(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 106714074 B
(45)授权公告日 2019.09.24

(21)申请号 201710163516.1

(22)申请日 2013.03.06

(65)同一申请的已公布的文献号
申请公布号 CN 106714074 A

(43)申请公布日 2017.05.24

(30)优先权数据
12305271.4 2012.03.06 EP

(62)分案原申请数据
201310070648.1 2013.03.06

(73)专利权人 杜比国际公司
地址 荷兰阿姆斯特丹

(72)发明人 P.贾克斯 J.贝姆 W.G.雷德曼

(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 宿小猛

(51)Int.Cl.
H04S 7/00(2006.01)

(56)对比文件
CN 1419796 A,2003.05.21,
CN 102326417 A,2012.01.18,
US 2010328419 A1,2010.12.30,
US 2009238371 A1,2009.09.24,
US 2010328423 A1,2010.12.30,

审查员 李雁

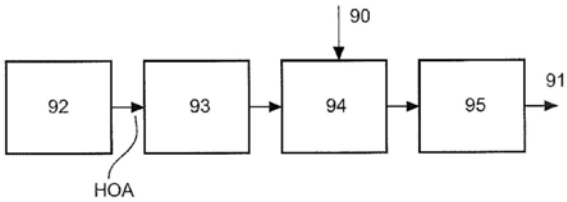
权利要求书1页 说明书8页 附图5页

(54)发明名称

用于回放更高阶立体混响音频信号的方法和设备

(57)摘要

本发明涉及用于回放更高阶立体混响音频信号的方法和设备。本发明允许面向空间声场音频的回放通过应用在EP 11305845.7中公开的空间弯曲处理对其链接的可见对象的系统性适配。在内容再现中使用的屏幕的基准尺寸(或从基准聆听位置的观看角度)被编码并作为元数据与内容一起被传输,或者解码器知道目标屏幕相对于固定基准屏幕尺寸的实际尺寸。解码器以如下方式弯曲声场:根据目标屏幕的尺寸和基准屏幕的尺寸的比率,压缩或拉伸屏幕方向上的全部声音对象。



CN 106714074 B

1.一种用于生成与目标屏幕尺寸相关联的扬声器信号的方法,所述方法包括:

接收包含编码的更高阶立体混响信号的比特流,所述编码的更高阶立体混响信号描述与制造屏幕尺寸相关联的声场;

解码所述编码的更高阶立体混响信号以获得表示所述声场的主要分量的第一组解码的更高阶立体混响信号和表示所述声场的环境分量的第二组解码的更高阶立体混响信号;

组合已经历了从所述制造屏幕尺寸到所述目标屏幕尺寸的适配的第一组解码的更高阶立体混响信号、和第二组解码的更高阶立体混响信号,以产生组合的一组解码的更高阶立体混响信号;以及

通过重现所述组合的一组解码的更高阶立体混响信号来生成所述扬声器信号,

其中,所述重现还包括将变换矩阵应用于所述组合的一组解码的更高阶立体混响信号,其中所述变换矩阵被从用于扬声器的规则间隔位置的第一模式矩阵、与基于所述目标屏幕尺寸和所述制造屏幕尺寸从所述规则间隔位置映射到的位置有关的第二矩阵、以及对角矩阵导出,所述对角矩阵具有增益函数的值作为所述对角矩阵的主对角的分量。

2.一种用于生成与目标屏幕尺寸相关联的扬声器信号的装置,所述装置包括:

接收器,用于获得包含编码的更高阶立体混响信号的比特流,所述编码的更高阶立体混响信号描述与制造屏幕尺寸相关联的声场;

音频解码器,用于对编码的更高阶立体混响信号进行解码,以获得表示所述声场的主要分量的第一组解码的更高阶立体混响信号和表示所述声场的环境分量的第二组解码的更高阶立体混响信号;

组合器,用于集成已经历了从所述制造屏幕尺寸到所述目标屏幕尺寸的适配的第一组解码的更高阶立体混响信号、和第二组解码的更高阶立体混响信号,以产生组合的一组解码的更高阶立体混响信号;以及

发生器,用于通过重现所述组合的一组解码的更高阶立体混响信号来产生所述扬声器信号,

其中,所述重现还包括将变换矩阵应用于所述组合的一组解码的更高阶立体混响信号,其中所述变换矩阵被从用于扬声器的规则间隔位置的第一模式矩阵、与基于所述目标屏幕尺寸和所述制造屏幕尺寸从所述规则间隔位置映射到的位置有关的第二矩阵、以及对角矩阵导出,所述对角矩阵具有增益函数的值作为所述对角矩阵的主对角的分量。

用于回放更高阶立体混响音频信号的方法和设备

[0001] 本申请是基于申请号为201310070648.1、申请日为2013年03月06日、发明名称为“用于回放更高阶立体混响音频信号的方法和设备”的专利申请的分案申请。

技术领域

[0002] 本发明涉及用于回放指定给视频信号的更高阶立体混响(Higher-Order Ambisonics)音频信号的方法和设备,该视频信号是对原始和不同的屏幕而生成的但是将被呈现在当前屏幕上。

背景技术

[0003] 一种存储和处理球形麦克风阵列的三维声场的方式是更高阶立体混响(HOA)表示。立体混响使用标准正交球形函数用于描述位于原点或空间中基准点(也称为蜜点)及其附近的区域中的声场。如此描述的精度通过立体混响阶N来确定,其中,有限数的立体混响系数描述声场。球形阵列的最大立体混响阶通过麦克风精盒(microphone capsule)的数量来限制,所述数量必须等于或大于立体混响系数的数量$O=(N+1)^2$。这样的立体混响表示的优势在于声场的再现可以单独地适配于近乎任何给定扬声器位置排列。

发明内容

[0004] 虽然促进空间音频的灵活通用的表示非常独立于扬声器设置,与不同尺寸屏幕上的音频回放的组合可以变得分散,因为空间声音回放未被相应地适配。

[0005] 立体和环绕声基于离散扬声器声道,并且涉及视频显示关于在哪里放置扬声器存在非常特定的规则。例如,在影院环境下,在屏幕的中央放置中央扬声器,而在屏幕的左侧和右侧放置左扬声器和右扬声器。因而,扬声器设置固有地随屏幕变动:对于小屏幕,扬声器彼此更接近,而对于巨屏,它们则远离。这样的优势是混音可以以非常连贯的方式完成:有关屏幕上的可视对象的声音对象可以可靠地放置在左声道、中央声道和右声道。因此,听众的体验在混音级上匹配声音艺术家的创造性意图。

[0006] 但是这样的优势同时也是基于声道系统的劣势:对于改变扬声器设置,灵活度非常有限。该劣势随着扬声器声道数量的增加而增加。例如,7.1和22.2格式要求各个扬声器的精确安装并且极其难以适配音频内容到次佳扬声器位置。

[0007] 基于声道系统的另一劣势是:优先效应限制了在左声道、中央声道和右声道之间摇移(pan)声音对象的能力,特别是对于类似影院环境下的大型聆听设置。对于偏心聆听位置,摇移的音频对象可以“落在”最靠近听众的扬声器上。因而,很多电影已经和重要的屏幕相关的声音混音,特别是排他地映射到中央声道上的对话,借此,获得屏幕上那些声音的非常稳定的定位,但是是以整体声音现场的次佳宽敞度为代价。

[0008] 通常对后置环绕声道选择类似的妥协:因为播放那些声道的扬声器的精确定位在制作时难以知晓,并且因为那些声道的密度相当低,通常仅环境声音和未校正项被混音到环绕声道。因而,环绕声道中的明显再现错误的概率可以降低,但是是以不能如实地在任何

地点而是在屏幕上(或甚至在如上所述的中央声道上)放置离散声音对象为代价。

[0009] 如上所述,在不同尺寸屏幕上的空间音频和视频回放的组合可以变得分散,因为空间声音回放未被相应地适配。取决于实际屏幕尺寸是否匹配在再现中使用的尺寸,声音对象的方向可以偏离屏幕上可视对象的方向。例如,如果已经在小屏幕的环境中进行了混音,则耦合到屏幕对象的声音对象(例如,演员的发音)将被定位在从混音器的位置看去相对狭窄的锥体。如果此内容受控于基于声场的表示并且在具有大很多的屏幕的影院环境中回放,则在屏幕的宽视场与屏幕相关声音对象的窄锥体之间存在明显失配。对象的可视图像的位置和对应声音的位置之间的巨大失配会分散观众注意力并且因而严重地影响电影的感知。

[0010] 更近地,已经提出音频现场的参数表示或面向对象表示,其通过单独音频对象和参数及特性的集合的组合来描述音频现场。例如,主要对处理波场综合系统已经提出面向对象现场描述,例如,在Sandra Brix、Thomas Sporer、Jan Plogsties于Proc.of 110th AES Convention,Paper 5314,2001年5月12-15日,在荷兰阿姆斯特丹发表的“CARROUSO-An European Approach to 3D-Audio”,以及在Ulrich Horbach、Etienne Corteel、Renato S.Pellegrini和Edo Hulsebos于Proc.of IEEE Intl.Conf.on Multimedia and Expo (ICME),pp.517-520,2002年8月,瑞士Lausanne,发表的“Real-Time Rendering of Dynamic Scenes Using Wave Field Synthesis”中。

[0011] EP 1518443 B1描述了应对使音频回放适配于可视屏幕尺寸的问题的两种不同途径。第一种途径对每一个声音对象取决于其对基准点方向和距离以及与相机和投影装备的孔径角(opening angle)和位置类似的参数,单独确定回放位置。实际上,在对象的可见性和相关混音之间的如此紧密的耦合不是典型的,相反,混音与相关可见对象的某些偏差可以实际上出于艺术原因而容忍。此外,区分直达声和环境声音是重要的。最后但不是最不重要的,物理相机和投影参数的合并相当复杂,并且这样的参数不总是可用。第二种途径(比较权利要求16)描述了根据以上步骤的声音对象的预计算,但是假设屏幕具有固定的基准尺寸。该方案要求全部位置参数(在笛卡尔坐标中)的线性缩放用于将屏幕适配到比基准屏幕大或小的屏幕上。然而,这意味着,适配到双倍尺寸屏幕也导致到声音对象的虚拟距离的翻倍。没有相对于基准座位(即,蜜点)中的听众的声音对象的角度位置中的任何改变,这只是听觉现场的“气息音”而已。对于角坐标中屏幕的相对尺寸(孔径角)的改变,不可能通过此途径产生可行的聆听结果。

[0012] 在EP 1318502 B1中描述面向对象声音现场描述格式的另一示例。这里,除了包括不同的声音对象及其特性之外,音频现场还包括关于要再现的房间的特性的信息以及关于基准屏幕的水平和垂直孔径角的信息。在解码器中,类似于EP 1518443 B1中的原理,确定实际可用屏幕的位置和尺寸,并且单独地最优化声音对象的回放以匹配基准屏幕。

[0013] 例如,在PCT/EP2011/068782中,已经对声场的通用空间表示提出了面向声场的类似更高阶立体混响HOA的音频格式,并且在记录和回放方面,面向声场的处理提供通用性和实践性之间极好的平衡,因为它可以缩放到实际上任意的空间分辨率,类似于面向对象格式的那样。另一方面,一些直接记录和再现技术存在,其对比对面向对象格式要求的完全合成的表示,允许得到真实声场的自然记录。显然,因为面向声场的音频内容不包括关于单独声音对象的任何信息,所以以上对使面向对象格式适配到不同屏幕尺寸而介绍的机制不能

得以应用。

[0014] 如今,仅有少量描述用于控制在面向声场的音频现场中包含的单独声音对象的相对位置的装置的公开可用。例如在Richard Schultz-Amling、FabianKuech、Oliver Thiergart、Markus Kallinger于2010年5月22-25日在英国伦敦的“Acoustical Zooming Based on a Parametric Sound Field Representation”,128th AES Convention,Paper 8120中描述的一系列算法要求将声场分解为有限数量的离散声音对象。可以控制这些声音对象的位置参数。此途径具有如下劣势:音频现场分解易出错并且在确定音频对象时的任何错误将极可能导致声音重现的人为产物。

[0015] 很多出版物涉及将HOA内容的回复最优化到“灵活回放布局”,例如,以上引用的Brix文章以及Franz Zotter、HannesPomberger、Markus Noisternig于2010年5月6-7日在法国巴黎的Proc.of the 2nd International Symposium on Ambisonics and Spherical Acoustics上的“Ambisonic Decoding With and Without Mode-Matching:A Case Study Using the Hemisphere”。这些技术处理使用不规则间隔的扬声器的问题,但是它们都没有对准改变音频现场的空间构成。

[0016] 本发明要解决的问题是空间音频内容对不同尺寸的视频屏幕的适配,该音频内容已被表示为声场分解的系数,从而屏幕上对象的声音回复位置与对应的可视位置匹配。通过在权利要求1中公开的方法解决此问题。在权利要求2中公开了使用此方法的设备。

[0017] 本发明允许面向空间声场的音频的回放对其链接可视对象的系统性适配。因而,满足了用于对电影的空间音频的可信再现的显然的先决条件。

[0018] 根据本发明,结合诸如在PCT/EP2011/068782和EP 11192988.0中公开的那些格式之类的面向声场的音频格式,通过应用在EP 11305845.7中公开的空间弯曲处理将面向声场的音频现场被适配到不同的视频屏幕尺寸。有利处理是编码并传输与内容一起在内容制作中使用的屏幕的基准尺寸(或者从基准聆听位置的视角)作为元数据。

[0019] 替代地,在编码中和对解码假定固定的基准屏幕尺寸,并且解码器知道目标屏幕的实际尺寸。解码器以如下方式弯曲声场:根据目标屏幕的尺寸和基准屏幕的尺寸的比率压缩或拉伸在屏幕的方向上的全部声音对象。这可以借助于例如如下说明的简单的双段分段线性弯曲函数(two-segment piecewise linear warping function)完成。与上述现有技术相反,此拉伸基本上限制于声音项的角位置,并且不需要导致声音对象距聆听区域的距离的改变。在下面描述发明的一些实施例,其允许控制声音现场的哪部分应当或不应当被操控。

[0020] 原则上,发明方法适用于回放指定给视频信号的原始更高阶立体混响音频信号的方法,上述视频信号是对原始和不同的屏幕而生成的但是将被呈现在当前屏幕上,所述方法包括如下步骤:

[0021] -解码所述更高阶立体混响音频信号以提供解码的音频信号;

[0022] -接收或建立从所述原始屏幕和所述当前屏幕在它们的宽度并且可能在它们的高度并且可能在它们的弯曲度之间的差异得出的再现适配信息;

[0023] -通过在空间域中弯曲它们来适配所述解码的音频信号,其中,所述再现适配信息控制所述弯曲,从而对于当前屏幕的观众和所述适配的解码的音频信号的听众,由所述适配的解码的音频信号表示的至少一个音频对象的感知位置匹配所述屏幕上的相关视频对

象的感知位置；

[0024] -对扬声器重现和输出适配的解码的音频信号。

[0025] 原则上，发明设备适用于回放指定给视频信号的原始更高阶立体混响音频信号，所述视频信号是对原始和不同的屏幕而生成的但是将被呈现在当前屏幕上，所述设备包括：

[0026] -适配于解码所述更高阶立体混响音频信号以提供解码的音频信号的装置；

[0027] -适配于接收或建立从所述原始屏幕和所述当前屏幕在它们的宽度并且可能在它们的高度并且可能在它们的弯曲度之间的差异得出的再现适配信息的装置；

[0028] -适配于通过在空间域中弯曲它们来适配所述解码的音频信号的装置，其中，所述再现适配信息控制所述弯曲，从而对于当前屏幕的观众和所述适配的解码的音频信号的听众，由所述适配的解码的音频信号表示的至少一个音频对象的感知位置匹配所述屏幕上的相关视频对象的感知位置；

[0029] -适配于对扬声器重现和输出适配的解码的音频信号的装置。

[0030] 本发明的有利的额外实施例在各自的从属权利要求中公开。

附图说明

[0031] 参考附图描述本发明的示例性实施例，其示出：

[0032] 图1示例工作室环境；

[0033] 图2示例影院环境；

[0034] 图3弯曲函数f(Φ)；

[0035] 图4权函数g(Φ)；

[0036] 图5原始权重；

[0037] 图6弯曲之后的权重；

[0038] 图7弯曲矩阵；

[0039] 图8已知HOA处理；

[0040] 图9根据本发明的处理。

具体实施方式

[0041] 图1示出具有基准点和屏幕的示例工作室环境，而图2示出具有基准点和屏幕的示例影院环境。不同的投影环境导致从基准点看去屏幕的不同孔径角。借助于现有技术的面向声场回放技术，在工作室环境中产生的音频内容(孔径角60°)将不匹配影院环境中的屏幕内容(孔径角90°)。在工作室环境中的孔径角60°必须同音频内容一起传输以便于允许内容对回放环境的不同特性的适配。

[0042] 为了易理解，这些图简化情形为2D场景。

[0043] 在更高阶立体混响理论中，经由傅里叶巴塞尔序列的系数$A_n^m(k)$描述空间音频场景。对于无源列(source-free volume)，声压被描述为球坐标的函数(半径r，倾角θ，方位角Φ和空间频率$k=\frac{\omega}{c}$(c是声音在空气中的速度))：

[0044] $$p(r,\theta,\phi,k)=\sum_{n=0}^{N}\sum_{m=-n}^{n}A_n^m(k)j_n(kr)\,Y_n^m(\theta,\phi)$$，

[0045] 其中，$j_n(kr)$是第一类的球型巴塞尔函数，其描述了径向从属性，$Y_n^m(\theta,\phi)$是球调和函数(SH,Spherical Harmonics)，其实际上是实数，而N是立体混响阶。

[0046] 通过在EP 11305845.7中公开的技术可以弯曲音频现场的空间构成。

[0047] 可以改变音频现场的二维或三维更高阶立体混响HOA表示中包含的声音对象的相对位置，其中，具有维度O_{in}的输入向量A_{in}确定输入信号的傅里叶序列的系数，而具有维度O_{out}的输出向量A_{out}确定相应改变的输出信号的傅里叶序列的系数。使用模式矩阵ψ_1的逆ψ_1^{-1}通过计算$s_{in}=\psi_1^{-1}A_{in}$将输入HOA系数的输入向量A_{in}解码为针对规则布置的扬声器位置的空间域中的输入信号s_{in}。通过计算$A_{out}=\Psi_2 s_{in}$在空间域中将输入信号s_{in}弯曲并解码为已适配输出HOA系数的输出向量A_{out}，其中根据弯曲函数f(Φ)修改模式矩阵ψ_2的模式向量，借助于该弯曲函数f(Φ)，原始扬声器位置的角度被一对一地映射到输出向量A_{out}中的目标扬声器位置的目标角。

[0048] 可以通过向虚拟扬声器输出信号s_{in}应用增益加权函数g(Φ)对抗(counter)扬声器密度的修改，导致信号s_{out}。原则上，可以指定任何加权函数g(Φ)。已经经验性地确定一个特别有利的变量为与弯曲函数f(Φ)的导数成比例：$g(\phi)=\frac{\mathbf{d}f_\phi(\phi)}{\mathbf{d}\phi}$。借助于此特定加权函数，假定适当高的内阶和输出阶，在特定弯曲角度的摇移函数f(Φ)的幅度保持等于在原始角Φ的原始摇移函数。因而，获得每个孔径角的同类声音平衡(幅度)。对于三维立体混响，增益函数在Φ方向和在θ方向上是

[0049] $$g(\theta,\phi)=\frac{\mathrm{d}f_\theta(\theta)}{\mathrm{d}\theta}\cdot\frac{\arccos\left((\cos f_\theta(\theta_{in}))^2+(\sin f_\theta(\theta_{in}))^2\cos\phi_\varepsilon\right)}{\arccos\left((\cos\theta_{in})^2+(\sin\theta_{in})^2\cos\phi_\varepsilon\right)}$$，其中，ϕ_ε是小方位角。

[0050] 通过使用尺寸$O_{warp}\times O_{warp}$变换矩阵$\mathbf{T}=\mathbf{diag(w)}\,\Psi_2\,\mathbf{diag(g)}\Psi_1^{-1}$可以共同地进行解码、加权和弯曲/解码，其中，diag(w)表示具有窗口向量值w作为其主对角的分量的对角矩阵，diag(g)表示具有增益函数值g作为其增益对角的分量的对角矩阵。为了变形变换矩阵T以获得尺寸$O_{out}\times O_{in}$，变换矩阵T的对应列和/或线被移除以进行空间弯曲操作$A_{out}=TA_{in}$。

[0051] 图3到图7图示了在二维(圆形)情况下的空间弯曲，并且示出了用于图1/2中的情形的分段线性弯曲函数的示例及其对13个规则布置的示例扬声器的摇移函数的影响。系统以1.5的因子拉伸在前方的声场以适配在影院中的较大屏幕。因此，来自其他方向的声音项被压缩。弯曲函数f(Φ)类似于具有单一实数参数的离散时间全通滤波器的相位响应，并且在图3中示出。对应的加权函数g(Φ)在图4中示出。

[0052] 图7描绘了13×65个单步变换弯曲矩阵T。矩阵的单独系数的对数性绝对值用根据所附灰度或阴影条的灰度或阴影型指示。已经对$N_{orig}=6$的输入HOA阶和$N_{warp}=32$的输出阶设计此示例矩阵。要求较高的输出阶以便于捕获由从低阶系数到高阶系数的变换展开的大部分信息。

[0053] 此特定弯曲矩阵的有用特性在于其有效部分为零。这允许在实现此操作时节约大量的计算能力。图5和图6图示了由一些平面波产生的波束图案的弯曲特性。两个图都是从

在Φ位置0、2/13π、4/13π、6/13π、...、22/13π和22/13π的相同十三个输入平面波得出的，全部具有一致的幅度“一”，且示出十三个角幅度分布，即，超定的结果向量s，规则解码操作$s=\Psi^{-1}A$，其中，HOA向量A是平面波的集合的或原始或弯曲的变量。圆之外的数字表示角度Φ。虚拟扬声器的数量可观地高于HOA参数的数量。用于来自前方的平面波的幅度分布或波束图案位于Φ＝0。

[0054] 图5示出原始HOA表示的权重和幅度分布。全部十三个分布都相似地形成并且突出主波瓣的相同宽度。图6示出对相同声音对象的权重和幅度分布，但是是在已经进行弯曲操作之后。对象已经从Φ＝0的前方远离并且该前方附近的主波瓣变得更加宽阔。通过更高阶$N_{warp}=32$的弯曲HOA向量促进波束图案的这些修改。用在空间中变化的局部阶创建混阶(mixed-order)信号。

[0055] 为了得出对将音频现场的回放适配于实际屏幕配置的合适弯曲特性$f(\Phi_{in})$，除了HOA系数之外还发送或提供了额外信息。例如，在混音处理中使用的基准屏幕的以下特性可以被包括在位流中：

[0056] • 屏幕中央的方向，

[0057] • 宽度，

[0058] • 基准屏幕的高度，

[0059] 全部都在从基准聆听位置测量(即，“蜜点”)的极化坐标中。

[0060] 另外，对特殊应用可以要求以下参数：

[0061] • 屏幕的形状，例如，它是平的或是球形，

[0062] • 屏幕的距离，

[0063] • 关于在立体3D视频投影情况中的最大和最小可视深度的信息。

[0064] 对于本领域技术人员已知这样的元数据如何编码。

[0065] 接着，假定编码的音频位流包括至少以上三个参数，中央的方向、基准屏幕的宽度和高度。为了理解，进一步假定实际屏幕的中央与基准屏幕的中央一致，例如，直接在听众的前方。此外，假定，仅以2D格式表示声场(相比于3D格式)并且这个的倾角的改变被忽略(例如，如当选择的HOA格式表示无垂直分量时，或者其中声音编辑认为图片和屏幕上声源的倾角之间的失配将足够小从而普通观察者将不会注意到它们。)对任意屏幕位置和3D情况的转变对于本领域技术人员是直接的。进一步，出于简单而假定屏幕结构是球型。

[0066] 借助这些假定，仅屏幕的宽度可以在内容和实际设置之间变化。下面，定义合适的两段分段式线性弯曲特性。由孔径角$2\Phi_{w,a}$定义实际屏幕宽度(即，$\Phi_{w,a}$描述半角)。由角度$\Phi_{w,r}$定义基准屏幕宽度，并且此值是在位流内传递的元信息的部分。对于在前方上(即，在视频屏幕上)的声音对象的可信再现，声音对象的全部位置(在极化坐标中)将由因子$\Phi_{w,a}/\Phi_{w,r}$操控。相反，在其他方向上的全部声音对象应当根据剩余空间而移动。弯曲特性导致

[0067] $$\phi_{out}=\begin{cases}\phi_{w,a}/\phi_{w,r}\cdot\phi_{in} & -\phi_{w,r}\le\phi_{in}\le\phi_{w,r}\\ \dfrac{(\pi-\phi_{w,a})}{(\pi-\phi_{w,r})}\cdot[\phi_{in}-\pi]+\pi & \text{否则}\end{cases},$$

[0068] 对获得此特性所要求的弯曲操作可以以在EP 11305845.7中公开的规则构建。例如，其结果是，可以得出单步线性弯曲运算符，该运算符在所操控的向量被输入HOA重现处

理之前被应用于每一个HOA向量。以上示例是很多可能弯曲特性中的一个。可以应用其他特性以便于找到复杂度和在操作之后剩余的失真量之间的平衡。例如，如果应用简单分段线性弯曲特性用于制造3D声场重现，则可以产生空间再现的典型枕型失真和桶形失真，但是如果因子$\Phi_{w,a}/\Phi_{w,r}$接近“一”，可以忽略空间重现的这样的失真。对于很大或很小的因子，可以应用更复杂的弯曲特性，其最小化空间失真。

[0069] 另外，如果所选择的HOA表示确实规定了倾角且声音编辑认为屏幕所对的垂直角是重要的，则可以对倾角应用基于屏幕的角高度θ_h(半高)和有关因子(例如，实际高度对基准高度的比率$\theta_{h,a}/\theta_{h,r}$)的类似等式作为弯曲操作符的部分。

[0070] 如另一示例，假定在听众前方，取代球形屏幕的纯平屏幕可能要求比上述示例性特性更为精致的弯曲特性。再次，这可以仅以宽度或仅以宽度+高度弯曲关注其自身。

[0071] 上述示例性实施例具有固定和极易于实现的优势。另一方面，不允许来自生产侧的适配处理的任何控制。以下实施例介绍用于以不同方式的更多控制的处理。

[0072] 实施例1：屏幕相关声音和其他声音之间的分离

[0073] 由于各种原因可能要求这样的控制技术。例如，不是音频现场中的全部声音对象直接与屏幕上的可见对象耦合，并且操控不同于环境声的直达声可以是有利的。可以在重现侧通过现场分析进行此区分。然而，通过向传输位流增加额外信息可以显著地改善和控制。理想化地，对实际屏幕特性适配何种声音项以及何种声音项不处理的决定应当留给进行声音混音的艺术家。

[0074] 向重现处理传输此信息的不同方式是可能的：

[0075] • 在位流内定义两个完全集合的HOA系数(信号)，一个用于描述有关可见项的对象而另一个用于表示独立或环境声音。在解码器中，仅第一HOA信号将经历对实际屏幕布局(geometry)的适配而另一个则未处理。在回放之前，组合所操控的第一HOA信号和未修改第二HOA信号。

[0076] 作为示例，声音工程师可以决定将类似对话的屏幕相关声音或特定的弗雷(Foley)项混合到第一信号中，并且将环境声音混合到第二新号中。以此方式，无论对音频/视频信号的回放使用哪个屏幕，环境将总是保持一致。

[0077] 这种处理具有额外优势，可以对特定类型的信号单独地最优化两个构成子信号的HOA阶，由此，用于屏幕相关声音对象(即，第一子信号)的HOA阶高于对环境信号分量(即，第二子声音)使用的HOA阶。

[0078] • 经由附着于时空频率瓦(tile)的标识，将声音的映射定义为屏幕相关的或屏幕独立的。出于此目的，例如，经由平面波分解确定HOA信号的空间特性。然后，对时间分段(分窗口)和时频转换输入每一个空间域信号。由此，将定义瓦的三维集合，其可以例如由声明该瓦的内容是否应当被适配于实际屏幕布局的二进制标识单独地标记。此子实施例比之前的子实施例更为有效，但是它限制了定义声音现场的哪个部分应当被操控或不被操控的灵活性。

[0079] 实施例2：动态适配

[0080] 在一些应用中，将要求以动态方式改变被发信号(signalled)的基准屏幕特性。例如，音频内容可以是从不同混音连结重新调整的内容分段的结果。在此情况下，描述基准屏幕参数的参数将随时间改变，并且动态地改变适配算法：对于屏幕参数的每一个改变，对应

地重新计算所应用的弯曲函数。

[0081] 另一应用示例起于混合对最终可视视频和音频现场的不同子部分已准备的不同HOA流。然后,有利的是,考虑处于共同位流中的多于一个(或以上用实施例1多于两个)HOA信号,每一个都具有其单独的屏幕特性。

[0082] 实施例3:替代实现

[0083] 取代在经由固定HOA解码器的解码之前的弯曲HOA表示,关于如何适配信号到实际屏幕特性的信息可以被集成到解码器设计中。此实现是对在以上示例性实施例中描述的基本实现的替代。然而,它不改变在位流内的屏幕特性的发信号。

[0084] 在图8中,HOA编码的信号被存储在存储设备82中。为了在影院中的呈现,来自设备82的HOA表示的信号在HOA解码器83中被HOA解码,穿过再现器85,并且对于一组扬声器输出为扬声器信号81。

[0085] 在图9中,HOA编码的信号被存储在存储设备92中。为了例如在影院中的呈现,来自设备92的HOA表示的信号在HOA解码器93中被HOA解码,穿过弯曲级94到再现器95,并且对于一组扬声器输出为扬声器信号91。弯曲级94接收上述再现适配信息90,并相应地使用它用于适配已解码的HOA信号。

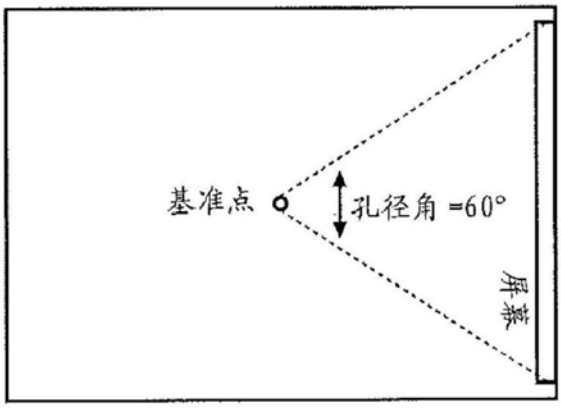


图1

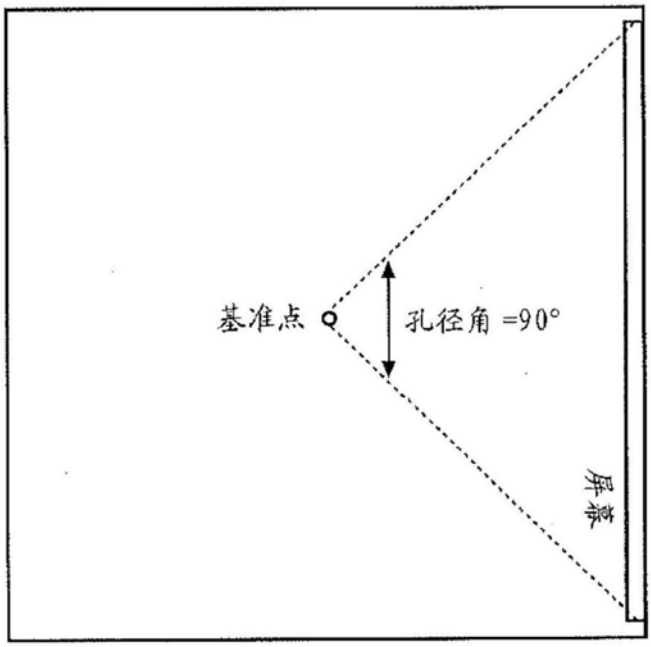


图2

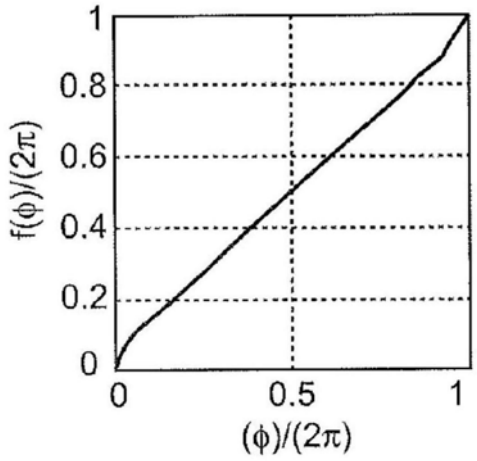


图3

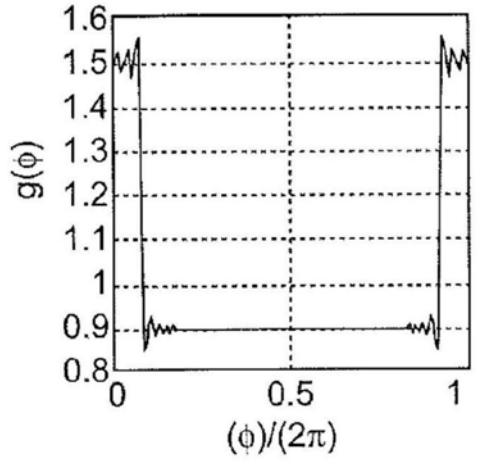


图4

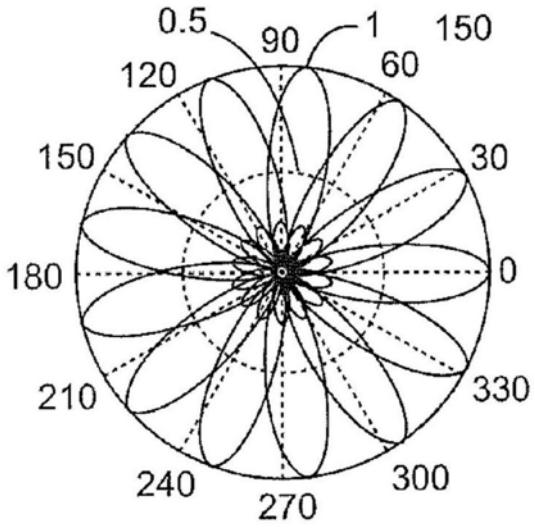

0.5
90
1
150
120
60
150
30
180
0
210
330
240
300
270

图5

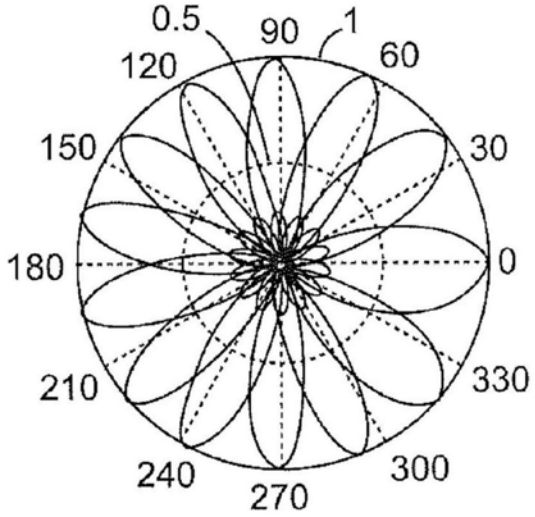

0.5
90
1
120
60
150
30
180
0
210
330
240
300
270

图6

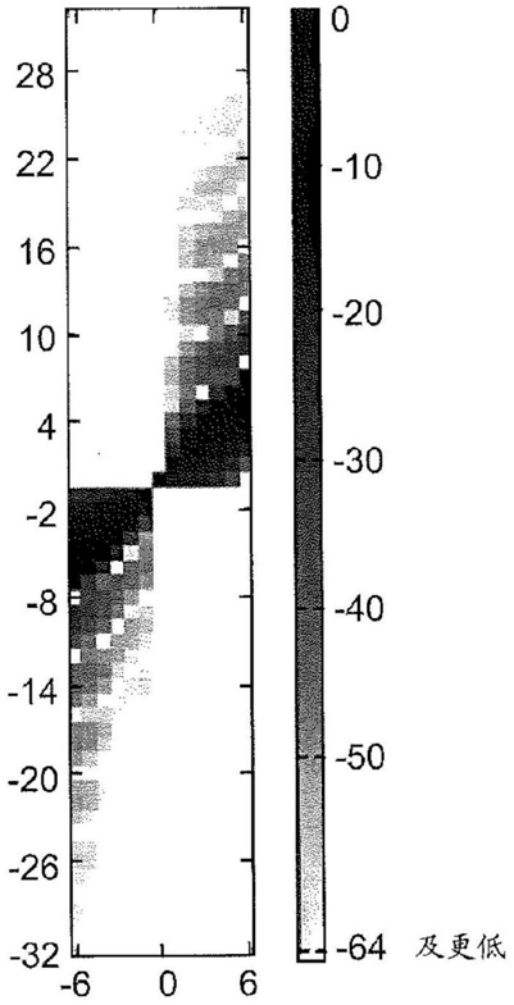


图7

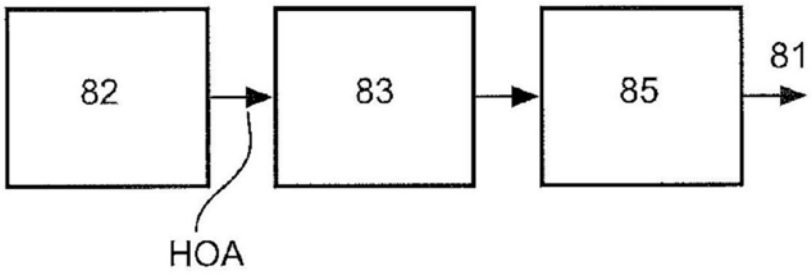


图8

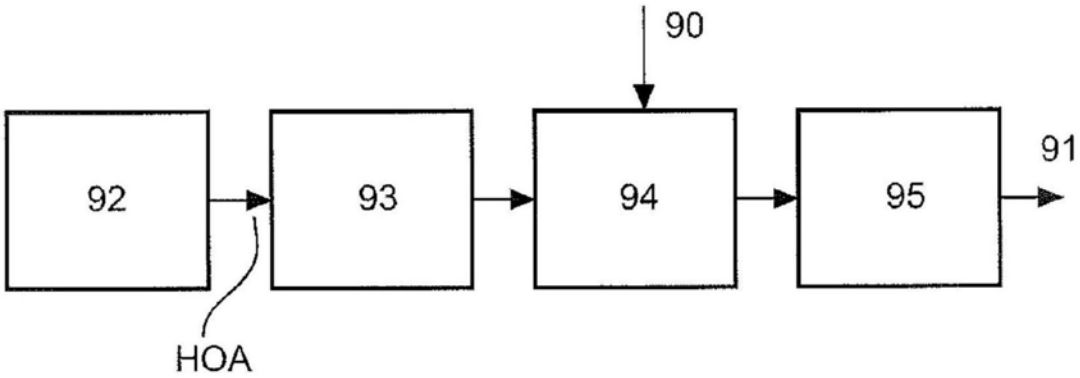


图9