0}^{\infty} C_j^* \int_0^{2\pi} \int_0^\pi Y_j^*(\theta, \phi) Y_i(\theta, \phi) \sin \theta d\theta d\phi \quad (16)$$

[0048] 只要本征函数 Y_j 和 Y_i 相互正交,则从一个子空间(索引j)到另一个子空间(索引i)的映射仅需要针对相同索引 $i=j$ 对谐波进行积分:

$$[0049] \quad \langle f(\theta, \phi) | Y_i(\theta, \phi) \rangle = \sum_{j=0}^{\infty} C_j^* \langle Y_j(\theta, \phi) | Y_i(\theta, \phi) \rangle \quad (17)$$

[0050] 一个本质的方面是,如果存在从连续描述到左矢量/右矢量符号的变化,则积分解(integral solution)可以被球谐的左矢量描述和右矢量描述之间的内积的和替代。一般来说,与连续基的内积可以被用于将基于右矢量的波描述 $|x\rangle$ 的离散表示映射到连续表示。例如, $x(\text{ra})$ 是位置基(即,半径)中的右矢量表示:

$$[0051] \quad \text{ra}: x(\text{ra}) = \langle \text{ra} | x \rangle \quad (18)$$

[0052] 针对不同种类的模式矩阵 Ψ 和 Ξ ,奇异值分解被用于处理任意种类的矩阵。

[0053] 奇异值分解

[0054] 奇异值分解(SVD,参见G.H.Golub、Ch.F.van Loan,“Matrix Computations”,约翰霍普金斯大学出版社,第3版,1996年10月11日)使得能够将具有m行和n列的任意矩阵A分解为三个矩阵U、 Σ 和 $V^\dagger$,参见方程(19)。在原始形式中,矩阵U和 $V^\dagger$ 分别是维度 $m \times m$ 和 $n \times n$ 的西矩阵(unitary matrix)。这样的矩阵是正交的,并且分别是从表示复单位矢量 $|u_i\rangle$ 和 $|v_i\rangle^\dagger = \langle v_i|$ 的正交的列构建的。来自复数空间的西矩阵与实空间中的正交矩阵是等效的,即,它们的列表示正交矢量基:

$$[0055] \quad A = U \Sigma V^\dagger \quad (19)$$

[0056] 矩阵U和V包含用于所有四个子空间的正交基。

[0057] • U的头r个列:A的列空间

[0058] • U的后 $m-r$ 个列: $A^\dagger$ 的零空间

[0059] • V的头r个列:A的行空间

[0060] • V的后 $n-r$ 个列:A的零空间

[0061] 矩阵 Σ 包含可以被用于表征A的行为的所有奇异值。一般来说, Σ 是具有多达r个对角元素 σ_i 的 $m \times n$ 矩形对角矩阵,其中,秩r给出A的线性独立的列和行的数量($r \leq (\min(m, n))$)。它包含降序的奇异值,即,在方程(20)和(21)中, σ_1 具有最高值, σ_r 具有最低值。

[0062] 在紧凑形式中,只有r个奇异值(即,U的r个列和 $V^\dagger$ 的r个行)是重构矩阵A必需的。

矩阵 U 、 Σ 和 $V^\dagger$ 的维度不同于原始形式。然而， Σ 矩阵总是得到二次形式。那么，对于 $m > n = r$ ，

$$[0063] \quad \begin{bmatrix} **** \\ **** \\ **** \\ A \\ **** \\ **** \\ **** \\ **** \end{bmatrix}_{m \times n} = \begin{bmatrix} **** \\ **** \\ **** \\ U \\ **** \\ **** \\ **** \\ **** \end{bmatrix}_{m \times n} \cdot \begin{bmatrix} \sigma_1 & 0 & \cdot & \cdot \\ 0 & \sigma_2 & 0 & \cdot \\ 0 & 0 & \cdot & 0 \\ \cdot & \cdot & 0 & \sigma_r \end{bmatrix}_{n \times n} \begin{bmatrix} **** \\ **** \\ V^\dagger \\ **** \end{bmatrix}_{n \times n} \quad (20)$$

[0064] 对于 $n > m = r$ ，

$$[0065] \quad \begin{bmatrix} **** \\ **** \\ **** \\ A \\ **** \end{bmatrix}_{m \times n} = \begin{bmatrix} **** \\ **** \\ U \\ **** \end{bmatrix}_{m \times m} \cdot \begin{bmatrix} \sigma_1 & 0 & \cdot & \cdot \\ 0 & \sigma_2 & 0 & \cdot \\ 0 & 0 & \cdot & 0 \\ \cdot & \cdot & 0 & \sigma_r \end{bmatrix}_{m \times m} \begin{bmatrix} **** \\ **** \\ V^\dagger \\ **** \end{bmatrix}_{m \times n} \quad (21)$$

[0066] 因此，可以通过低秩逼近来非常高效地实现SVD，参见以上提及的Golub/van Loan教科书。这种逼近精确地描述了原始矩阵，但是包含多达 r 个秩-1矩阵。就Dirac符号而言，矩阵 A 可以由 r 个秩-1外积表示：

$$[0067] \quad A = \sum_{i=1}^r \sigma_i |u_i\rangle\langle v_i| \quad (22)$$

[0068] 当查看方程(11)中的编码器解码器链时，不仅存在用于编码器的模式矩阵（比如，矩阵 Ξ ），而且模式矩阵（比如，矩阵 Ψ ）的逆矩阵或另一个复杂的解码器矩阵也要被考虑。对于一般的矩阵 A ，可以通过执行方形矩阵 Σ 的逆以及 U 和 $V^\dagger$ 的共轭复数转置来直接从SVD检查 A 的伪逆 A^+ ，其结果是：

$$[0069] \quad A^+ = V \Sigma^{-1} U^\dagger \quad (23)$$

[0070] 对于方程(22)的基于矢量的描述，伪逆 A^+ 通过执行 $|u_i\rangle$ 和 $\langle v_i|$ 的共轭转置而得到，而奇异值 σ_i 必须被求逆。作为结果得到的伪逆看上去如下：

$$[0071] \quad A^+ = \sum_{i=1}^r \left(\frac{1}{\sigma_i}\right) |v_i\rangle\langle u_i| \quad (24)$$

[0072] 如果不同矩阵的基于SVD的分解与基于矢量的描述组合（参见方程(8)和(10)），则对于编码处理得到：

$$[0073] \quad |a_s\rangle = \sum_{s_i=1}^{r_s} \sigma_{s_i} |u_{s_i}\rangle\langle v_{s_i}| \cdot |x\rangle = \sum_{s_i=1}^{r_s} \sigma_{s_i} |u_{s_i}\rangle\langle v_{s_i}| x \rangle \quad (25)$$

[0074] 对于解码器，当考虑伪逆矩阵 Ψ^+ （方程(24)）时：

$$[0075] \quad |y\rangle = \left(\sum_{l=1}^r \left(\frac{1}{\sigma_{s_l}} \right) |v_{s_l}\rangle \langle u_{s_l}| \right) |a_l\rangle \quad (26)$$

[0076] 如果假定来自编码器的高保真立体声声场描述 $|a_s\rangle$ 与用于解码器的 $|a_l\rangle$ 几乎相同, 并且维度 $r_s = r_l = r$, 则对于输入信号 $|x\rangle$ 和输出信号 $|y\rangle$, 组合的方程看上去如下:

$$[0077] \quad |y\rangle = \left(\sum_{l=1}^r \left(\frac{1}{\sigma_{s_l}} \right) |v_{s_l}\rangle \langle u_{s_l}| \right) \sum_{s_l=1}^r \sigma_{s_l} |u_{s_l}\rangle \langle v_{s_l}| x\rangle \quad (27)$$

发明内容

[0078] 然而, 编码器解码器链的这个组合的描述具有如下所述的一些特定问题。

[0079] 对高保真立体声矩阵的影响

[0080] 高阶高保真立体声 (H0A) 模式矩阵 Ξ 和 Ψ 受到声源或扬声器的位置 (参见方程 (6)) 和它们的高保真立体声阶次的直接影响。如果几何结构是规则的, 即, 源或扬声器位置之间的相互角距离几乎相等, 则方程 (27) 可以被求解。

[0081] 但是在实际的应用中, 通常不是如此。因此, 执行 Ξ 和 Ψ 的 SVD 并且调查它们在对应的矩阵 Σ 中的奇异值是有道理的, 因为它反映了 Ξ 和 Ψ 的数值行为。 Σ 是具有实奇异值的正定矩阵。但是尽管如此, 即使存在多达 r 个奇异值, 这些值之间的数值关系对于声场的再现也是非常重要的, 因为在解码器侧必须构建矩阵的逆或伪逆。用于测量该行为的合适的量是 A 的条件数 (condition number)。条件数 $\kappa(A)$ 被定义为最小奇异值与最大奇异值之比:

$$[0082] \quad \kappa(A) = \frac{\sigma_r}{\sigma_1} \quad (28)$$

[0083] 逆问题

[0084] 病态矩阵是有问题的, 因为它们具有大的 $\kappa(A)$ 。在求逆或求伪逆的情况下, 病态矩阵导致很小的奇异值 σ_1 变得非常占主导的问题。在 P.Ch.Hansen, "Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion", 工业和应用数学学会 (SIAM), 1998 中, 两种基本类型的问题通过描述奇异值如何衰减而被区分 (第 1.1 章第 2-3 页)。

[0085] • 秩亏问题, 在该秩亏问题中, 矩阵在大奇异值和小奇异值的聚类之间具有差距 (非逐渐衰减);

[0086] • 离散不适定问题, 在该离散不适定问题中, 平均地, 矩阵的所有奇异值都逐渐衰减到零, 即, 在奇异值谱中没有差距。

[0087] 关于编码器侧的麦克风的几何结构以及解码器侧的扬声器几何结构, 主要发生第一个秩亏问题。然而, 在记录期间修改一些麦克风的位置比控制客户端的所有可能的扬声器位置更简单。尤其是在解码器侧, 模式矩阵的逆或伪逆将被执行, 这对于较高的模式分量导致数值问题和过分强调的值 (参见以上提及的 Hansen 的书)。

[0088] 信号有关的相关性

[0089] 减少上述求逆问题可以例如通过降低模式矩阵的秩, 即, 通过避免最小奇异值来实现。但是然后, 阈值将被用于最小的可能的值 σ_r (参见方程 (20) 和 (21))。在以上提及的

Hansen的书中描述了用于这种最低奇异值的最佳值。Hansen提出了 $\sigma_{opt} = \frac{1}{\sqrt{SNR}}$, 其取决于输入信号的特性(这里用 $|x\rangle$ 描述)。从方程(27)可以看出, 该信号对再现具有影响, 但是信号相关性在解码器中不能被控制。

[0090] 非正交基的问题

[0091] 根据方程(25)和(26), 在不同基的每个系统中描述了在HOA编码和HOA解码器之间传输的状态矢量 $|a_s\rangle$ 。然而, 如果正交基被使用, 则状态不改变。那么, 模式分量可以从一个基投影到另一个基。所以, 原则上, 每个扬声器设置或声音描述应构建在正交基系统上, 因为这允许这些基之间的矢量表示的改变, 例如, 在高保真立体声中, 从3D空间投影到2D子空间中。

[0092] 然而, 经常存在具有病态矩阵的设置, 在病态矩阵中, 基矢量是近似线性相关的。所以, 原则上, 非正交基将被处理。这使得当HOA声场描述应被采用到不同的扬声器设置上时或者当期望在编码器侧或解码器侧处理不同的HOA阶次和维度时必要的从一个子空间到另一个子空间的改变复杂化。

[0093] 投影到稀疏扬声器集上的典型问题是, 声音能量在扬声器附近高, 并且在这些扬声器之间的距离很时低。所以, 不同扬声器之间的位置需要相应地对能量进行平衡的平移函数。

[0094] 上述问题可以通过本发明的处理来绕开, 并且通过权利要求1中公开的方法来解决。利用该方法的设备在权利要求2中被公开。

[0095] 根据本发明, 考虑到最低的模式矩阵秩以及截断的奇异值分解, 用于编码过程的互逆基(reciprocal basis)与用于解码过程的原始基被组合使用。因为双正交系统被表示, 所以确保编码器矩阵和解码器矩阵的乘积保持至少用于最低的模式矩阵秩的单位矩阵。

[0096] 这通过将基于右矢量的描述变为基于对偶空间(具有互逆基矢量的左矢量空间)的表示来实现, 在对偶空间中, 每一个矢量是右矢量的伴随(adjoint)。它通过使用模式矩阵的伪逆的伴随来实现。“伴随”意指复数共轭转置。

[0097] 因此, 伪逆的伴随已经被用于编码器侧以及伴随解码器矩阵。对于所述处理, 正交互逆基矢量被使用, 以便关于基变化而不变。此外, 这种类型的处理允许考虑与输入信号相关的影响, 对于正则化过程中的 σ_i 导致降噪最佳阈值。

[0098] 原则上, 本发明的方法适合于使用奇异值分解的高阶高保真立体声编码和解码, 所述方法包括以下步骤:

[0099] -接收音频输入信号;

[0100] -基于声源的方向值和所述音频输入信号的高保真立体声阶次, 形成对应的球谐右矢量和对应的编码器模式矩阵;

[0101] -对所述编码器模式矩阵执行奇异值分解, 其中, 两个对应的编码器酉矩阵、对应的包含奇异值的编码器对角矩阵以及相关的编码器模式矩阵秩被输出;

[0102] -从所述音频输入信号、所述奇异值和所述编码器模式矩阵秩确定阈值;

[0103] -将所述奇异值中的至少一个与所述阈值进行比较, 并且确定对应的最终的编码器模式矩阵秩;

- [0104] -基于扬声器的方向值和解码器高保真立体声阶次,形成对应的用于位于与所述方向值对应的方向上的特定扬声器的球谐右矢量以及对应的解码器模式矩阵;
- [0105] -对所述解码器模式矩阵执行奇异值分解,其中,两个对应的解码器酉矩阵以及对应的包含奇异值的解码器对角矩阵被输出,并且所述解码器模式矩阵的对应的最终的秩被确定;
- [0106] -从所述最终的编码器模式矩阵秩和所述最终的解码器模式矩阵秩确定最终的模式矩阵秩;
- [0107] -从所述编码器酉矩阵、所述编码器对角矩阵和所述最终的模式矩阵秩计算所述编码器模式矩阵的伴随伪逆,得到高保真立体声右矢量,
- [0108] 并且根据所述最终的模式矩阵秩减少所述高保真立体声右矢量的分量的数量,以便提供改动的高保真立体声右矢量;
- [0109] -从所述改动的高保真右矢量、所述解码器酉矩阵、所述解码器对角矩阵和所述最终的模式矩阵秩计算伴随解码器模式矩阵,得到所有扬声器的输出信号的右矢量。
- [0110] 原则上,本发明设备适于使用奇异值分解的高阶高保真立体声编码和解码,所述设备包括适于执行以下操作的部件:
- [0111] -接收音频输入信号;
- [0112] -基于声源的方向值和所述音频输入信号的高保真立体声阶次,形成对应的球谐右矢量和对应的编码器模式矩阵;
- [0113] -对所述编码器模式矩阵执行奇异值分解,其中,两个对应的编码器酉矩阵、对应的包含奇异值的编码器对角矩阵以及相关的编码器模式矩阵秩被输出;
- [0114] -从所述音频输入信号、所述奇异值和所述编码器模式矩阵秩确定阈值;
- [0115] -将所述奇异值中的至少一个与所述阈值进行比较,并且确定对应的最终的编码器模式矩阵秩;
- [0116] -基于扬声器的方向值和解码器高保真立体声阶次,形成对应的用于位于与所述方向值对应的方向上的特定扬声器的球谐右矢量以及对应的解码器模式矩阵;
- [0117] -对所述解码器模式矩阵执行奇异值分解,其中,两个对应的解码器酉矩阵以及对应的包含奇异值的解码器对角矩阵被输出,并且所述解码器模式矩阵的对应的最终的秩被确定;
- [0118] -从所述最终的编码器模式矩阵秩和所述最终的解码器模式矩阵秩确定最终的模式矩阵秩;
- [0119] -从所述编码器酉矩阵、所述编码器对角矩阵和所述最终的模式矩阵秩计算所述编码器模式矩阵的伴随伪逆,得到高保真立体声右矢量,
- [0120] 并且根据所述最终的模式矩阵秩减少所述高保真立体声右矢量的分量的数量,以便提供改动的高保真立体声右矢量;
- [0121] -从所述改动的高保真右矢量、所述解码器酉矩阵、所述解码器对角矩阵和所述最终的模式矩阵秩计算伴随解码器模式矩阵,得到所有扬声器的输出信号的右矢量。
- [0122] 本发明的有利的附加的实施例在各个从属权利要求中被公开。

附图说明

- [0123] 参照附图描述本发明的示例性实施例,在附图中示出了:
- [0124] 图1 基于SVD的HOA编码器和解码器的框图;
- [0125] 图2 包括线性泛函平移的HOA编码器和解码器的框图;
- [0126] 图3 包括矩阵平移的HOA编码器和解码器的框图;
- [0127] 图4 用于确定阈值 σ_e 的流程图;
- [0128] 图5 在降低的模式矩阵秩 r_{fin_e} 的情况下对奇异值的重新计算以及 $|a'_s\rangle$ 的计算;
- [0129] 图6在有平移或没有平移时,在降低的模式矩阵秩 r_{fin_e} 和 r_{fin_d} 的情况下对奇异值的重新计算以及 $|y(\Omega_1)\rangle$ 的计算。

具体实施方式

[0130] 图1中描绘了关于本发明的基于SVD的HOA处理的、具有编码器部分和解码器部分的框图。两个部分都在使用SVD,以便产生互逆基矢量。相对于已知的模式匹配解有所改变,例如,与方程(27)相关的改变。

[0131] HOA编码器

[0132] 为了用互逆基矢量工作,基于右矢量的描述被变到左矢量空间,在左矢量空间中,每个矢量是右矢量的Hermitean共轭或伴随。它通过使用模式矩阵的伪逆来实现。然后,根据方程(8),基于(对偶)左矢量的高保真立体声矢量也可以重新用(对偶)模式矩阵表示:

$$[0133] \quad \Xi_d: \langle a_s | = \langle x | \Xi_d = \langle x | \Xi^+ \quad (29)$$

[0134] 所得到的编码器侧的高保真立体声矢量 $\langle a_s |$ 现在为左矢量语义。然而,统一的描述是期望的,即,返回到右矢量语义。代替 Ξ 的伪逆,使用 Ξ_d^+ 或 $\Xi^{+†}$ 的Hermitean共轭:

$$[0135] \quad |a_s\rangle = \Xi_d^+ |x\rangle = \Xi^{+†} |x\rangle \quad (30)$$

[0136] 根据方程(24),

$$[0137] \quad \Xi^{+†} = \left(\sum_{i=1}^{r_s} \left(\frac{1}{\sigma_{s_i}} \right) |v_{s_i}\rangle \langle u_{s_i}| \right)^{\dagger} = \sum_{i=1}^{r_s} \left(\frac{1}{\sigma_{s_i}} \right) |u_{s_i}\rangle \langle v_{s_i}| \quad (31)$$

[0138] 其中,所有奇异值都是实数,并且 σ_{s_i} 的复数共轭可以被忽视。

[0139] 这导致高保真立体声分量的以下描述:

$$[0140] \quad |a_s\rangle = \sum_{i=1}^{r_s} \left(\frac{1}{\sigma_{s_i}} \right) |u_{s_i}\rangle \langle v_{s_i}| x \rangle \quad (32)$$

[0141] 用于源端的基于矢量的描述揭示了 $|a_s\rangle$ 取决于逆 σ_{s_i} 。如果对于编码器侧这样做,则它将变到解码器侧的对应的对偶基矢量。

[0142] HOA解码器

[0143] 在解码器最初基于伪逆的情况下,为了导出扬声器信号 $|y\rangle$ 得到:

$$[0144] \quad |a_l\rangle = \Psi^{+†} |y\rangle \quad (33)$$

[0145] 即,扬声器信号为:

$$[0146] \quad |y\rangle = (\Psi^{*†})^* \cdot |a_l\rangle = \Psi^\dagger \cdot |a_l\rangle \quad (34)$$

[0147] 考虑方程(22), 解码器方程导致:

$$[0148] \quad |y\rangle = (\sum_{i=1}^r \sigma_{i_l} |u_{i_l}\rangle \langle v_{i_l}|)^\dagger |a_l\rangle \quad (35)$$

[0149] 因此, 不是构建伪逆, 而是在方程(35)中只保留伴随运算(用“ $\dagger$ ”表示)。这意味着在解码器中需要更少的算术运算, 因为只需要切换虚部的符号, 并且转置仅仅是修正的存储器访问:

$$[0150] \quad |y\rangle = (\sum_{i=1}^r \sigma_{i_l} \cdot |v_{i_l}\rangle \langle u_{i_l}|) |a_l\rangle \quad (36)$$

[0151] 如果假定编码器和解码器的高保真立体声表示是几乎相同的, 即, $|a_s\rangle = |a_l\rangle$, 则通过方程(32), 整个编码器解码器链得到以下相关性:

$$[0152] \quad |y\rangle = \sum_{i=1}^r \left(\frac{\sigma_{i_l}}{\sigma_{s_l}} \right) |v_{i_l}\rangle \langle u_{i_l} | u_{s_l} \rangle \langle v_{s_l} | x \rangle \quad (37)$$

$$[0153] \quad |y\rangle = \sum_{i=1}^r \left(\frac{\sigma_{i_l}}{\sigma_{s_l}} \right) \langle u_{i_l} | u_{s_l} \rangle \cdot |v_{i_l}\rangle \langle v_{s_l} | x \rangle \quad (38)$$

[0154] 在真实的情景下, 来自方程(11)的平移矩阵G和有限的高保真立体声阶次要被考虑。后者导致被用于描述声场的基矢量的有限数量的线性组合。此外, 基矢量的线性独立性受到附加的误差源(比如数值舍入误差或测量误差)的影响。从实践角度来看, 这可以通过数值秩来绕开(参见以上提及的Hansen的书的第3.1章), 这确保所有的基矢量在某些容限内都是线性独立的。

[0155] 为了对噪声更鲁棒, 考虑输入信号的SNR, 其中输入信号的SNR影响编码器右矢量以及所计算的输入的高保真立体声表示。所以, 如果必要的话, 即, 对于将被求逆的病态模式矩阵, 根据编码器中的输入信号的SNR来对 σ_i 值进行正则化。

[0156] 编码器中的正则化

[0157] 正则化可以通过不同的方式来执行, 例如, 通过使用经由截断的SVD的阈值。SVD提供降序的 σ_i , 其中, 具有最低水平或最高索引的 σ_i (被表示为 σ_r)包含非常频繁地切换的并且导致噪声效应和SNR的分量(参看方程(20)和(21)以及以上提及的Hansen教科书)。因此, 截断SVD(TSVD)将所有的 σ_i 值与阈值进行比较, 并且忽视超出阈值 σ_ϵ 的噪声分量。阈值 σ_ϵ 可以是固定的或者可以根据输入信号的SNR被最佳地修正。

[0158] 矩阵的迹(trace)是指所有对角矩阵元素的和。

[0159] TSVD方框(图1至3中的10、20、30)具有以下任务:

[0160] • 计算模式矩阵秩r;

[0161] • 移除低于阈值的噪声分量, 并且设置最终的模式矩阵秩 r_{fin} 。

[0162] 所述处理对复矩阵 Ξ 和 Ψ 进行处理。然而, 为了对实数值 σ_i 进行正则化, 这些矩阵不能被直接使用。适当的值来自 Ξ 与其伴随 $\Xi^\dagger$ 之间的乘积。所得到的矩阵是具有实数对角本征值的二次矩阵, 所述实数对角本征值等同于适当的奇异值的二次值。如果可以用矩阵 Σ^2 的迹(如下面的表达式39所示)描述的所有本征值之和保持固定, 则系统的物理性质守

恒。这也适用于矩阵 Ψ 。

$$[0163] \quad \text{trace}(\Sigma^2) = \sum_{i=1}^r \sigma_i^2 \quad (39)$$

[0164] 因此,编码器侧的方框ONB_s(图1-3中的15、25、35)或解码器侧的方框ONB_l(图1-3中的19、29、39)修改奇异值,以使得正则化之前和之后的 $\text{trace}(\Sigma^2)$ 守恒(参见图5和图6):

[0165] • 修正其余的 σ_i (对于 $i=1 \dots r_{fin}$),以使得原始矩阵和目标截断矩阵 Σ_t 的迹保持固定($\text{trace}(\Sigma^2) = \text{trace}(\Sigma_t^2)$)。

[0166] • 计算满足以下表达式的常数值 $\Delta\sigma$:

$$[0167] \quad \sum_{i=1}^r \sigma_i^2 = \sum_{i=1}^{r_{fin}} (\sigma_i + \Delta\sigma)^2 \quad (40)$$

[0168] 如果正常数量的奇异值和减少数量的奇异值之间的差值被调用($\Delta E = \text{trace}(\Sigma) - \text{trace}(\Sigma)_{r_{fin}}$),则所得到的值如下:

$$[0169] \quad \Delta\sigma = \frac{1}{r_{fin}} \left(-\sum_{i=1}^{r_{fin}} \sigma_i + \sqrt{\left[\sum_{i=1}^{r_{fin}} \sigma_i \right]^2 + r_{fin} \Delta E} \right) = \frac{1}{r_{fin_d}} \left(-\text{trace}(\Sigma)_{r_{fin}} + \sqrt{\text{trace}(\Sigma)_{r_{fin}}^2 + r_{fin_d} \Delta E} \right) \quad (41)$$

[0170] • 对截断的矩阵 Σ_t 重新计算所有的新的奇异值 $\sigma_{i,t}$:

$$[0171] \quad \sigma_{i,t} = \sigma_i + \Delta\sigma \quad (42)$$

[0172] 另外,如果用于适当的 $|a\rangle$ 的基(参见方程(30)或(33))被改变为对应的SVD相关的 $\{U^\dagger\}$ 基,则可以实现编码器和解码器的简化,导致:

$$[0173] \quad |a'\rangle = \sum_{i=1}^{r_{fin}} \langle u_i | \left[\sum_{i=1}^{r_{fin}} \sigma_{i,t} |u_i\rangle \langle v_i| \right] |a\rangle = \sum_{i=1}^{r_{fin}} \sigma_{i,t} \langle v_i | a \rangle \quad (43)$$

[0174] (备注:如果 σ_i 和 $|a\rangle$ 在没有附加的编码器或解码器索引的情况下被使用,则它们是指编码器侧和/或解码器侧)。该基是正交的,以使得它保持 $|a\rangle$ 的范数。即,代替 $|a\rangle$,正则化可以使用 $|a'\rangle$, $|a'\rangle$ 需要矩阵 $|a'\rangle$ 和 V ,但是不需要更大的矩阵 U 。

[0175] 在 $\{U^\dagger\}$ 基中使用缩小的右矢量 $|a'\rangle$,这具有秩真正被降低的优点。

[0176] 因此,在本发明中,SVD在两侧都被使用,不仅用于执行单个矩阵 Ξ 和 Ψ 的正交基和奇异值,而且还用于得到它们的秩 r_{fin} 。

[0177] 分量改动

[0178] 通过考虑 Ξ 的源秩或者通过相对于阈值或最终的源秩忽视对应的 σ_s 中的一些,可以减少分量的数量,并且可以提供更鲁棒的编码矩阵。因此,根据解码器侧的分量的对应数量对所传输的高保真立体声分量的数量的改动被执行。通常,它取决于高保真立体声阶次 O 。这里,从用于编码器矩阵 Ξ 的SVD方框得到的最终的模式矩阵秩 r_{fin_e} 和从用于解码器矩阵 Ψ 的SVD方框得到的最终的模式矩阵秩 r_{fin_d} 要被考虑。在改动#压缩步骤/级16中,如下改动分量的数量:

[0179] • $r_{fin_e} = r_{fin_d}$; 不改变—不压缩;

[0180] • $r_{fin_e} < r_{fin_d}$: 压缩, 忽视解码器矩阵 $\Psi^\dagger$ 中的 $r_{fin_e} - r_{fin_d}$ 个列 $\Rightarrow$ 编码器和解码器操作减少;

[0181] • $r_{fin_e} > r_{fin_d}$: 在传输之前取消高保真立体声状态矢量的 $r_{fin_e} - r_{fin_d}$ 的分量, 即, 压缩。忽视编码器矩阵 Ξ 中的 $r_{fin_e} - r_{fin_d}$ 个行 $\Rightarrow$ 编码器和解码器操作减少。

[0182] 结果是, 将在编码器侧和解码器侧使用的最终的模式矩阵秩 r_{fin} 是 r_{fin_d} 和 r_{fin_e} 中的较小的一个。

[0183] 因此, 如果在编码器和解码器之间存在用于交换另一端的秩的双向信号, 则可以使用秩差异来改进可能的压缩并且减少编码器和解码器中的操作的数量。

[0184] 考虑平移函数

[0185] 由于关于为稀疏和不规则的扬声器设置得到的能量分布的问题, 前面提及了平移函数 f_s, f_l 或平移矩阵 G 的使用, 参见方程 (11)。这些问题必须对通常可以在高保真立体声中使用的有限阶次进行处理 (参见章节“对高保真立体声矩阵的影响”到“非正交基的问题”)。

[0186] 关于对平移矩阵 G 的要求, 在编码之后, 假定一些声学源的声场处于由高保真立体声状态矢量 $|a_s\rangle$ 表示的良好状态。然而, 在解码器侧, 状态已经被准备得如何是完全未知的。即, 完全不知道系统的当前状态。因此, 为了保持方程 (9) 和 (8) 之间的内积, 采取互逆基。

[0187] 使用已经在编码器侧的伪逆提供了以下优点:

[0188] • 互逆基的使用满足了编码器基和解码器基之间的双正交性 ($\langle x^i | x_j \rangle = \delta_j^i$);

[0189] • 编码/解码链中的操作的数量减少;

[0190] • 在关于 SNR 行为的数值方面得到改善;

[0191] • 经修改的模式矩阵中的列是正交的, 而不仅仅是线性独立的;

[0192] • 它简化了基的改变;

[0193] • 使用秩-1 逼近导致更少的存储器工作和减少的操作数量, 尤其在最终的秩较低的时候。一般来说, 对于 $M \times N$ 矩阵, 只需要 $M+N$ 个操作, 而不是 $M*N$ 个操作;

[0194] • 它简化了解码器侧的改动, 因为解码器中的伪逆可以被避免;

[0195] • 具有数值不稳定的 σ 的逆问题可以被绕开。

[0196] 在图 1 中, 在编码器或解码器侧, 声源的 $s=1, \dots, S$ 不同的方向值 Ω_s 和高保真立体声阶次 N_s 被输入到步骤或级 11, 步骤或级 11 从其形成球谐的对应右矢量 $|Y(\Omega_s)\rangle$ 和具有维度 $0 \times S$ 的编码器模式矩阵 $\Xi_{0 \times S}$ 。矩阵 $\Xi_{0 \times S}$ 是对应于输入信号矢量 $|x(\Omega_s)\rangle$ 而产生的, 输入信号矢量 $|x(\Omega_s)\rangle$ 包括关于不同方向 Ω_s 的 S 个源信号。因此, 矩阵 $\Xi_{0 \times S}$ 是球谐右矢量 $|Y(\Omega_s)\rangle$ 的集合。因为不仅信号 $x(\Omega_s)$ 还有位置随着时间而变化, 所以计算矩阵 $\Xi_{0 \times S}$ 可以被动态地执行。该矩阵具有用于源的非正交基 $NONB_s$ 。根据输入信号 $|x(\Omega_s)\rangle$ 和秩值 r_s , 在步骤或级 12 中确定特定的奇异阈值 σ_ϵ 。编码器模式矩阵 $\Xi_{0 \times S}$ 和阈值 σ_ϵ 被馈送到截断奇异值分解 TSVD 处理 10 (参看以上的章节“奇异值分解”), 该处理在步骤或级 13 中执行对模式矩阵 $\Xi_{0 \times S}$ 的奇异值分解, 以便得到其奇异值, 由此, 一方面, 酉矩阵 U 和 $V^\dagger$ 以及包含 r_s 个奇异值 $\sigma_1 \dots \sigma_{r_s}$ 的对角矩阵 Σ 被输出, 另一方面, 相关的编码器模式矩阵秩 r_s 被确定 (备注: σ_i 是来自 $SVD(\Xi) = U \Sigma V^\dagger$ 的矩阵 Σ 的第 i 奇异值)。

[0197] 在步骤/级12中,根据章节“编码器中的正则化”来确定阈值 σ_e 。阈值 σ_e 可以将所使用的 σ_{s_i} 值的数量限制为截断的或最终的编码器模式矩阵秩 r_{fin_e} 。阈值 σ_e 可以被设置为预定义的值,或者可以被改为输入信号的信噪比SNR: $\sigma_{e,opt} = \frac{1}{\sqrt{SNR}}$,由此全部S个源信号 $|x(\Omega_s)\rangle$ 的SNR在预定义数量的采样值上被测量。

[0198] 在比较器步骤或级14中,将来自矩阵 Σ 的奇异值 σ_r 与阈值 σ_e 进行比较,并且根据该比较,计算截断的或最终的编码器模式矩阵秩 r_{fin_e} ,该秩 r_{fin_e} 根据章节“编码器中的正则化”来修改其余的 σ_{s_i} 值。最终的编码器模式矩阵秩 r_{fin_e} 被馈送到步骤或级16。

[0199] 关于解码器侧,在步骤或级18中,对应于方框17中的相关信号 $|y(\Omega_1)\rangle$ 的扬声器位置,根据扬声器的 $l=1, \dots, L$ 方向值 Ω_1 和解码器高保真立体声阶次 N_1 ,确定关于方向 Ω_1 上的特定扬声器的对应的球谐右矢量 $|Y(\Omega_1)\rangle$ 以及对应的具有维度 $Q \times L$ 的解码器模式矩阵 $\Psi_{0 \times L}$ 。类似于编码器矩阵 $\Xi_{0 \times S}$,解码器矩阵 $\Psi_{0 \times L}$ 是用于所有方向 Ω_1 的球谐右矢量 $|Y(\Omega_1)\rangle$ 的集合。 $\Psi_{0 \times L}$ 的计算被动态地执行。

[0200] 在步骤或级19中,对解码器模式矩阵 $\Psi_{0 \times L}$ 执行奇异值分解处理,并且将所得到的酉矩阵 U 和 $V^\dagger$ 以及对角矩阵 Σ 馈送到方框17。此外,最终的解码器模式矩阵秩 r_{fin_d} 被计算并且被馈送到步骤/级16。

[0201] 在步骤或级16中,如上所述的那样,根据最终的编码器模式矩阵秩 r_{fin_e} 和最终的解码器模式矩阵秩 r_{fin_d} 确定最终的矩阵秩 r_{fin} 。将最终的模式矩阵秩 r_{fin} 馈送到步骤/级15和步骤/级17。

[0202] 编码器侧矩阵 U_s 、 $V_s^\dagger$ 、 Σ_s 、秩值 r_s 、最终的模式矩阵秩值 r_{fin} 以及所有源信号的时间相关的输入信号右矢量 $|x(\Omega_s)\rangle$ 被馈送到步骤或级15,步骤或级15使用方程(32)从这些与 $\Xi_{0 \times S}$ 相关的输入值计算编码器模式矩阵的伴随伪逆 $(\Xi^+)^{\dagger}$ 。该矩阵具有维度 $r_{fin_e} \times S$ 和用于源的正交基 ONB_s 。当处理复矩阵及其伴随时,考虑: $\Xi_{0 \times S}^{\dagger} \Xi_{0 \times S} = \text{trace}(\Sigma^2) = \sum_{l=1}^r \sigma_{s_l}^2$ 。步骤/级15输出对应的时间相关的高保真立体声右矢量或状态矢量 $|a's\rangle$,参见上面的章节“HOA编码器”。

[0203] 在步骤或级16中,如上面的章节“分量改动”中所描述的,使用最终的模式矩阵秩 r_{fin} 来减少 $|a's\rangle$ 的分量的数量,以便可能地减少所传输的信息量,在改动之后得到时间相关的高保真立体声右矢量或状态矢量 $|a'1\rangle$ 。

[0204] 根据保真立体声右矢量或状态矢量 $|a'1\rangle$,根据从模式矩阵 $\Psi_{0 \times L}$ 导出的秩值 r_1 以及解码器侧矩阵 $U_l^\dagger$ 、 V_l 、 Σ_l ,以及根据来自步骤/级16的最终的模式矩阵秩值 r_{fin} ,计算具有维度 $L \times r_{fin_d}$ 和用于扬声器的正交基 ONB_l 的伴随解码器模式矩阵 $(\Psi)^{\dagger}$,得到所有扬声器的时间相关的输出信号的右矢量 $|y(\Omega_1)\rangle$,参见上面的章节“HOA解码器”。解码是利用平常的模式矩阵的共轭转置执行的,所述共轭转置依赖于特定的扬声器位置。对于附加的渲染,应使用特定的平移矩阵。

[0205] 解码器由步骤/级18、19和17表示。编码器由其他步骤/级表示。

[0206] 图1的步骤/级11至19原则上分别对应于图2中的步骤/级21至29以及图3中的步

骤/级31至39。

[0207] 另外,在图2中,在步骤或级211中计算的用于编码器侧的平移函数 f_s 以及在步骤或级218中计算的用于解码器侧的平移函数 f_{1281} 被用于线性泛函平移。平移函数 f_s 是用于步骤/级21的附加输入信号,平移函数 f_1 是用于步骤/级28的附加输入信号。在上面的章节“考虑平移函数”中描述了使用这种平移函数的原因。

[0208] 与图1相比,在图3中,平移矩阵 G 控制对在步骤/级37的输出处的所有扬声器的时间相关的输出信号的初步右矢量的平移处理371。这导致所有扬声器的时间相关的输出信号的改动的右矢量 $|y(\Omega_1)\rangle$ 。

[0209] 图4更详细地示出了用于基于编码器模式矩阵 Ξ_{0XS} 的奇异值分解SVD处理来确定阈值 σ_e 的处理。SVD处理传递矩阵 Σ (在其递减对角中包含从 σ_1 到 σ_{r_s} 的所有奇异值 σ_i ,参见方程(20)和(21))以及矩阵 Σ 的秩 r_s 。在使用固定阈值的情况(方框41)下,在从 $i=1$ 开始并且可以一直到 $i=r_s$ 的、由变量 i 控制的循环(方框42和43)内,检查(方框45)在这些 σ_i 值之间是否存在量值差距。这种差距被假定为当奇异值 σ_{i+1} 的量值明显小于其在前的奇异值 σ_i 的量值(例如,小于 $1/10$)时发生。当检测到这种差距时,循环停止,并且阈值 σ_e 被设置为当前奇异值 σ_i 。在 $i=r_s$ 的情况(方框44)下,达到最低的奇异值 $\sigma_i=\sigma_r$,回路被退出,并且 σ_e 被设定为 σ_r (方框46)。

[0210] 在不使用固定阈值的情况(方框41)下,调查所有 S 个源信号的 T 个采样的块 $X=[|x(\Omega_s, t=0)\rangle, \dots, |x(\Omega_s, t=T)\rangle]$ (=矩阵 $S \times T$)。计算 X 的信噪比SNR(方框48),并且将阈值 σ_e 设置为 $\sigma_e = \frac{1}{\sqrt{SNR}}$ (方框49)。

[0211] 图5示出了在步骤/级15、25、35内,在降低的模式矩阵秩 r_{fin} 的情况下的奇异值的重新计算以及 $|a'_s\rangle$ 的计算。来自图1/2/3的方框10/20/30的编码器对角矩阵 Σ_s 被馈送到步骤或级51、步骤或级52、步骤或级54,其中步骤或级51使用值 r_s 来计算总能量 $trace(\Sigma^2) = \sum_{i=1}^{r_s} \sigma_{s_i}^2$,步骤或级52使用值 r_{fin_e} 来计算减小的总能量 $trace(\Sigma_{r_{fin_e}}^2) = \sum_{i=1}^{r_{fin_e}} \sigma_{s_i}^2$ 。总能量值和减小的总能量值之间的差值 ΔE 、值 $trace(\Sigma_{r_{fin_e}})$ 和值 r_{fin_e} 被馈送到步骤或级53,步骤或级53计算 $\Delta\sigma = \frac{1}{r_{fin_e}} \left(-trace(\Sigma_{r_{fin_e}}) + \sqrt{[trace(\Sigma_{r_{fin_e}})]^2 + r_{fin_e} \Delta E} \right)$ 。

[0212] 为了确保由 $trace(\Sigma^2) = \sum_{i=1}^r \sigma_{s_i}^2$ 描述的能量被保持为使得结果在物理上是有意义的,值 $\Delta\sigma$ 是需要的。如果在编码器或解码器侧,能量由于矩阵简化而减小,则这样的能量损失被用值 $\Delta\sigma$ 补偿,值 $\Delta\sigma$ 被以相等的方式分布给所有剩余的矩阵元素,即, $\sum_{i=1}^{r_{fin}} (\sigma_i + \Delta\sigma)^2 = \sum_{i=1}^r (\sigma_i)^2$ 。

[0213] 步骤或级54从 Σ_s 、 $\Delta\sigma$ 和 r_{fin_e} 计算 $\Sigma_t^+ = \sum_{i=1}^{r_{fin_e}} \frac{1}{(\sigma_{s_i} + \Delta\sigma)} I$ 。

[0214] 输入信号矢量 $|x(\Omega_s)\rangle$ 被乘以矩阵 $V_s^\dagger$ 。结果乘以 Σ_t^+ 。后一乘法的结果是右矢量 $|a'_s\rangle$ 。

[0215] 图6示出了在步骤/级17、27、37内,在有或没有平移时在降低的模式矩阵秩 r_{fin} 的情况下的奇异值的重新计算以及扬声器信号 $|y(\Omega_1)\rangle$ 的计算。来自图1/2/3中的方框19/29/39的解码器对角矩阵 Σ_1 被馈送到步骤或级61、步骤或级62以及步骤或级64,其中步骤

或级61使用值 r_1 来计算总能量 $trace(\Sigma^2) = \sum_{i=1}^{r_1} \sigma_{s_i}^2$, 步骤或级62使用值 r_{find} 来计算减小的总能量 $trace(\Sigma_{r_{find}}^2) = \sum_{i=1}^{r_{find}} \sigma_{s_i}^2$ 。总能量值和减小的总能量值之间的差值 ΔE 、值 $trace(\Sigma_{r_{find}})$ 和值 r_{find} 被馈送到步骤或级63, 步骤或级63计算:

$$[0216] \quad \Delta\sigma = \frac{1}{r_{find}} (-trace(\Sigma_{r_{find}}) + \sqrt{(trace(\Sigma_{r_{find}}))^2 + r_{find}\Delta E})$$

$$[0217] \quad \text{步骤或级64从 } \Sigma_1, \Delta\sigma \text{ 和 } r_{find} \text{ 计算 } \Sigma_t = \sum_{i=1}^{r_{find}} \frac{1}{(\sigma_{s_i} + \Delta\sigma)} I$$

[0218] 右矢量 $|a'_s\rangle$ 被乘以矩阵 Σ_t 。结果被乘以矩阵 V 。后一乘法的结果是所有扬声器的时间相关的输出信号的右矢量 $|y(\Omega_1)\rangle$ 。

[0219] 本发明处理可以由单个处理器或电子电路执行, 或者由并行操作和/或对本发明的处理的不同部分操作的几个处理器或电子电路执行。

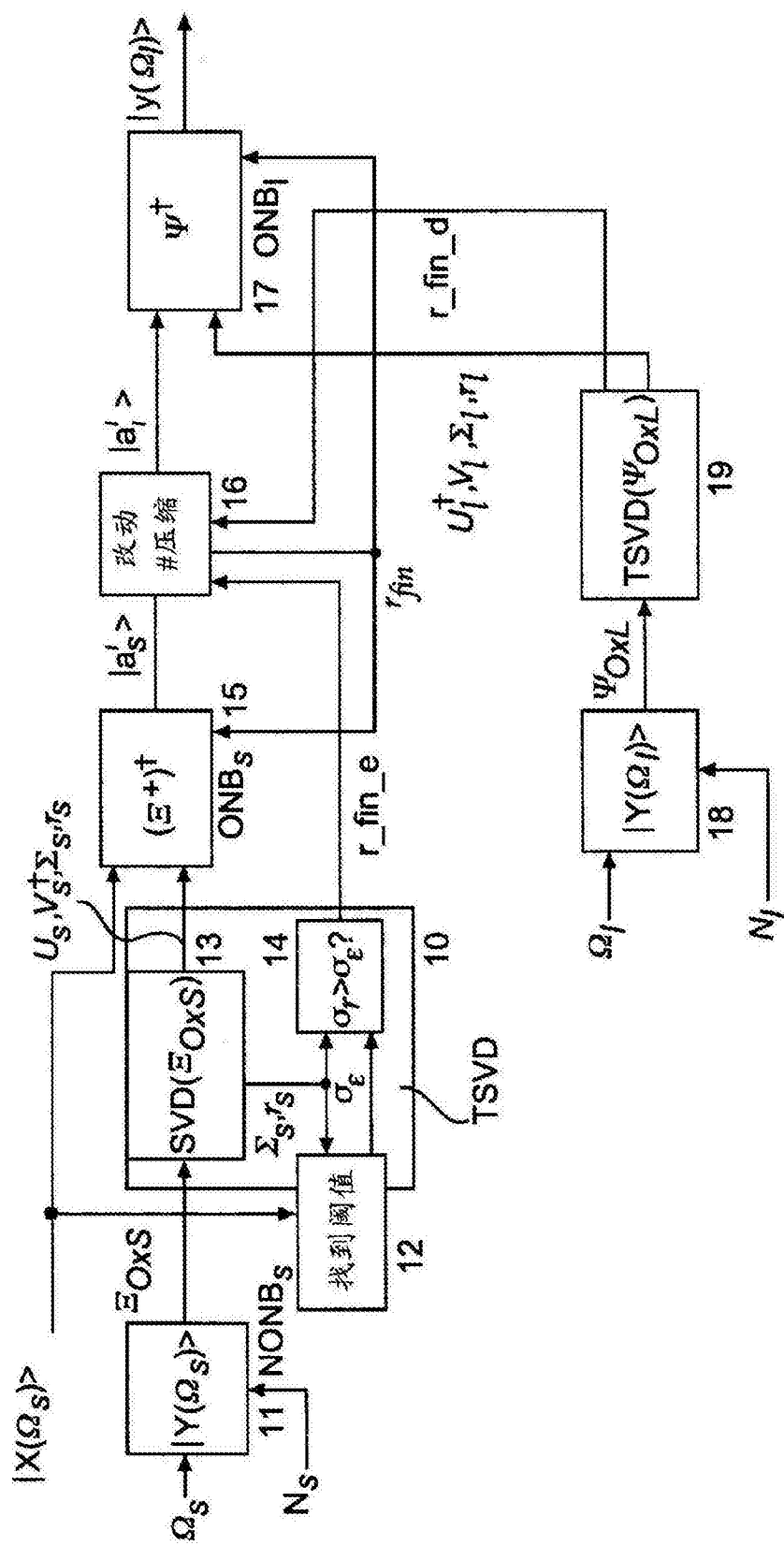


图1

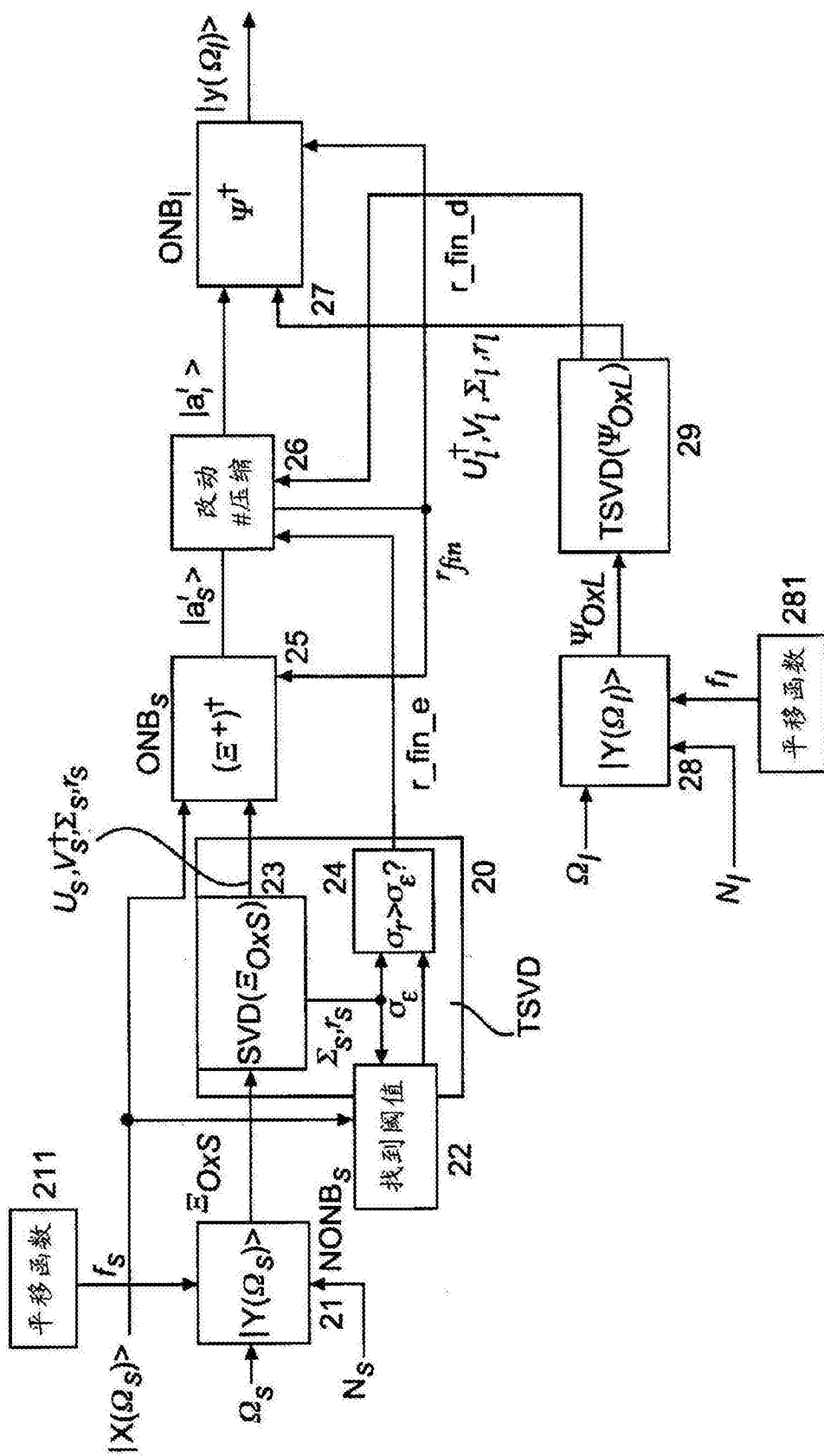


图2

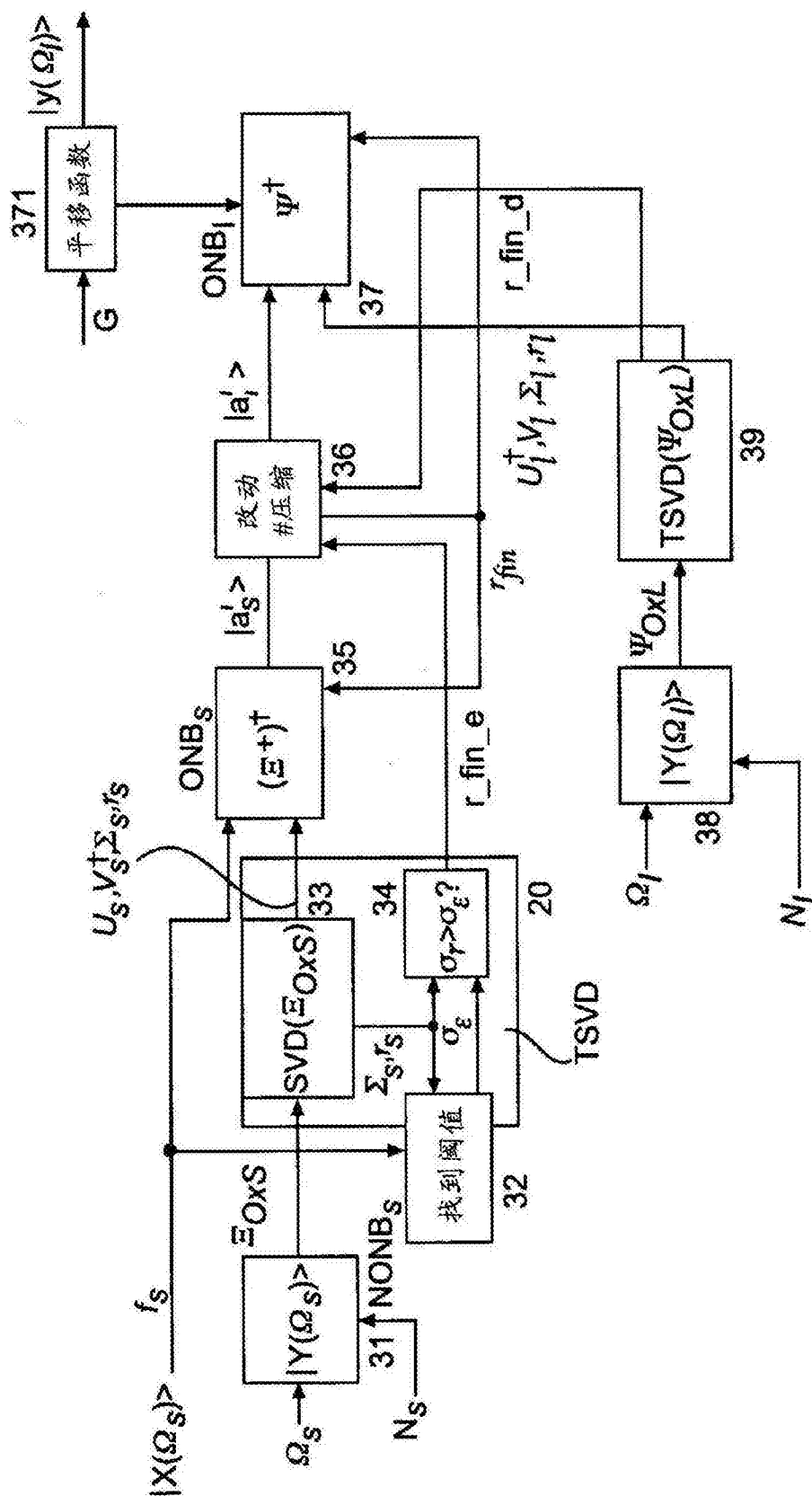


图3

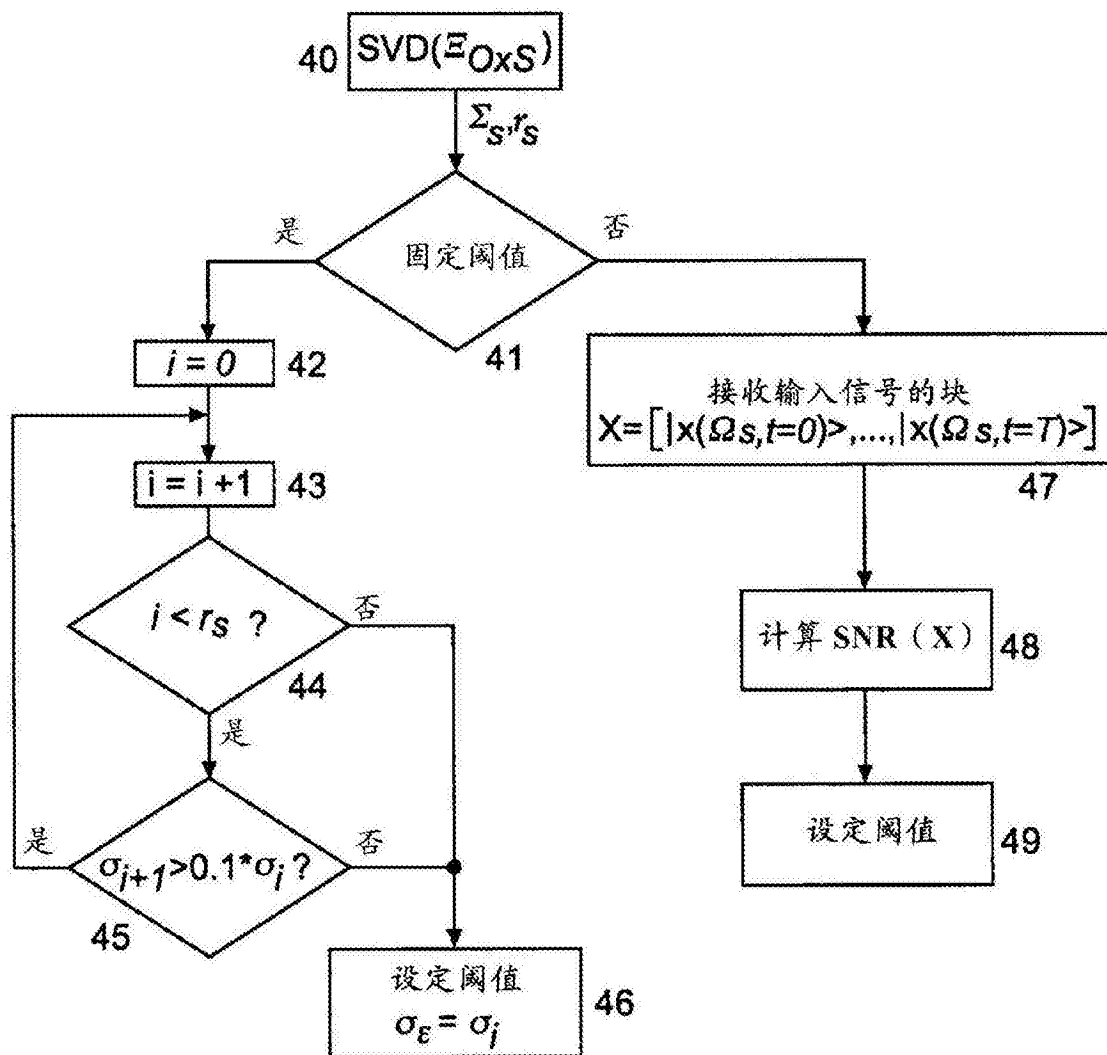


图4

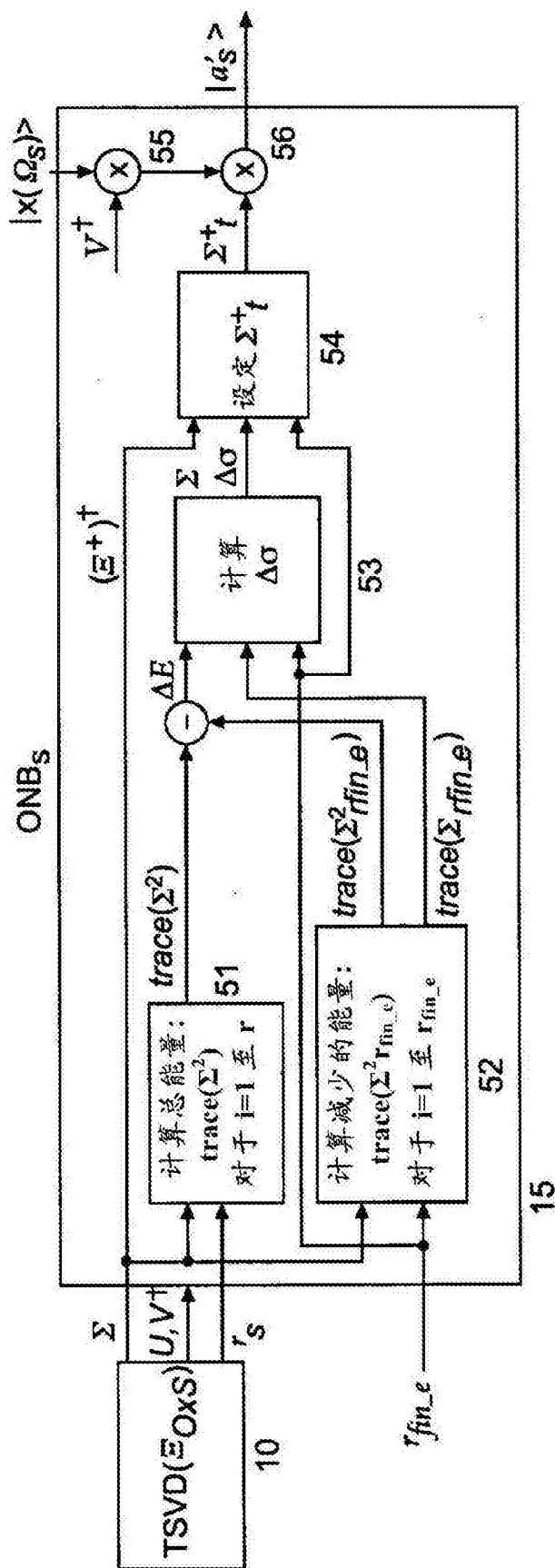


图5

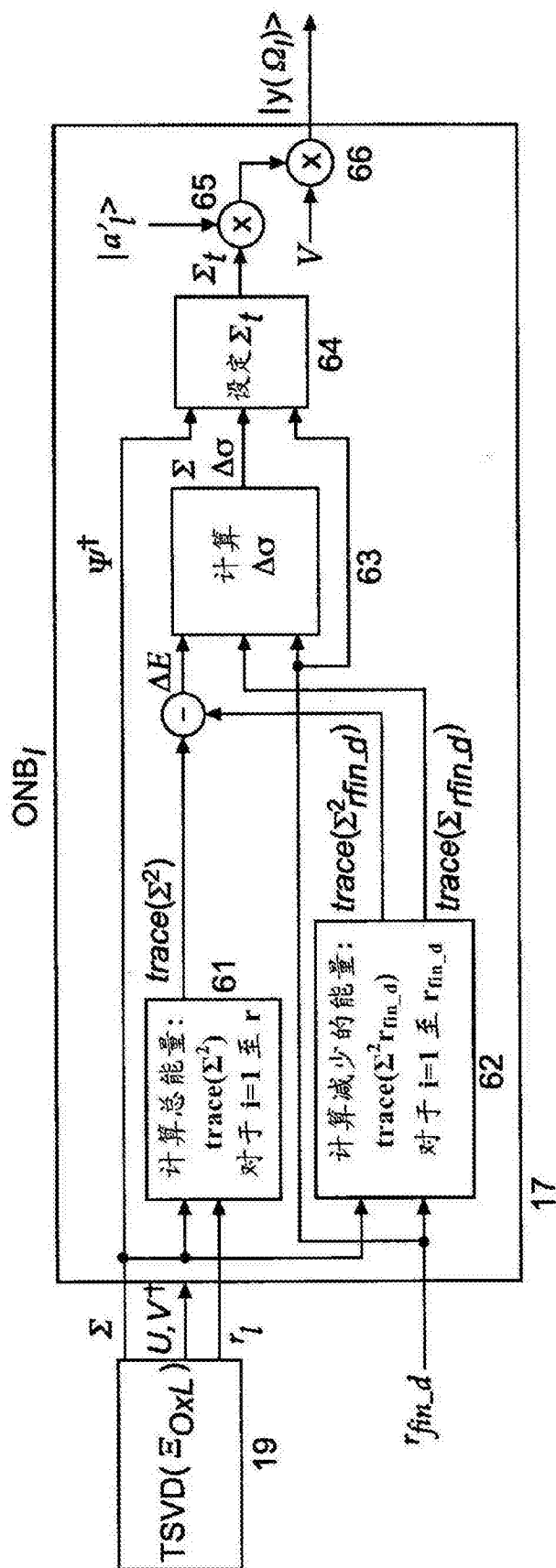


图6