

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200880016796.2

[51] Int. Cl.

G11B 27/038 (2006.01)

H04S 3/00 (2006.01)

H04R 3/12 (2006.01)

[43] 公开日 2010 年 3 月 24 日

[11] 公开号 CN 101681663A

[22] 申请日 2008.5.21

[21] 申请号 200880016796.2

[30] 优先权

[32] 2007.5.22 [33] EP [31] 07108601.1

[86] 国际申请 PCT/IB2008/051998 2008.5.21

[87] 国际公布 WO2008/142651 英 2008.11.27

[85] 进入国家阶段日期 2009.11.20

[71] 申请人 皇家飞利浦电子股份有限公司

地址 荷兰艾恩德霍芬

[72] 发明人 A·S·哈马

S·L·J·D·E·范德帕

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 龚海军 刘红

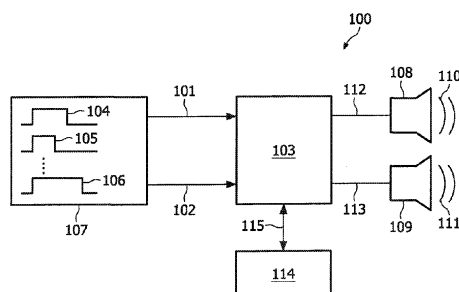
权利要求书 4 页 说明书 17 页 附图 5 页

[54] 发明名称

处理音频数据的设备和方法

[57] 摘要

根据本发明的一个示例性实施例，提供了一种用于处理音频数据(101, 102)的设备(100)，其中该设备(100)包括操纵单元(103)(特别地，重采样单元)，其适用于以修改第一音频项目(104)的转换部分的与时间有关的音频属性的方式选择性地操纵(特别地，重采样)第一音频项目(104)的转换部分(特别地，也可以以逼真的方式模拟运动的时间延迟效果)。



1. 一种用于处理音频数据(101, 102)的设备(100), 其中该设备(100)包括操纵单元(103), 其适用于以在所述音频数据(101, 102)的第一音频项目(104)的转换部分中选择性地修改所述音频数据(101, 102)的第一音频项目(104)的与时间有关的音频属性的方式操纵所述音频数据(101, 102)的第一音频项目(104)的转换部分。

2. 根据权利要求1的设备(100), 其中所述第一音频项目(104)的转换部分是所述第一音频项目(104)的结尾部分。

3. 根据权利要求2的设备(100), 其中所述操纵单元(103)适用于以减小由所述第一音频项目(104)的经操纵的结尾部分的节奏、音高和频率组成的组中的至少一个的方式操纵所述第一音频项目(104)的结尾部分。

4. 根据权利要求1的设备(100), 其中所述操纵单元(103)适用于以在所述音频数据(101, 102)的第二音频项目(105)的转换部分中选择性地修改所述音频数据(101, 102)的第二音频项目(105)的与时间有关的音频属性的方式操纵所述音频数据(101, 102)的第二音频项目(105)的转换部分。

5. 根据权利要求4的设备(100), 其中所述第二音频项目(105)的转换部分是所述第二音频项目(105)的开头部分。

6. 根据权利要求5的设备(100), 其中所述操纵单元(103)适用于以增大由所述第二音频项目(105)的经操纵的开头部分的节奏和频率组成的组中的至少一个的方式操纵所述第二音频项目(105)的开头部分。

7. 根据权利要求1的设备(100), 其中所述操纵单元(103)适用于排他地操纵所述第一音频项目(104)的所述转换部分或多个转换部分, 而所述第一音频项目(104)的剩余部分保持未被操纵。

8. 根据权利要求4的设备(100), 其中所述操纵单元(103)适用于以协同方式操纵所述第一音频项目(104)的转换部分和所述第二音频项目(105)的转换部分, 以便再现所述第一音频项目(104)和随后的第二音频项目(105)。

9. 根据权利要求1的设备(100), 其中所述操纵单元(103)适用于以生成再现所述第一音频项目(104)的音频源在所述转换部分期间正在移动的可听体验的方式处理所述第一音频项目(104)。

10. 根据权利要求9的设备(100),其中所述操纵单元(103)适用于生成再现所述第一音频项目(104)的音频源在所述第一音频项目(104)的结尾部分期间正在离开的可听体验。

11. 根据权利要求4和9的设备(100),其中所述操纵单元(103)适用于以生成再现所述第二音频项目(105)的音频源在所述转换部分期间正在移动的可听体验的方式处理所述第二音频项目(105)。

12. 根据权利要求11的设备(100),其中所述操纵单元(103)适用于生成再现所述第二音频项目(105)的音频源在所述第二音频项目(105)的开头部分期间正在接近的可听体验。

13. 根据权利要求11的设备(100),其中所述操纵单元(103)适用于按照以下顺序生成所述第一音频项目(104)的结尾部分和所述第二音频项目(105)的开头部分之间的转换:

处理所述第二音频项目(105)的转换部分,使得所述第二音频项目(105)的转换部分的再现为可感知为源于遥远的起始位置;

处理所述第一音频项目(104)的转换部分,使得所述第一音频项目(104)的转换部分的再现可感知为源于正在从中央位置移动到遥远的最终位置的位置;

在处理所述第一音频项目(104)的转换部分的同时,处理所述第二音频项目(105)的转换部分,使得所述第二音频项目(105)的转换部分的再现可感知为源于正在从遥远的起始位置移动到中央位置的位置;

随后处理所述第一音频项目(104)的转换部分,使得所述第一音频项目(104)的转换部分的声音被减弱。

14. 根据权利要求1的设备(100),其中所述操纵单元(103)适用于以在所述转换部分内逐渐修改所述音频数据(101,102)的与时间有关的音频属性的方式操纵所述转换部分。

15. 根据权利要求1的设备(100),其中所述操纵单元(103)适用于以修改所述音频数据(101,102)的与时间有关的音频属性的方式操纵所述转换部分,以在所述转换部分中生成与声学多普勒效应一致的可听体验。

16. 根据权利要求1的设备(100),其中所述操纵单元(103)适用于以实现所述第一音频项目(104)的转换部分和中央部分间的平滑连接的方式操纵所述转换部分。

17. 根据权利要求1的设备(100),其中所述操纵单元(103)适用

于以附加地在所述第一音频项目(104)的转换部分中选择性地修改所述音频数据(101, 102)的响度的方式操纵所述第一音频项目(104)的转换部分。

18. 根据权利要求1的设备(100), 其中所述操纵单元(103)适用于以在所述第一音频项目(104)的转换部分中选择性地修改所述音频数据(101, 102)的时间延迟音频属性的方式操纵所述第一音频项目(104)的转换部分。

19. 根据权利要求1的设备(100), 包括适用于再现经处理的音频数据(112, 113)的音频再现单元(108, 109), 特别地包括由头戴式耳机、耳塞和扬声器组成的组中的一个。

20. 根据权利要求1的设备(100), 其中所述第一音频项目(104)包括由音乐项目、语音项目和视听项目组成的组中的至少一个。

21. 根据权利要求1的设备(100), 适用于由自动唱片选播系统、用于在播放列表中搜索音频项目的系统、广播频道切换系统、公共因特网网页切换系统、电话信道切换系统、音频项目重放开始系统和音频项目重放停止系统组成的组中的至少一个。

22. 根据权利要求1的设备(100), 实现为由音频环绕系统、移动电话、头戴耳机、头戴式耳机重放装置、扬声器重放装置、助听器、电视设备、录像机、监视器、游戏设备、膝上型计算机、音频播放器、DVD 播放器、CD 播放器、基于硬盘的媒体播放器、无线电设备、互联网无线电设备、公共娱乐设备、MP3 播放器、hi-fi 系统、车载娱乐设备、汽车娱乐设备、医疗通信系统、体佩式设备、语音通信设备、家庭影院系统、家庭剧场系统、平板电视装置、气氛创建设备、亚低音扬声器和音乐厅系统组成的组中的至少一个。

23. 处理音频数据(101, 102)的方法, 其中该方法包括以在所述音频数据(101, 102)的第一音频项目(104)的转换部分中选择性地修改所述音频数据(101, 102)的第一音频项目(104)的与时间有关的音频属性的方式操纵所述音频数据(101, 102)的第一音频项目(104)的转换部分。

24. 一种计算机可读介质, 其中存储了处理音频数据(101, 102)的计算机程序, 当被处理器(103)执行时, 所述计算机程序适合于完成或控制根据权利要求23的方法。

25. 一种处理音频数据(101, 102)的程序元件, 其中当被处理器

(103) 执行时, 所述程序元件适合于完成或控制根据权利要求 23 的方法。

处理音频数据的设备和方法

技术领域

本发明涉及用于处理音频数据的设备。

除此之外，本发明涉及处理音频数据的方法。

并且，本发明涉及程序单元。

进一步地，本发明涉及计算机可读介质。

背景技术

音频重放（playback）设备变得越来越重要。特别地，越来越多的用户购买基于头戴式耳机的音频播放器和基于扬声器的音频环绕系统。

当通过音频播放器逐个重放不同的音频项目时，期望在两个相继的轨道之间具有明显无缝的转换（transition）。这可被表示为“混音（mix）”。在“交叉渐变（cross-fade）”期间，可能在从一个轨道到另一个轨道的转换阶段期间交叉渐变轨道。在自动化系统中，为了提供轨道间的无缝转换，离去的轨道的放大率（amplification）典型地将以与进来的轨道的放大率增大相同的速率减小。

包括混音和交叉渐变以实现连续歌曲间平滑转换的、允许歌曲能够自动重放的方法是已知的。这样的技术可被表示为自动 DJ。当提供了播放列表时，不可能根据定义播放在播放列表中的所有歌曲，使得在转换期间音频质量的主观感知是合适的。

传统的自动 DJ 系统允许盲目地执行交叉渐变，允许节奏和和声的冲突。这可能给出感知上地令人不愉快（“差劲的 DJ”）的体验。在普通用户定义的播放列表的情况下，不相称转换的出现率（occurrence）甚至比由专业的唱片选播者（disc jockey）制作的播放列表中的更大。

另一种传统系统基于在两个重放项目间留出短暂停顿的规则，使得和声的混音不会发生，并且节奏的连续性被破坏。即，声音被减弱。该方法有效地使两个重放列表项目在时间上分离，并且如果暂停足够长，不会有旋律或和声的不连续的体验。任何自动 DJ 效果在这样的概念中明显不存在。

当听音频播放列表、记录或其他音乐收藏品时，用户通常做的是例如通过分别按播放器上的“下一首”或“上一首”按钮从一个项目向前或向后

跳转到另一个项目。这可以在音频项目的开头和结尾之间的任何地方实施。这在音频播放器中实现的方式是当前项目的声音被减弱而新的轨道开始播放。

从一个音频轨道移动到另一个的更先进的方法是旨在以以下方式混音两个轨道的自动 DJ 系统：类似于舞曲唱片选播者如何将一个项目的结尾结合到另一个的开头来完成从一个轨道移动到另一个。这两个信号可被同步并且信号逐渐地被交叉渐变，以给出从一个项目到另一个平滑转换的印象。

US 2005/0047614 A1 公开了一种用于在诸如环绕环境这样的多声道（multi-channel）音频环境中加强歌曲间转换的系统和方法。在该方法中，通过在转换期间独立操纵（manipulate）每个节目的各个声道的音量，给正在结束的节目运动的错觉（illusion of motion）以创建歌曲正在退出的印象，而给正在开始的节目运动（motion）以创建歌曲正在进入的印象。

然而，根据 US 2005/0047614 A1 的两段音频之间的转换因为移动被以过分简单化的方法模拟的缘故对于人类收听者来说听上去仍然可能显得不自然。

发明内容

本发明的一个目的是提供一种允许音频项目的开头或结尾处的适当音频体验的音频系统。

为了达到上述目标，提供了根据独立权利要求的用于处理音频数据的设备、处理音频数据的方法、程序单元和计算机可读介质。有益的实施例在从属权利要求中限定。

根据本发明的示例性实施例，提供了一种用于处理音频数据的设备，其中该设备包括操纵单元（特别地，重采样单元），适用于以修改音频数据第一音频项目的转换部分的与时间有关的音频属性的方式选择性地操纵（特别地，重采样）音频数据第一音频项目的转换部分（特别地，也可以以逼真的方式模拟移动的时间延迟效果）。

根据本发明的另一个示例性实施例，提供了一种处理音频数据的方法，其中该方法包括以修改音频数据的第一音频项目的转换部分的与时间有关的音频属性的方式选择性地操纵音频数据的第一音频项目的转换部分。

根据本发明的又一个示例性实施例，提供了一种程序单元（例如源代

码或可执行代码形式的软件程序), 当被处理器执行时, 其适用于控制或实施具有上述特征的数据处理方法。

根据本发明的再一个示例性实施例, 提供了一种计算机可读介质(例如 CD、DVD、USB 棒、软盘或硬盘), 在其中存储了计算机程序, 当它被处理器执行时, 适用于控制或完成具有上述特征的数据处理方法。

根据本发明实施例被实施的用于音频节奏操纵(tempo manipulation)和/或频率改变(frequency alteration)目的的数据处理可通过计算机程序即通过软件实现, 或者通过使用一个或多个特殊的电子优化电路即以硬件实现, 或者以混合的形式即借助于软件部件和硬件部件来实现。

在本申请的上下文中, 术语“操纵”可特别地表示重新计算音频数据流或音频数据片段(audio data piece)特定部分以选择性地修改这个部分的时间或频率相关的属性, 即对关于声音表现的节奏和音高的可听体验有影响的参数。因此, 诸如节奏和/或音高这样的属性可通过这样的操纵得以修改, 特别地以获得多普勒效应。因此, 操纵或重采样可通过重新计算具有与原始记录的文件中的属性不同的属性的声音文件中的样本来进行。这可包括以某种方式移除样本、修改可获得的频率范围、引入暂停、增加或减少音调的再现次数等以改善音频片段间转换的感知。特别地, 因为允许结尾和开头轨道的感知解耦的音高转换效果可避免相继的音频片段间的节奏和和声冲突。

音频项目的术语“转换部分”可特别地表示该音频项目的开头部分和/或结尾部分, 在所述部分, 转换发生在所述音频项目与另一个(在前的或随后的)音频项目之间或者该音频项目与无声时间间隔之间。

术语“与时间有关的音频属性”可特别地表示可以以特定方式调整时间特性和相应的音频参数, 例如以强调淡入(fading in)或淡出(fading out)音频片段的印象。这可以包括被称为所谓的声学多普勒效应的、用于指示音频项目淡入或淡出的直观测量的频率变化。

根据本发明的一个示例性实施例, 音频片段的转换部分可选择地被处理以改善对于人耳来说的该音频项目和在前或随后的音频信息之间转换的感知。通过在淡入和/或淡出期间改变与时间有关的音频重放属性, 可以生成接近或离开声源的印象, 这在心理上可分别被关联为新歌曲的开始或当前被重放歌曲的结束。

因此, 根据示例性实施例, 可以使用于自动地选播音乐(DJing)的动态混音(mixing)成为可能。在自动唱片选播系统中, 可以使歌曲转换

不出现令人烦人的中断。这一般可以通过交叉渐变两首连续的歌曲来完成。为了获得平滑的转换，要求歌曲的节奏和旋律在混音区域被调准，并且这些歌曲具有在混音区域中匹配的和声属性。这通常在可在另一首之后播放的歌曲上施加约束。根据一个示例性实施例，调准节奏、旋律和和声的需要通过在转换期间对每首歌曲施加采样频率的不同的滑动的变化而得以克服。滑动的采样频率可创建被混音的两首歌曲的自然解耦，使得节奏、旋律和和声冲突不重要。因此，本发明的实施例可克服不是每个播放列表（或者每对歌曲）都能采用自动 DJ 方法来交叉渐变这一限制。本发明实施例所依据的认识是除了通过暂停的时间上的分离之外，还有其他可能的方法使两个播放列表项目感知上分离。出于这个目的，可以使用一个或两个音频信号的频谱的动态系统的操纵。特别地，可以执行一种方法，其中在歌曲的混音区域中，执行歌曲的操纵/重采样，使得一首歌曲具有滑动下降的频率和节奏，而另一首歌曲具有滑动上升的节奏和频率。因此，自动 DJ 应用和强制转换中音频项目的时间上的操纵可得以使用，并且可基于这样的考虑：可引起导致频率滑移效果的足够强大的多普勒频移效应。因此，使得自动 DJ 应用的动态混音成为可能。可使在自动 DJ 系统中被混音的两首歌曲的自然解耦成为可能，使得这些歌曲不需要在节奏、旋律、和声内容等方面相似。这可以通过在转换期间操纵这两首歌曲而得以创建，使得正在结束的歌曲的节奏和/或频率从原始频率滑动地下降到较低的频率，而正在开始的歌曲的节奏和/或频率平滑地向具有不同频率轮廓（frequency contour）的原始频率下降。这也能作为空间转换效果的副产品而得以实现。可以创建两首歌曲的虚拟源的移动错觉，并且可以生成多普勒效应。取决于创建源移动错觉的方法，这常常还可以产生多普勒效应，即多普勒效应是移动效果的结果。

接下来，将解释用于处理音频数据的设备的另一个示例性实施例。然而，这些实施例也应用于处理音频数据的方法、程序元件和计算机可读介质。

第一音频项目的转换部分可以是第一音频项目的结尾部分。换句话说，通过以渐进或逐步的方式调节时间属性，可以执行操纵以平滑地淡出第一音频项目的结尾。

附加地或可选地，第一音频项目的转换部分可以是第一音频项目的开头部分。换句话说，通过以渐进或逐步的方式调节时间属性，可以执行操纵以淡入第一音频项目的开头。因此，可以仅操纵音频项目的开头部分、

仅处理音频项目的结尾部分或者处理音频项目的开头部分和结尾部分这两者。也可以以这样的方式处理音频项目的中间部分，例如，用户可以在第一首歌曲的中间停止重放，并从第二首歌曲的开头或从其中间的任何地方开始播放第二首歌曲。换句话说，音频项目的自然开头或自然结尾可以或不可以与转换部分一致/结合在一起。因此，根据本发明示例性实施例的选择性的时间操纵也可以在歌曲的中间进行。

特别地，操纵单元可适用于以由第一音频项目受操纵的结尾部分的节奏和频率组成的组中的至少一个滑出（glide out）的方式操纵第一音频项目的结尾部分。因此，通过考虑当重放这样的音频内容时对音频感知有影响的这样的与时间有关的音频参数，可以获得声学多普勒效应的印象，正如从离开的救护车喇叭可知，它不仅幅度下降，而且频率也下降（应当注意，离开的救护车喇叭声音的频率比接近的救护车的声音低，但是频率没有减小（滑动），除非救护车正相对于观察者加速或减速）。特别地，当操纵淡出音频项目的结尾部分时，节奏和/或频率可被减小。

虽然本发明的实施例可以专注于提供接连再现的音频项目间的平滑转换，但可以处理确切地仅一个音频项目，例如将在结尾部分轻柔地减弱的音频项目。

然而，操纵单元也可以适用于以修改第二音频项目（可能接在第一音频项目之后）的转换部分的与时间有关的音频属性的方式操纵第二音频项目（可能接在第一音频项目之后）的转换部分。因此，可以通过考虑两个转换部分之间的与时间有关的音频属性来使第一音频项目和第二音频项目之间的转换平滑。在转换部分（一个或多个）期间，第一音频项目和第二音频项目这两者可以都被同时重放，但是具有不同的音频参数。

特别地，第二音频项目的转换部分可以是第二音频项目的开头部分。则操纵单元可适用于以使由第二音频项目受操纵的开头部分的节奏和频率组成的组中的至少一个滑入/淡入的方式操纵第二音频项目的开头部分。为了这样的淡入效果，（以渐进或逐步的方式）增大节奏和频率直到第二音频项目的转换部分已被完成可以是合适的。

操纵单元可适用于选择性地操纵第一音频项目的仅一个转换部分（开头部分或结尾部分）或多个转换部分（开头部分和结尾部分），而第一音频项目的剩余（中央）部分将保持不被采样，即是不改变。因此，在平滑地淡入随后将被重放的音频信号后，原始数据将被重放，使得在转换区（regime）完成后不发生音频失真（audio artefact）。

操纵单元可适用于以协同方式操纵第一音频项目的转换部分和第二音频项目的转换部分。因此，淡出项目的节奏和频率的降低（引起离开音频源的多普勒效应）可以以协调的方式与其中节奏和频率被增大（接近的音频源的多普勒效应）的随后音频信号的淡入相结合。这可以允许甚至在起源非常不同的音频内容之间的听觉上合适的转换部分，使得将被混音的两首歌曲不必在节奏、旋律或和声冲突方面彼此对应。

操纵单元还可用作运动体验（motion experience）生成单元，其适用于以生成再现第一音频项目的音频源在转换期间正在移动的可听体验的方式处理第一音频项目。然而，这样的移动音频源印象不必受限于音频项目响度的简单变化（对于接近的对象增大响度，对于离开的对象减小响度），但是可以通过考虑产生与音频源的实际运动有关的交叉声道时延的时间修改来进一步改进这样的运动感知。特别地，声学多普勒效应不仅修改离开或接近的声音源的响度，还修改频率、节奏和其他与时间有关的音频参数。通过考虑这样的与时间有关的属性，与简单的响度调节系统相比，被重放的音频数据的移动将被感知为显著地更加自然，或者更明确地更加接近移动的声音源的感知。

这样的运动体验生成单元可适用于生成再现第一音频项目的音频源在第一音频项目的结尾部分期间正在离开的可听体验。因此，相应音频项目部分的操纵可以以模拟离开的声音源的声学多普勒效应的方式来进行。

运动体验生成单元可进一步适用于以生成再现第二音频项目的音频源在转换部分期间正在移动（特别地，在第二音频数据的开头部分期间正在接近）的可听体验的方式处理第二音频项目。换句话说，在这样的实施例中，第二音频项目的开头部分的处理可以以人耳能够感知到接近的音频源的声学多普勒效应的印象的方式来进行。

从心理学视角看来，淡出与离开的声音源有关以及淡入与接近的声音源有关是非常直观的。

运动体验生成单元可适用于根据下面的测量生成第一音频项目结尾部分和第二音频项目开头部分之间的转换。首先，可处理第二音频项目转换部分的第一部分，使得第二音频项目转换部分的再现可感知为源于遥远的开始位置。换句话说，第二音频项目被打开并且将被感知为来自位于远处的声音源，这可以通过小音量和相应的方向属性来模拟。随后，第一音频项目转换部分的第一部分可以以第一音频项目转换部分的再现可感知为源于从中央位置移动到遥远的最终位置的位置的方式来处理。换句话

说,在第一音频项目的中央部分的重放期间,该音频数据将以人类听众具有发出第一音频项目的声音源位于中央位置的印象的方式来配置。为了指示第一音频项目随后将淡出,可以虚拟地(virtually)将在转换部分的第一部分中发出第一音频项目的声音源从该中央位置移动到遥远的最终位置。该运动可以逐渐地来进行。同时,随着发出第一音频项目的虚拟声音源的离开,第二音频项目的第二转换部分可以以第二音频项目的转换部分的第二部分的再现可感知为源于正从遥远的开始位置(例如逐渐地)移动到中央位置的位置(与发出第一音频项目的(虚拟)声音源预先所处位置相同的位置,或者另一个位置)的方式来处理。因此,由于第二音频项目将被淡入,人类听众将获得发出指示第二音频项目的声波的虚拟音频源正在接近第二音频项目的主要部分将被再现的位置的印象。随后,处理第一音频项目的转换部分的第三部分,使得第一音频项目的转换部分的声音被减弱。因此,在第二音频项目已(虚拟地)接近最终或中间的位置后,可以(渐进地或以逐步的方式)降低第一音频项目的音量,从而完成淡出过程。可选地,然后发出第二音频项目的主要部分的虚拟声音源可被再次重新定位,或者可维持在中央位置。

“中央位置”可指头戴式耳机信号如何在音频的“中央部分”期间从原始音频信号生成的方式。例如,当没有进行转换时,左信号未经处理地到达左耳,而右信号未经处理地到达右耳。在音频轨道的“中央部分”,可使用被表示为“中央位置(重现/再现/)”的处理模型。在中央位置,表现(立体声信号的)原始左、右音频声道的信号可典型地被直接发送至左、右头戴式耳机,或者对信号施加某个与转换期间的处理无关的处理。这种类型的附加处理可与频谱均衡、空间扩展(spatial widening)、动态压缩、多声道-立体声变换(在原始音频数据具有非立体声格式的情况下)、或者在音频轨道的中央部分期间独立于在转换部分期间所使用的转换方法施加的其他类型的音频处理效果和增强相关。

设备可包括适用于再现经处理的音频数据的音频再现单元。这样的(物理的或真实的)音频再现单元可以是例如头戴式耳机、耳机或扬声器,其被提供用于重放的经处理的音频数据。音频数据可以以收听重放的音频数据的用户得到(虚拟的)音频重放单元被置于另一个位置的印象的方式来处理。

第一音频项目可以是音乐项目(例如音乐剪辑或CD上的音乐轨道)、语音项目(例如电话对话的一部分)或者可以是视频/视听项目(例如音

乐视频、电影等)。因此,本发明的实施例可被实施于所有其中必须处理音频数据的领域中,特别是实施于其中将以平滑方式把两个音频项目相互连接的领域中。

本发明示例性实施例的示例性应用领域为自动唱片选播系统、用于在播放列表中搜索音频项目的系统、广播频道切换系统、公共因特网页面切换系统、电话信道切换系统、音频项目重放开始系统和音频项目重放停止系统。用于在播放列表中搜索音频项目的系统可允许针对特定的音频项目搜索或扫描播放列表,并随后重放这样的音频项目。在两个相继的这样的音频项目之间的转换部分,可实施本发明的实施例。另外,当在不同电视或无线电频道之间切换时,即在广播频道切换系统中切换时,可根据本发明的示例性实施例来进行在前频道的淡出和随后频道的淡入。当操作计算机的用户在不同因特网页面间切换从而使用公共因特网页面切换系统时,也可以根据本发明的示例性实施例来进行相同的淡出、淡入。在电话对话期间,当进行不同信道或通信方之间的切换时,可以针对这样的电话信道切换系统实施本发明的实施例。同样针对简单地开始或停止音频重放,即针对无声的重放模式和高声的重放模式之间的切换,可以实施本发明的实施例。

本发明的实施例可与附加的可能性相结合以使用空间转换结果来创建两首歌曲之间空间分离的错觉。“交叉渐变”的两首歌曲可具有不同的移动轨迹,使得现存的源(第一首歌曲)向例如左侧离开,而新的歌曲(第二源)从右边移动进声像(sound image)。

在使两个项目分离中使用上升和声模式(harmonic pattern)和下降和声模式可以也具有来自实验心理学的有力支持,其中观察到两个音调综合体的不同频率调制轨迹引起这两个音调综合体分离为两个不同的感知流(例如参见 A.S. Bregman (1990), “Auditory Scheme Analysis: The Perceptual Organization of Sound”, Cambridge, MA: Bradford Books, MIT Press)。

操纵与时间有关的音频参数的效果是歌曲在混音区域中被在感知上解耦,使得它们不再被感知为不和谐的。因此,使用该方法,对于确保节奏、旋律或和声匹配,只需给予较低的特别关注。这允许任何任意歌曲对的混音,进而允许需要通过根据本发明示例性实施例的自动 DJ 方法重放的任何播放列表。

本发明的示例性实施例可以应用于通过混音两首连续歌曲的开头和

结尾来创建歌曲转换以获得平滑转换的应用中，例如应用于自动 DJ 应用中。

根据本发明的另一个示例性实施例，可使转换效果和普通收听之间的空间转换成为可能。空间转换效果可被用于音频项目之间的强制转换。所述转换效果基于典型地在基于模型的重现场景中的音频流的动态专门化（dynamic specialisation）。不期望在普通头戴式耳机收听中运行基于模型的空间处理，因此可以定义转换，用于普通收听到转换重现以及返回。

因此，可使用音频信号的空间操纵来完成从一个轨道到另一个轨道的移动。目标可以是给出一个轨道物理地离开而另一个轨道进入的感知。例如，以这样的方式：当前音乐轨道文件远离地飞到右手侧而另一个轨道从左手侧滑进。当在音频播放器列表的情境下完成了该操作时，它给出了对播放列表的非常强烈的空间印象。在空间坐标中的这种类型的音频播放列表项目的表现将在音频技术中提供新的应用。

在头戴式耳机收听中，清楚地定义了什么是左、什么是右。一个明显的解决方案是使用例如标准的幅度平移规则（amplitude panning rule），以立体声图像（stereo image）逐渐衰减并仅移动到右耳信号、并同时增大始于左耳的另一个轨道的音量的方式改变平衡的立体声图像。然而，以这种方式获得的转换效果既不是非常令人感兴趣，又没有给出轨道改变的非常强烈的空间印象。一个问题可能是，立体声音频记录的两个声道可以包含取决于记录的生产的类型非常不同的听觉线索。

通常，立体声音频项目的两个声道是关联的。然而，例如在幅度平移或立体声混响中创建的这个关联并不具有与任何可识别的空间属性（例如音频源的距离，或者例如个体音乐乐器的声音到达的明确角度）的直接关系。因此，产生令人信服的空间音频轨道变化的挑战在于，因为音频轨道在第一地方（place）中没有空间位置，则将它远远地扔到右边的某处是不合适的。使用基于虚拟扬声器收听者系统的重现场景可能遇到这样的挑战。然而，可以考虑普通收听场景（在头戴式耳机中，或立体声或多声道扬声器再现场景）和轨道转换效果之间的转换。

接下来，将解释与音频项目之间的空间转换相关的实施例。可以提供一种用于在头戴式耳机收听中实施从一个音频流到另一个的强制转换中的直观空间音频效果的方法。例如当用户在全面检查播放列表或浏览无线电频道列表中按“下一首”或“上一首”按钮时，所建议的效果向收听体验提供新的空间维度。该方法基于将立体声信号映射至虚拟扬声器收听者模

型，其中能够使空间转换直观和清楚。

可提供一种使用音频信号的空间操纵来从一个轨道移动到另一个轨道的方法，以给出一个轨道物理地离开而另一个进入的感知。例如，以这样的方式：当前的音乐轨道朝第一方向离去，而另一个轨道从可与第一方向相反的第二方向滑进。当在音频播放列表的情境中执行该方法时，它给出了非常强烈的播放列表空间印象。例如，用户可记得第一首歌曲就在第二首歌曲的左手侧，而另一首歌曲很远地在右边的某处。自然地，该场景可直接延伸到诸如北、东、南和西这样的方向以给用户音频材料的两维的表现。因此，可使一维、两维或者甚至三维的空间效果成为可能。因此，可以将立体声音频材料的两个音频声道定位至其中扬声器和收听者耳朵具有明确几何位置的模拟扬声器收听者场景。一旦执行该操作，可以将虚拟扬声器移动到任意位置以创建期望的空间效果。在从一个音频项目换到另一个中，可执行模拟，使得两个播放第一音频项目的虚拟扬声器从用户耳朵远远地移动到左边，而播放另一个项目的另一对扬声器可从右边被带入到合适的或最优的重放位置。因此，可以提供不同空间音频收听场景的几何特性，并且可使用虚拟声学环境中的声音传播的模拟。

当一个音频项目必须结束，而另一个必须开始时，创建了第一音频项目朝一个方向远离收听者而移动、第二音频项目朝向收听者而移动的听觉图像。可提供一种在强制转换和头戴式耳机收听期间转换音频的方法。该方法可包括通过模拟虚拟扬声器在某个位置开始新项目，将当前项目从头戴式耳机移动到虚拟扬声器配置，将当前项目移动到目标位置，以及同时将新项目的扬声器位置移动到该虚拟扬声器位置，将新项目从扬声器位置移动到头戴式耳机收听，并减弱当前项目的声音。

当在播放列表上预览项目以便项目（虚拟地）传递到收听者前面时，或当暂时减弱一个项目的声音时，也可以使用该方法。

用于处理音频数据的设备可被实现为由音频环绕系统、移动电话、头戴耳机、扬声器、助听器、电视设备、录像机、监视器、游戏设备、膝上型电脑、音频播放器、DVD 播放器、CD 播放器、基于硬盘的媒体播放器、互联网无线电设备、公共娱乐设备、MP3 播放器、hi-fi 系统、车载娱乐设备、汽车娱乐设备、医疗通信系统、体佩式（body-worn）设备、语音通信设备、家庭影院系统、家庭剧场系统、平板电视、气氛创建设备、亚低音扬声器（subwoofer）和音乐厅系统组成的组中的至少一个。其他应用也是可能的。

然而,虽然根据本发明实施例的系统主要意欲改善声音或音频数据的质量,但也可以将该系统应用于音频数据和视觉数据的结合。例如,本发明的实施例可被实现在视听应用中,视听应用例如视频播放器或家庭影院系统,其中发生不同视听项目(例如音乐剪辑或视频序列)间的转换。

根据下文将要描述的实施例的实例,本发明的上述方面和其他方面将变得清楚明白,并且将参考这些实施例的实例进行阐述。

附图说明

下面将参考实施例的实例更详细地描述本发明,本发明不限于实施例的实例图1图解了根据本发明示例性实施例的音频数据处理设备。

图2至图5图解了转换至和自转换模型,该模型通过基于本发明示例性实施例的转换模型的声音重现的参数操纵来实现。

图6图解了作为扬声器收听者模型的特殊情形的通用头戴式耳机的几何描述。

图7图解了以两声道扬声器收听配置对收听者的模拟。

图8示出了表现从虚拟麦克风对移开的一个音频轨道的扬声器对,以及播放另一个轨道的新的扬声器对被移动到收听位置。

图9图解了根据本发明示例性实施例的立体声扬声器收听中的轨道转换。

具体实施方式

附图中的图解是示意性的。在不同的附图中,为类似或相同的元件提供相同的附图标记。

在下文中,参考图1,将说明根据本发明一个示例性实施例的用于处理音频数据101、102的设备100。

图1中所示的设备100包括诸如CD、硬盘等的音频数据源107。在音频数据源107上存储了多个音乐轨道,例如第一音频项目104、第二音频项目105和第三音频项目106(例如三个音乐片段)。

一旦接收到相应的控制信号,音频数据101、102(例如用于扬声器左侧和右侧的数据)可从音频数据源107发送到诸如微处理器或中央处理单元(CPU)这样的控制单元103。

控制单元103与用户接口单元114双向通信,并且能够与用户接口单元114交换信号115。用户接口单元114包括诸如LCD显示器或等离子体

设备这样的显示元件，并且包括诸如按钮、小键盘、操纵杆或者甚至语音识别系统的麦克风这样的输入元件。人类用户能够控制控制单元 103 的操作，并且因此可以调节设备 100 的用户偏好。例如，人类用户可以完全地对播放列表的项目进行切换。并且，控制单元 103 能够输出相应的重放或经处理的信息。

在以下文将更详细地描述的方式处理音频数据 101、102 之后，第一经处理的音频数据 112 被应用于第一扬声器 108 以供重放，由此生成声波 110，获得第二经处理的音频数据 113，音频数据 113 可由连接的第二扬声器 109 再现，第二扬声器 109 能够生成声波 111。

在第一音频项目 104 将被再现并且随后第二音频项目 105 将被再现的场景中，期望在在前的第一音频项目 104 和随后的第二音频项目 105 之间具有平滑或无缝的转换部分。出于这个目的，控制单元 103 可作为操纵单元，用于以修改第一音频项目 104 和第二音频项目 105 之间的转换部分的与时间有关的音频属性的方式操纵第一音频项目 104 和第二音频项目 105 之间的转换部分。更特别地，可以处理第一音频项目 104 的结尾部分和第二音频项目 105 的起始部分或开头部分。因此，可获得第一音频项目 104 滑出或淡出、第二音频项目 105 滑入或淡入的可听感知。出于这个目的，第一和第二音频项目 104、105 的时间属性仅在转换部分可被调整，而第一和第二音频项目 104、105 的中央部分可不加修改地被重放。这可包括修改音频数据 101、102 的频率和节奏值，使得滑出的第一音频项目 104 将根据声学多普勒效应得以操纵，从而人类收听者对经操纵的第一音频项目 104 的感知是音量和频率/节奏两者在结尾部分都被减小。

因此，根据声学多普勒效应操纵第二音频项目 105 的起始部分，使得第二音频项目 105 的开头部分的感知的可听效果是增大的响度和增大的频率/节奏。通过采取这种措施，可获得非常直观的淡入特性。

可同时或以重叠的方式重放第一音频项目 104 的经操纵的结尾部分和第二音频项目 105 的经操纵的开头部分。

协调或调整第一音频项目 104 的结尾部分和第二音频项目 105 的开头部分的时间特性的变化，以获得合适的声音。

特别地，控制单元 103 还可以生成这样的感知：在重放第一音频项目 104 的结尾部分的期间根据第一音频项目 104 的结尾部分发出声波的虚拟音频源离开了。更特别地，这样的运动实验（experiment）生成特征可生成重放第二音频项目 105 的开头部分的虚拟重放设备接近人类收听者的

可听感知。

图 1 的系统可被用作自动 DJ 系统。

本发明的实施例基于这样的领悟 (insight): 任何空间转换效果都隐含地或明显地基于扬声器 - 收听者系统的模型。该模型可被用来控制由音频作品的原始音频信号的数字滤波实现的动态重现操作。在普通收听场景中, 音频信号可通过再现系统的扬声器直接重放。根据一个示例性实施例, 扬声器系统可以是从小体声耳机到诸如 5.1 环绕音频系统或波场合成系统 (wave field synthesis system) 这样的多声道扬声器系统范围内的任何配置。

根据一个示例性实施例, 提供了一种用于从普通收听到在空间轨道转换效果中使用的重现模型的转换以及反转换回普通收听模式的通用方法。在这样的实施例中, 可能的是, 普通收听场景通常可以识别为在空间转换效果中使用的重现模型的特殊情形。因此, 到和从转换模式的转换可以通过基于转换模型的声音重现的参数操纵来进行。这在图 2 至图 5 中得以解释并且将在下文中更详细地描述。

图 2 示出了方案 200。

方案 200 示出了音频作品 201, 它在普通收听 202 中的音频再现路径中被重放。音频再现系统用附图标记 203 表示, 可被实现为头戴式耳机、立体声系统或 5.1 系统。

另外, 虚拟扬声器 - 收听者模型用附图标记 204 表示, 并且包括代表普通收听的模型的特殊情形 205、转换效果的音频再现路径 206 和转换效果的另一个音频再现路径 207。

图 3 显示了方案 300。在方案 300 中, 还显示了第二音频作品 301。

从图 3 可以得出, 在转换开始时, 第一音频作品 201 通过代表转换模型的普通收听的模型的特殊情形 205 而被发送。从代表普通收听的模型的特殊情形 205 到转换效果的音频再现路径 206 的转换开始, 并且它基于虚拟扬声器 - 收听者模型 204 的参数的参数操纵。第二音频作品 301 的动态转换重现可在这个阶段通过转换效果的另一个音频再现路径 207 开始。

图 4 随后示出了方案 400。

在连续的转换中, 采用虚拟扬声器 - 收听者模型 204 重现第一音频作品 201 和第二音频作品 301 这两者以实现期望的动态空间转换效果。典型地, 以第一音频作品 201 似乎远离收听者而第二音频作品 301 正在接近收听者的方式再现第一音频作品 201。

在图 5 中示出随后的方案 500。

参考图 5，以其结束代表普通收听场景的等同模式的方式修改第二音频作品 301 的动态重现。换句话说，第二音频作品 301 从转换效果的音频再现路径 207 被移动到代表普通收听的模式的特殊情形 205。最终，对于第二音频作品 301，再现被从虚拟扬声器收听者重现场景的特殊模式转换到图 2 的普通音频再现场景。

根据本发明的一个示例性实施例，可以使用一种模型，在该模型中，使用虚拟的麦克风来捕获从虚拟扬声器播放的信号 $x(n)$ ，使得所捕获的信号由下式给出：

$$y(n)=x(n)*\delta(dT)/d^2 \text{ 和 } T=F/c ,$$

其中星号表示卷积， d 是虚拟扬声器和麦克风之间以米为单位的距离，其中 F 是采样频率， c 是声音的速度。实际上，对应于分数时间指数 (fractional time index) dT 的信号值可使用诸如拉格朗日插值滤波器 (Lagrange interpolator filter) 这样的分数延迟滤波器来实现。

图 6 示出了涉及作为扬声器 - 收听者模型的特殊情形的通用头戴式耳机收听的几何描述的排列 610。

图 6 显示了用于再现音频内容的头戴式耳机 600。还示出了左虚拟扬声器 601 和右虚拟扬声器 602。进一步地，示出了左虚拟麦克风 603 和右虚拟麦克风 604。无穷大的距离由附图标记 605 表示。

基于前面的讨论，立体声声道之间的串音 (crosstalk) 或关联 (correlation) 被看作是同时发生的，使得几何声学意义上信号间的关联不被建模为从一个音频声道到另一个音频声道的泄漏。

在本发明的一个实施例中的普通收听模式是头戴式耳机收听。图 6 中图示了作为提出的扬声器 - 收听者模型的特殊情形的根据排列 610 的这样的通用头戴式音频收听场景的几何描述。声音从左、右虚拟扬声器 601、602 播放，理论上左、右虚拟扬声器 601、602 彼此相离无穷远地放置。声音被放置在左、右虚拟扬声器 601、602 附近的左、右虚拟麦克风 603、604 捕获。捕获的信号然后通过头戴式耳机 600 向用户重放。来自原始左、右声道的立体声记录的合成在头戴式耳机收听中准确地产生了原始信号。该几何描述的无穷大的距离仅是建立没有两个信号间串音的模型的一个实施例，可以通过给予麦克风 (或扬声器，或两者) 减少或消除串音的方向性属性 (directivity property) 来获得相似的结果。

根据一个示例性实施例，仅考虑了自由场（free field）中的全方向虚拟扬声器和麦克风。然而，本发明的实施例还包括方向性和声场模拟的使用。技术人员已知将更多的实际方向性属性和房间模型包括在声学模型中所需的措施。实际上，不需要或不可能在即使具有全方向换能器（transducer）的源之间具有无穷大的距离。自由声场条件下和对于全方向源的声音（单位为分贝）的衰减由下式给出：

$$L_R = 20 \log_{10}(R)$$

例如，20 米的间隔已经给出了 26dB 的串音衰减，这对在典型的立体声音频材料中的空间图像具有可忽略的影响。这样的表现在感知上类似于原始立体声再现，并且也不立即提供直观的特殊轨道转换方法。然而，可以做出另一个变换，该变换将左、右虚拟扬声器 601、602 和左、右虚拟麦克风 603、604 的位置移动至图 7 中图示的另一个设置 700，图 7 附加地显示了人类收听者的头 701。

在图 7 中，左、右虚拟扬声器 601、602 被移动至典型的扬声器收听中左、右扬声器的位置。左、右虚拟麦克风 603、604 被移动至典型的收听情形下表示收听者耳朵位置的位置。

因此，图 7 示出了对在两声道的扬声器收听系统中收听者头 701 的模拟。

在从图 6 场景到图 7 场景的转换中左虚拟扬声器 601 和左虚拟麦克风 603 之间的距离保持不变。因此，立体声音频再现的整个音量保持近似相同。然而，这对于当前的实施例不绝对必要。

图 8 示意性地显示了方案 800，其包括将被重放的音频数据的第一音频项目 104 和第二音频项目 105。

表现第一音频项目 104 的左、右虚拟扬声器 601、602 对将被从左、右虚拟麦克风 603、604 转移走，与第二音频项目 105 有关的新扬声器 801、802 对被移向收听位置。

在典型的应用中，从一个音频项目 A 跳到音频项目 B 可以进行如下步骤。该序列可从其中用户正在收听项目 A 的情形开始。

1. 将项目 B 的扬声器装置放置到开始位置。该开始位置可以例如是在右侧远离用户耳朵的位置。
2. 将项目 A 从头戴式耳机收听（图 6）移动到扬声器收听（图 7），并将虚拟扬声器放置于收听位置。

3. 将项目 A 移动到目标位置（例如左边远离用户耳朵的某处），并且同时将项目 B 从开始位置移动到收听位置。
4. 将表现项目 B 的扬声器从扬声器模拟移动到头戴式耳机模拟配置。
5. 减弱项目 A 的声音。

类似的算法也可被用于在播放列表中快速地扫描或搜索音频项目。在这种情况下，音频项目序列从右边流向左边（或者反之亦然）以给予用户播放列表内容的概览（预览），或帮助识别特定项目。在这个特殊的应用中，发出头戴式耳机收听模拟是有用的，使得项目在扬声器重放配置中被重放。该选择提供了传递给收听者的音频项目的平滑流。在这种类型的场景中，播放列表也可以被表示为二维或三维图，在该图上用户自由地在左/右、前/后、上/下或它们的组合的方向上导航。

类似的实施例也可以被直接应用于涉及不同音频流间转换的其他可能应用。例如，可以在改变无线电或 TV 频道、具有背景音频的因特网页面中应用所述类似实施例，可以在个人计算机中从一个音频应用改变到另一个中应用所述类似实施例，等等。

类似的场景也可以被用来为仅涉及一个项目的转换创建新类型的效果。例如，当开始和停止音频项目的重放时，或在暂时减弱音频项目的声音时，可使用空间转换效果。

另外，用于空间转换的相同机制还可以被用于各种不同的电话应用中以在不同通话者（talker）之间切换。

在另一个实施例中，再现系统可以是如图 9 中所示的立体声扬声器系统 900。

图 9 显示了重放第二音频项目 105 的虚拟扬声器 901、902 和重放第二音频项目 105 的虚拟扬声器 903、904。并且显示了左、右附加扬声器 905、906。因此图 9 显示了在立体声扬声器收听中的轨道转换。通过使用本领域技术人员已知的 3D 音频重现技术处理供给左、右附加扬声器 905、906 的音频信号来创建虚拟扬声器 901 至 904。

在图 9 的场景中，通过以使重现的虚拟扬声器的位置和方向属性与真实的扬声器一致的方式移动包含虚拟扬声器 901 至 904 的“泡（bubble）”，获得到其中直接通过左、右附加扬声器 905、906 播放信号的普通音频收听的转换。

就处理而言，对于从通过虚拟扬声器收听者系统的第二音频项目 105 的重放到通过立体声装置的真实左、右附加扬声器 905、906 的重放的转

换，可以给出以下描述。动态重现算法基于输入信号的线性数字滤波，其可由以下不同的方程式描述：

$$y(n)_l = x(n)_l * h(n,t)_{ll} + x(n)_r * h(n,t)_{rl}$$

$$y(n)_r = x(n)_l * h(n,t)_{rl} + x(n)_r * h(n,t)_{rr}$$

其中星号表示卷积，重现滤波器（rendering filter）由冲激响应表示。该重现模型的一个特殊情形是直接的左到左（ll）和右到右（rr）滤波器被减小至单位增益并且串音项（左到右（lr）和右到左（rl））消失。该特殊的情况与使用扬声器的普通收听相同。在动态重现中，转换因此能通过使用实施系数从原始重现滤波器到代表特殊情形的函数的平滑演变的动态转换路径从任何空间重现场景得以实现。

应当注意术语“包括”不排除其他元件或特征，“一”或“一个”不排除多个。同样，联合不同实施例而描述的元件可被组合。

还应当注意在权利要求中的附图标记将不被解释为限制权利要求的范围。

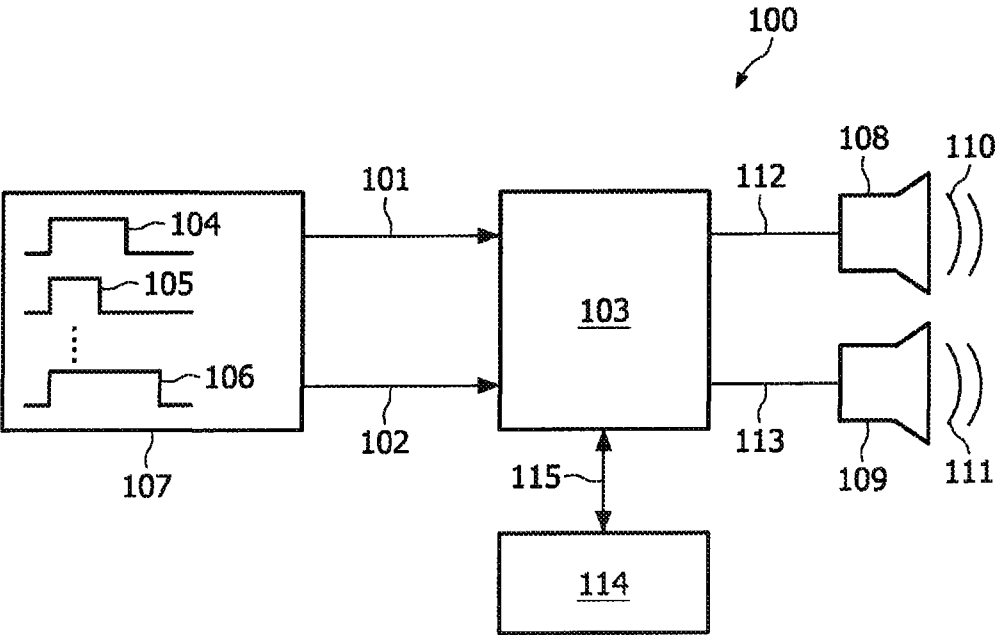


图 1

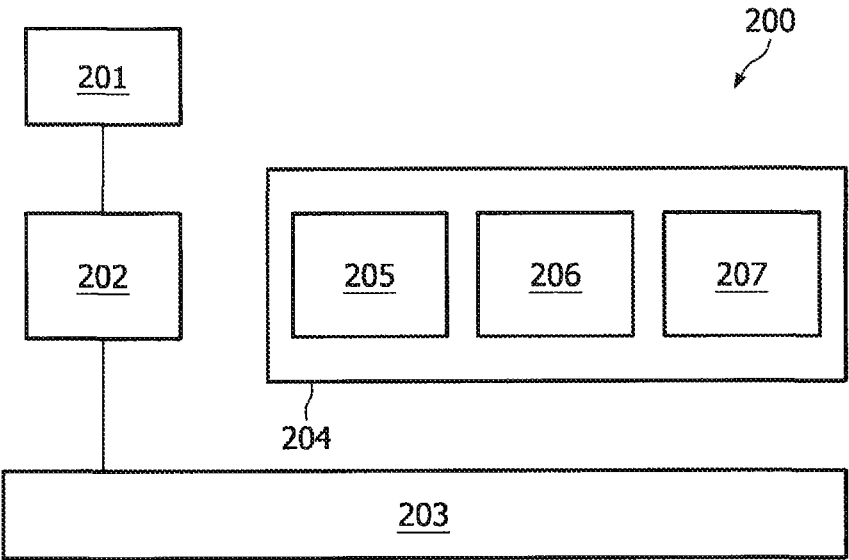


图 2

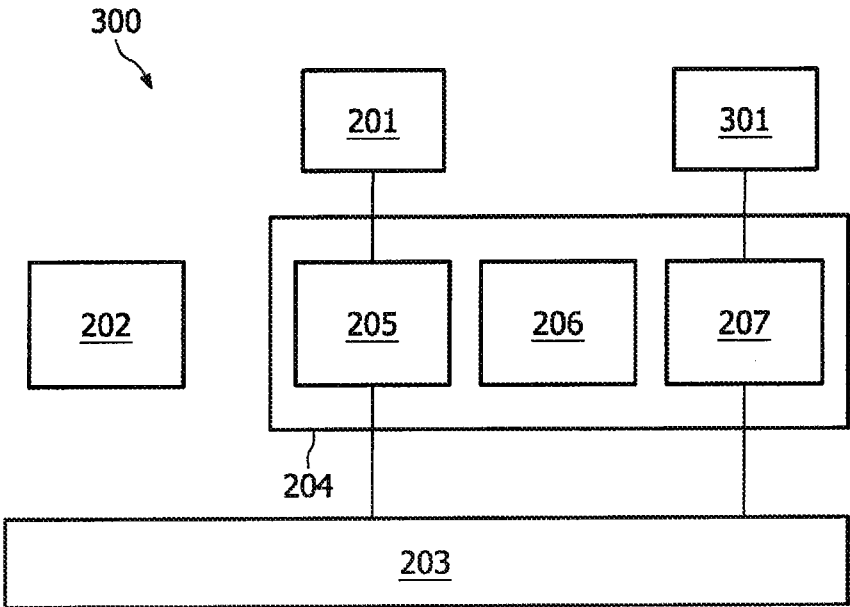


图 3

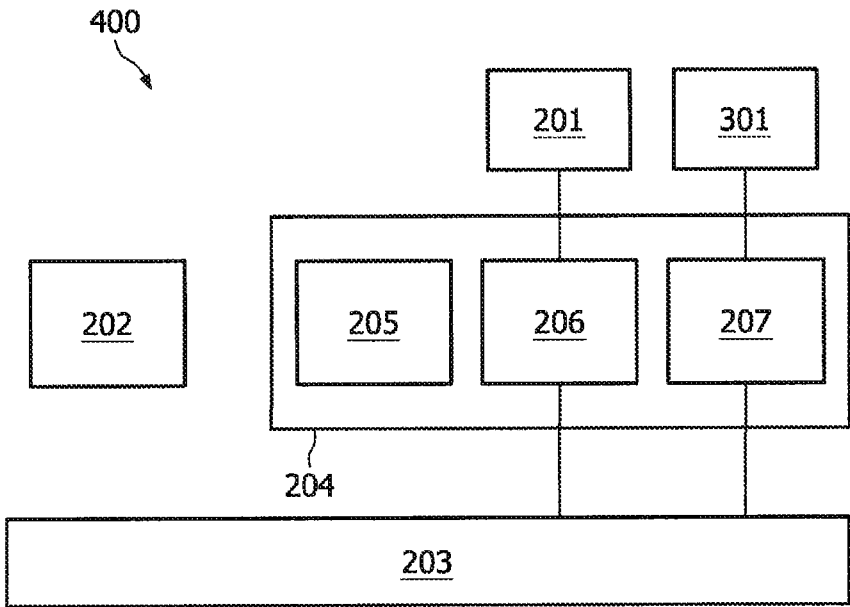


图 4

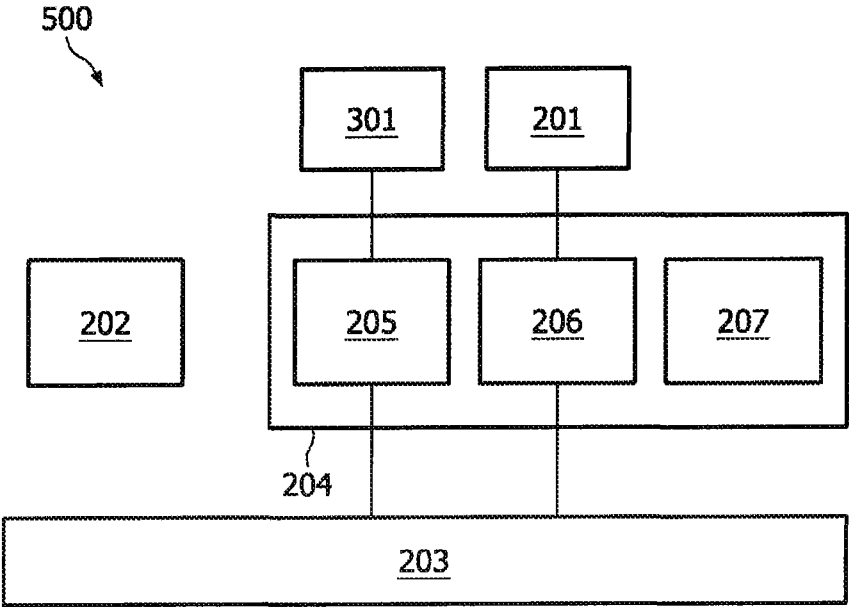


图 5

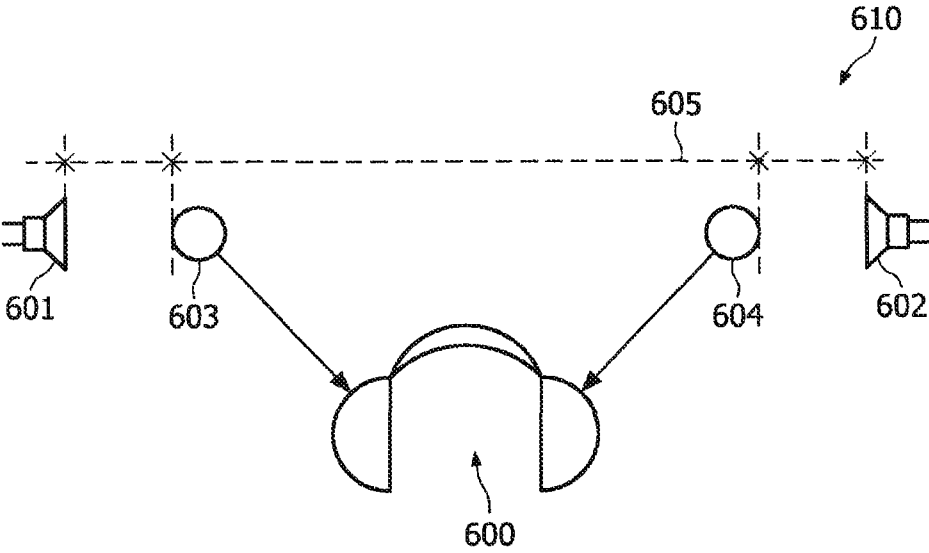


图 6

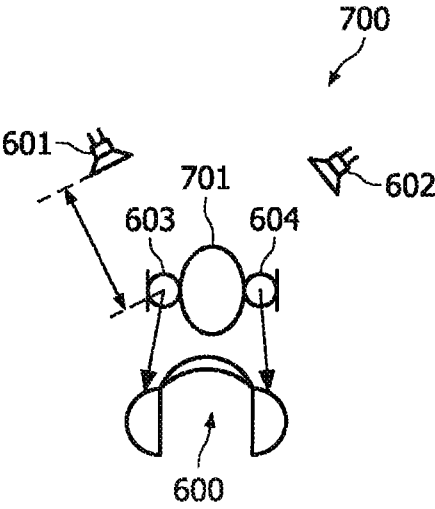


图 7

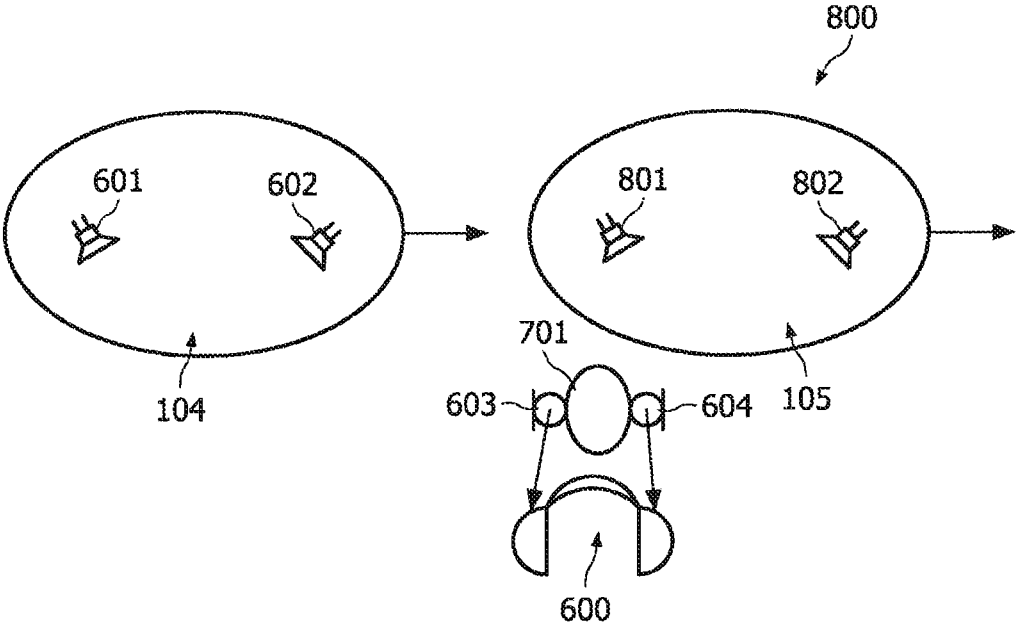


图 8

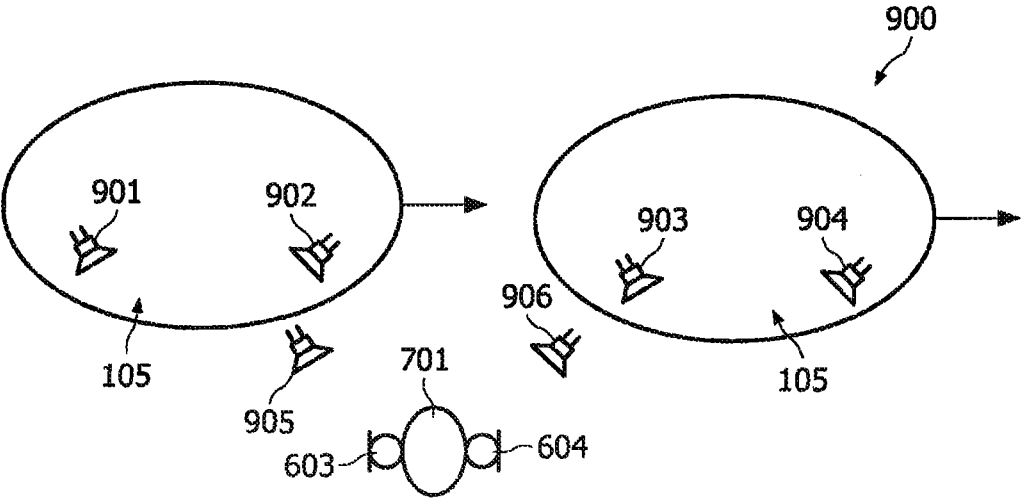


图 9