



(12)发明专利

(10)授权公告号 CN 105940380 B

(45)授权公告日 2019.03.15

(21)申请号 201580006162.9

(22)申请日 2015.02.27

(65)同一申请的已公布的文献号  
申请公布号 CN 105940380 A

(43)申请公布日 2016.09.14

(30)优先权数据  
14/194,192 2014.02.28 US

(85)PCT国际申请进入国家阶段日  
2016.07.27

(86)PCT国际申请的申请数据  
PCT/US2015/017992 2015.02.27

(87)PCT国际申请的公布数据  
W02015/131040 EN 2015.09.03

(73)专利权人 谷歌有限责任公司  
地址 美国加利福尼亚州

(72)发明人 威廉·巴斯蒂安·克雷杰  
图拉伊·扎基扎德·沙别斯塔雷

(74)专利代理机构 中原信达知识产权代理有限  
责任公司 11219  
代理人 李佳 穆德骏

(51)Int.Cl.  
G06F 11/27(2006.01)  
G06F 3/16(2006.01)

(56)对比文件  
US 2009/0044122 A1,2009.02.12,全文.  
mathieu lagrange等.Long Interpolation  
of Audio Signals Using Linear Prediction  
in Sinusoidal Modeling.《Journal of the  
Audio Engineering》.2005,第53卷(第10期),第  
891-905页.

审查员 赵小娟

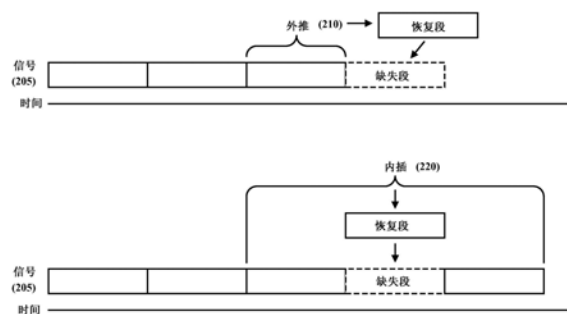
权利要求书3页 说明书12页 附图8页

(54)发明名称

跨缺失数据的正弦内插

(57)摘要

提供了用于隐藏音频信号中的缺失段和/或不连续性从而恢复信号的连续性的方法和系统。本方法和系统是针对音频信号设计的且以音频信号为目标,基于正弦波的内插和外推操作,并且不依赖于正弦波是谐波的假设。连同其它优点,本方法和系统相比于现有音频隐藏方法的改进在于本系统和方法促进了异步内插,如果信号是谐波则使用对应于时域波形内插的内插程序,并且具有对音频信号有效的峰值选择程序。



1. 一种用于恢复音频信号的连续性的计算机实现方法,所述方法包括:  
计算音频信号的段的频谱,所述段位于所述音频信号的缺失段的相对侧;  
确定所计算频谱的幅度峰值;  
根据幅度,将所确定的峰值排序;  
根据按照幅度被排序的所述峰值,从最高峰值开始,确定在所述缺失段的相对侧上的近邻的峰值,从而形成频谱峰值对;以及  
使用所述频谱峰值对,跨所述音频信号的所述缺失段执行复系数内插,以生成用于所述缺失段的恢复数据;以及  
使用所生成的恢复数据来恢复所述音频信号的连续性。
2. 根据权利要求1所述的方法,其中,所述内插是正弦内插。
3. 根据权利要求1所述的方法,其中,每个峰值对包括所述缺失段的第一侧上的第一峰值和所述缺失段的第二侧上的第二峰值。
4. 根据权利要求1所述的方法,其中,从最高峰值开始,针对每个所述峰值确定近邻的峰值。
5. 根据权利要求1所述的方法,其中,使用所述恢复数据包括:使用位于所述缺失段的相对侧上的所述段的数据来衰落所述恢复数据。
6. 根据权利要求5所述的方法,其中,使用重叠相加操作,以位于所述缺失段的相对侧上的所述段的数据来衰落所述恢复数据。
7. 根据权利要求1所述的方法,还包括:  
对所述信号执行快速傅立叶变换以生成所述信号的频域表示;  
平滑所述信号的幅度频谱;  
确定所平滑的频谱的局部斜率的符号;  
基于所述局部斜率的所述符号从正变成负的点,来生成所述峰值的估计;以及  
通过在对应的所估计峰值的邻域中搜索未平滑的信号,来确定每个所述峰值的真实位置。
8. 根据权利要求7所述的方法,其中,通过对所述频谱应用低通滤波器,来平滑所述信号的幅度频谱。
9. 根据权利要求1所述的方法,其中,从内插间隔的范围中选择所执行的所述内插的间隔,以使由内插器产生的瞬时波形和根据已知信号段估计的瞬时波形的对准度量最大化。
10. 根据权利要求9所述的方法,其中,选择所述内插间隔的范围,以将接收设备的抖动缓冲器填充至与预期分组丢失一致的水平。
11. 一种用于恢复音频信号的连续性的计算机实现方法,所述方法包括:  
计算音频信号的第一段的频谱和所述音频信号的第二段的频谱,其中,所述第一段位于所述信号的缺失段的第一侧上,以及所述第二段位于所述信号的缺失段的第二侧上;  
识别所述第一段和所述第二段的频谱中的多个对象,其中,所述多个对象对应于峰值;  
根据对应峰值的振幅,将所识别的对象排序(510);  
确定所排序对象的索引;以及  
针对所述对象中的至少一个,基于所排序对象的所述索引,从最高幅度开始,识别匹配对象;

使用所述至少一个对象和所述匹配对象,跨所述音频信号的所述缺失段执行复数系数内插,以生成所述缺失段的恢复数据;

使用所生成的恢复数据来恢复所述音频信号的连续性。

12. 根据权利要求11所述的方法,其中,根据所排序对象的所述索引,从具有最大峰值的所述对象开始且持续至具有最小峰值的所述对象,针对所述对象中的至少一个识别匹配对象。

13. 根据权利要求11所述的方法,其中,识别所述第一段和所述第二段的频谱中的所述多个对象包括:创建对象列表,所述第一段的频谱和所述第二段的频谱两者的每个峰值具有一个对象。

14. 根据权利要求13所述的方法,其中,针对所述列表中的每个,所述对象列表包括所述对应峰值的振幅、所述对应峰值所位于的频率、指定所述对应峰值是属于第一频谱还是第二频谱的标签、以及与所述对应峰值所属的频谱互补的频谱中的匹配峰值的位置。

15. 根据权利要求11所述的方法,其中,针对所述对象中的至少一个识别匹配对象包括:

通过在所有其它对象上执行搜索,来针对对象识别候选匹配对象;

从所述候选匹配对象中排除已经具有匹配对象的所有对象;

针对每个剩余候选匹配对象,评估匹配准则;以及

基于所评估的匹配准则,来将所述剩余候选匹配对象中的一个选择为所述对象的所述匹配对象。

16. 根据权利要求15所述的方法,其中,基于所述候选匹配对象具有对于所评估的匹配准则的最高值,而将所述剩余候选匹配对象中的一个选为所述匹配对象。

17. 根据权利要求15所述的方法,还包括:将所述剩余候选匹配对象局限于具有来自互补频谱的峰值的对象和具有在频率上接近的峰值的对象。

18. 根据权利要求15所述的方法,还包括:

识别没有候选匹配对象的对象;以及

创建虚拟峰值,所述虚拟峰值具有与所述对象的频率相同的频率,且在所述相同频率处具有零振幅或互补频谱的振幅。

19. 根据权利要求15所述的方法,还包括:

识别没有候选匹配对象的对象;以及

从所述多个对象中排除所述对象。

20. 一种用于恢复音频信号的连续性的计算机实现方法,所述方法包括:

计算音频信号的段的频谱,所述段位于所述音频信号的缺失段的相对侧上;

确定所计算频谱的幅度峰值;

根据幅度,将所确定的峰值排序;

根据按照幅度被排序的所述峰值,从最高峰值开始,确定在所述缺失段的相对侧上的近邻的峰值,从而形成频谱峰值对;以及

跨所述音频信号的所述缺失段的间隔执行内插,以生成用于所述缺失段的恢复数据,其中,基于将根据相位偏移的振幅而被加权的对应相位偏移集合对准,来确定所述间隔的长度;以及

使用所生成的恢复数据来恢复所述音频信号的连续性。

## 跨缺失数据的正弦内插

### 背景技术

[0001] 不同设备间的分组丢失和时钟漂移可以导致音频信号中的缺失段或不连续性。对个体而言,信号中的零的插入和不连续性两者都是明显可听到的。虽然已经提出了用于恢复信号的连续性的各种方法,但这种方法一般仅针对语音信号,并且一般并不很好地适用于音频。

### 发明内容

[0002] 本发明内容以简化形式介绍了概念的选择,以便提供对本公开的某些方面的基本理解。本发明内容并不是本公开的全面概述,并且并不意图识别本公开的关键或重要元素或描绘本公开的范围。本发明内容仅仅提出了本公开的某些概念,以作为下面提供的具体实施方式的前序。

[0003] 本公开一般涉及用于信号处理的方法和系统。更具体地,本公开的各方面涉及使用内插和外推操作来隐藏音频信号中的缺失段和/或不连续性。

[0004] 本公开的一个实施例涉及一种计算机实现方法,其包括:计算音频信号的段的频谱,该段位于音频信号的缺失段的相对侧;确定所计算的频谱的幅度峰值;根据幅度将所确定的峰值排序;根据该峰值来确定在缺失段的相对侧的近邻的峰值,从而形成频谱峰值对;以及使用该频谱峰值对来跨音频信号的缺失段执行内插以生成用于该缺失段的恢复数据。

[0005] 在另一实施例中,本计算机实现方法还包括使用位于缺失段的相对侧的段的数据来衰落该恢复数据。

[0006] 在另一实施例中,本计算机实现方法还包括:平滑信号的幅度频谱;确定所平滑的频谱的局部斜率的符号;基于局部斜率的符号从正变成负的点来生成峰值的估计;以及通过在对应的所估计的峰值的邻域中搜索未平滑的信号来确定每个峰值的真实位置。

[0007] 在另一实施例中,从内插间隔的范围中选择所执行的内插的水平,以使由内插器产生的瞬时波形和根据已知信号段估计的瞬时波形的对准度量最大化。

[0008] 本公开的另一实施例涉及一种计算机实现方法,其包括:计算音频信号的第一段的频谱和音频信号的第二段的频谱,其中,该第一段位于信号的缺失段的第一侧,以及第二段位于信号的缺失段的第二侧;识别第一段和第二段的频谱中的多个对象,其中,所述多个对象对应于峰值;根据对应峰值的振幅将所识别的对象分类;确定已分类对象的索引;以及针对对象中的至少一个识别匹配对象。

[0009] 在另一实施例中,本计算机实现方法中的识别第一段和第二段的频谱中的所述多个对象的步骤包括:创建对象列表,对于第一段的频谱和第二段的频谱两者的每个峰值具有一个对象。

[0010] 在另一实施例中,本计算机实现方法中针对对象中的至少一个识别匹配对象的步骤包括:通过在所有其它对象内执行搜索来针对对象识别候选匹配对象;从候选匹配对象中排除已经具有匹配对象的所有对象;针对每个剩余候选匹配对象评估匹配准则;以及基于所评估的匹配准则来将剩余候选匹配对象中的一个选为对于对象的匹配对象。

[0011] 在另一实施例中,本计算机实现方法还包括将剩余候选匹配对象局限于具有来自互补频谱的峰值的对象和具有在频率上接近的峰值的对象。

[0012] 在另一实施例中,本计算机实现方法还包括:识别没有候选匹配对象的对象,并且创建具有与对象的频率相同的频率且在相同频率下具有零振幅或互补频谱的振幅的虚拟峰值。

[0013] 在另一实施例中,计算机实现方法还包括识别没有候选匹配对象的队形,并将该对象从所述多个对象中排除。

[0014] 本公开的另一实施例涉及一种计算机实现方法,其包括:计算音频信号的段的频谱,该段位于音频信号的缺失段的相对侧;确定所计算的频谱的幅度峰值;根据幅度将所确定的峰值排序;根据该峰值来确定在缺失段的相对侧的近邻的峰值,从而形成频谱峰值对;以及跨音频信号的缺失段的间隔执行内插,其中,基于对准根据相位偏移的振幅所加权的对应相位偏移集合,来确定间隔的长度。

[0015] 在一个或多个其它实施例中,本文所述的方法和系统可选地可包括以下附加特征中的一个或多个:内插是正弦内插;每对峰值由缺失段的第一侧的第一峰值和缺失段的第二侧的第二峰值构成;从最高峰值开始针对每个峰值确定近邻的峰值;使用重叠相加操作以位于缺失段的相对侧的段的数据来衰落恢复数据;通过对频谱应用低通滤波器来平滑信号的幅度频谱;选择内插间隔的范围以将接收设备的抖动缓冲器填充至与预期分组丢失一致的水平;根据已分类对象的索引,从具有最大峰值的对象开始且持续至具有最小峰值的对象,针对对象中的至少一个识别匹配对象;针对列表中的每个对象,对象的列表包括对应峰值的振幅、对应峰值位于该处的频率、指定对应峰值是属于第一频谱还是第二频谱的标签以及与对应峰值所在的频谱互补的频谱中的匹配峰值的位置;和/或基于候选匹配对象对于所评估的匹配准则具有最高值而将剩余候选匹配对象中的一个选为匹配对象。

[0016] 根据下面给出的具体实施方式,本公开的进一步适用范围将变得显而易见。然而,应理解的是本具体实施方式和特定示例虽然指示了优选实施例,但仅仅是以举例说明的方式给出的,因为根据本具体实施方式,对于本领域的技术人员而言本公开的精神和范围内的各种改变和修改将变得显而易见。

## 附图说明

[0017] 根据结合所附权利要求和附图进行的以下具体实施方式的研究,本公开的这些及其它目标、特征和特性对于本领域的技术人员而言将变得更加显而易见,所述权利要求和附图全部构成本说明书的一部分。在所述附图中:

[0018] 图1是图示出根据本文所述的一个或多个实施例的用于使用内插和外推操作来隐藏音频信号中的缺失段和/或不连续性的示例性应用的示意图。

[0019] 图2是图示出根据本文所述的一个或多个实施例的对信号进行内插和外推操作的简化示例的示意图。

[0020] 图3是图示出根据本文所述的一个或多个实施例的其中数据音频信号的段的过去和未来末端两者处均可用的示例性内插方法的流程图。

[0021] 图4是图示出根据本文所述的一个或多个实施例的用于峰值提取的示例性方法的流程图。

[0022] 图5是图示出根据本文所述的一个或多个实施例的用于选择配对互补峰值的示例性方法的流程图。

[0023] 图6是图示出根据本文所述的一个或多个实施例的用于将对象配对的示例性方法的流程图。

[0024] 图7是图示出根据本文所述的一个或多个实施例的用于对应于音频信号的缺失段的过去和未来频谱中的正弦峰值的对象的示例性数据的表格。

[0025] 图8是图示出根据本文所述的一个或多个实施例的被布置成用于基于正弦波的内插和外推法来隐藏音频信号中的缺失段和/或不连续性的示例性计算设备的框图。

[0026] 在本文中提供的标题仅仅是为了方便起见,并且不一定影响本公开中要求保护的内容的范围或意义。

[0027] 在附图中,为了便于理解和方便起见相同的参考标号和任何缩写识别具有相同或类似结构或功能的元件或动作。在以下具体实施方式的过程中将详细地描述附图。

## 具体实施方式

### [0028] 概述

[0029] 现在将描述各种示例和实施例。以下描述提供了特定细节,以供透彻理解并实现这些示例的描述。然而,本领域的技术人员将理解的是本文所述的一个或多个实施例可在没有这些细节中的许多的情况下被实践。同样地,本领域的技术人员还将理解的是本公开的一个或多个实施例能够包括在本文中并未详细地描述的许多其它显而易见的特征。另外,下面可能未详细地示出或描述某些众所周知的结构或功能,以免不必要地使相关描述含糊难懂。

[0030] 本公开的实施例涉及用于隐藏(例如,恢复)音频信号中的缺失段和/或不连续性并从而恢复信号的连续性的方法和系统。该方法和系统是基于内插和外推操作,并且还可用来伸长或缩短音频段。

[0031] 恢复音频信号中的缺失段和/或不连续性的方法有时被称为分组丢失隐藏法。常规分组丢失隐藏法通常仅针对语音,因为语音信号在不可能进行重传的实时应用中是普遍的。然而,随着基于因特网的应用的发展,用于一般音频信号的分组丢失隐藏已变得更加重要。仅针对语音的现有分组丢失隐藏法严重地依赖于语音信号的节距结构,并且结果,这些方法对于音乐而言表现并不好,在音乐中,信号结构一般更加复杂,并且可以用于语音的简单模型一般未很好地表示该信号结构。因此,本公开的方法和系统被设计成对于音频(在更一般的意义上与仅语音相比)应用表现很好。

[0032] 鉴于上述现有方法的不足,本公开的实施例提供了能够实时操作的用于隐藏音频信号中的分组丢失和不连续性的方法和系统。

[0033] 如在本文中更详细地描述的,在进行外推后进行内插,并且在进行内插后是所观察到的数据。内插和外推操作两者将信号表示为正弦波的和。根据一个或多个实施例,对于外推,可在最后观察到的数据中检测到正弦波集合,而对于内插,可在缺失段之前观察或生成的数据中检测到第一集合的正弦波,并在缺失数据段之后的观察中检测到第二集合的正弦波。

[0034] 根据本文所述的至少一个实施例,本公开的内插法在将在缺失段之前和之后观察

到的正弦波配对之前将正弦波按其振幅分类。正弦波被描述为复指数的实部。对正弦波的复数增益和频率进行线性内插。

[0035] 如下面将更详细地描述的,内插法能够在两个模式下操作:同步模式和异步模式。在同步模式下,缺失数据的段的长度是固定的(并且正常与原始信号一致)。如果内插模型在同步模式下并未很好地拟合数据,则复指数的振幅将具有偏离,这是可听到的。在异步模式下,调整缺失数据段的长度以降低这些偏离的重要性。

[0036] 图1图示出根据本公开的一个或多个实施例的用于音频信号的外推和内插的示例性应用。例如,两个用户(例如,音乐家)120a和120b可在不同的物理位置(例如,远程环境)处演奏音乐,其中,用户120a在“位置1”处,以及用户120b在“位置2”处。另外,用户120a和120b可通过例如有线连接或网络进行相互通信,并且用户120a和120b中的每一个可具有分别用于再现和捕捉音频(例如,音乐)信号的扬声器105和扩音器110。多个信号源105(例如,扬声器)可位于多个传感器120之间(例如,扩音器、个体、其它音频捕捉设备等)。

[0037] 当跨通信网络发射数据(例如,音频)分组时,有时在网络的接收端处接收到分组之前,分组可能在延长的时间段内缺失,或者完全丢失。当分组在传输期间缺失时,必须做某些事以补偿此类数据丢失。例如,一个现有方法是插入零来代替缺失数据。然而,这种方法提供了不适当的解决方案,因为用户经受退化的性能和音频质量。

[0038] 针对恢复语音信号的其它常规分组丢失隐藏法通常借助于自回归(AR)模型(例如,滤波器)或其某种近似对信号的接收段进行外推。用于语音的自回归模型通常可以视为短期模型与长期模型的级联。短期模型描述频谱包络,而长期模型描述节距结构。对于分组丢失隐藏而言更重要的是长期模型,其通常捕捉范围从2.5至15毫秒(ms)的相关性。在某些情况下,为长期模型分配单位滤波增益。分组丢失隐藏的某些现有方法通过在缺失样本的间隙上运行具有零激励的AR模型(例如,滤波器)来进行缺失段的最小均方误差外推估计,该初始滤波状态基于间隙之前的信号。

[0039] 如果在开始隐藏操作时信号的过去段和未来段两者是已知的,则能够使用内插。按照惯例在通常涉及到较短间隙的信号恢复中使用内插法(例如,针对已损坏记录)。

[0040] 基于AR信号模型的一个现有内插方法描述了一种完全编码方案,其将编码器视为用于卡尔曼估计器的测量值集合,这自然地在分组丢失时导致了最佳内插,并且能够虑及与间隔相关的未来和过去的已知数据。然而,这种现有方法并不兼容结合现有编码器的操作。

[0041] 某些现有方法探索除了在数据缺失的段上进行外推和内插的AR模型之外的方法。例如,一个方法在根据间隙前后的信号的频谱所识别的正弦波之间使用用于内插的众所周知的正弦模型。其它方法将信号建模为适当基矢量与AR模型的和,并且然后同时优化基矢量的增益和AR模型信号贡献。该基矢量可以是正弦波,但是在这种方法中未提供用于基矢量的有效选择过程。

[0042] 一般地,现有内插法在计算上是昂贵的。例如,一个方法尝试使用线性预测器来预测跨缺失数据的非常长的段的正弦波的时间-频率轨迹。在实践中,不能使得此类轨迹的预测可靠。在另一方法中,结合隐式马尔可夫模型(HMM)使用谐波正弦模型在已知的语音段之间进行内插,以跟踪语音特征的演进。谐波假设意味着这种方法对音频信号并不是有效的。

[0043] 为了克服音频隐藏的现有方法的各种缺点,本公开的实施例提供了用于在缺失段

上进行外推和内插的改进方法和系统。如在本文中更详细地描述的,本方法和系统是针对音频信号设计的且以音频信号为目标,基于正弦波的内插和外推,并且不依赖于正弦波是谐波的假设。

[0044] 连同相比于现有方法的其它优点和改进一起,本公开的方法和系统(i)促进异步内插,(ii)如果信号是谐波,则使用对应于时域波形内插的内插程序,以及(iii)具有对音频信号有效的峰值选择程序。

[0045] 下面提出了对本公开的内插和外推使用基于正弦的方法的动机,提供了关于本方法的总体架构的细节,并且进一步描述了组成本系统的各种部件。

[0046] 傅立叶变换表示了基于复指数的信号(例如,音频信号)。由于该信号是实数,所以可以将复指数的扩展重新布置为正弦波的扩展。同样地,可以将频谱峰值解释为峰值频率下的正弦波的增益。而傅立叶变换以稳态正弦波的形式扩展信号(在离散傅立叶变换的情况下,正弦波在有限时间段上延伸),此观点的自然延伸是将信号视为具有连续变化增益和频率的正弦波的总和。然后可以在窗口化信号段上的傅立叶变换解释为这些正弦波的状态的近似快照,每个峰值表示正弦波。在此快照中幅度从峰值的逐渐滚降(也可将其视为快照的有限分辨率)是使用有限窗口长度的副作用。然后可以在快照之间内插单个正弦波。此原理构成了根据本文所述的一个或多个实施例的用于在音频信号的缺失段上进行内插的方法和系统的部分基础。

[0047] 如下面将更详细地描述的,本公开的算法旨在恢复连续缺失数据样本的段(其中,此类段可以是音频信号的缺失段或不连续性)。虽然数据在缺失数据段的过去末端处几乎始终可用,但对于段的未来侧而言情况不一定如此。根据至少一个实施例,可在开始算法之前指定将要恢复的数据段的长度。该恢复段可以短于缺失数据段,并且因此允许尚未接收到缺失段的未来侧的数据的情况。在那种情况下,算法可执行外推。在数据可用于缺失段的未来侧的情形中,可在缺失段上执行内插。同样地,根据本方法的至少一个实施例,在后续的外推操作集合之后是内插操作,并且该内插操作完成了缺失数据段的恢复。

[0048] 图2是根据本文所述的一个或多个实施例的对信号205(例如,音频信号)进行内插(220)和外推(210)的简化示例。外推(210)的应用不同于一般在基于AR模型的系统中使用的应用,其中,被外推的信号经受被外推的数据与缺失数据段的未来侧的信号数据之间的合并操作。

[0049] 图3图示出根据本文所述的一个或多个实施例的示例性内插操作300,其中,数据在将恢复的音频信号段的过去和未来末端两者处(例如,在音频信号的缺失段前后连续)可用。

[0050] 在方框305处,如果尚不可用,则可计算位于缺失段之前和之后(例如,直接或顺序位于缺失段之前和之后)的音频信号各段(例如,第一和第二段)的复短期频谱。在至少一个示例中,可在缺失段之前顺序出现(例如,在时间上较早)的音频信号的段视为在缺失段的第一侧(或过去末端),同时可将在缺失段之后顺序出现(例如,在时间上较晚)的段视为在缺失段的第二侧(或未来末端)。

[0051] 在方框310处,可确定在方框305处计算的频谱的幅度峰值。

[0052] 在方框315处,可根据幅度将在方框310处确定的峰值排序(例如,布置、列出等)。例如,根据至少一个实施例,可以以列表形式放置在方框310处确定的峰值并通过递增/递

减幅度来分类。应注意的是可将缺失段的过去末端和未来末端的峰值包含在单个列表中。例如,方框310处的列表可包括用于每个峰值的信息,峰值属于缺失段的过去末端还是未来末端。

[0053] 在方框320处,针对每个峰值,从最高峰值开始,可根据某些条件确定在(音频信号的)缺失数据段的另一侧的近邻。根据至少一个实施例,在方框320的执行期间可跳过已被选作近邻的峰值。根据幅度对峰值进行分类确保了最高峰值具有适当的近邻。应注意的是在方框320处的操作可留下对于可听性能而言不关键的不配对的某些小峰值(下面更详细地描述用以处理这些剩余峰值的一个方法)。

[0054] 在方框325处,可将结果得到的缺失数据段的未来末端和过去末端处的频谱峰值对用作用于跨该段进行内插的基础。

[0055] 在方框330处,可通过在内插阶段的过去末端和未来末端处的重叠相加操作,使用现有数据来衰落该数据。

[0056] 根据一个或多个实施例,除了以具有相位的适当旋进的过去末端频谱来替换未来末端频谱,本公开的外推操作类似于内插操作。正弦频率被假设为在外推期间是恒定的。根据至少一个实施例,如果在外推操作之后是另一外推,则可在外推操作的未来末端处省略该衰落操作。另一方面,如果在外推操作之后是内插,则可能需要进行衰落,以虑及并非未来-过去峰值对的一部分的小峰值。

[0057] 以下描述提供了关于本公开的分组丢失隐藏法的组成部分的附加细节。

[0058] 频谱估计和峰值提取

[0059] 为了能够内插正弦波,需要在缺失段的两个端(未来末端和过去末端)处对其进行检测,而外推依赖于在段的过去末端处正检测到的正弦波。虽然可使用各种方法来执行此检测,但重要的是以直观上促进准确的正弦波的内插的精度来找到峰值的振幅和频率。例如,对此目的至少8Hz的频率分辨率表现很好。

[0060] 根据至少一个实施例,可使用32秒(ms)的窗口。在16kHz下,此窗口对应于512个样本。可使用Hann窗口将信号窗口化并以零来填补该信号,以获得可对其执行快速傅立叶变换(FFT)的2048个样本的序列。

[0061] 图4图示出用于峰值提取的示例性过程。峰值提取过程400对幅度频谱进行操作,并且根据至少一个实施例,包括方框405—420。在方框405处,可对幅度频谱进行低通滤波(例如,平滑)以确保不考虑某些局部峰值。在至少一个实施方式中,可以约0.02秒的带宽来执行方框405处的对频谱的此低通滤波。

[0062] 在方框410处,可确定来自方框405的所平滑的频谱的局部斜率的符号。

[0063] 在方框415处,可基于斜率符号从正变为负的点来生成峰值的初始估计。

[0064] 在方框420处,可在初始峰值估计的邻域中搜索原始未平滑的信号中的真实峰值位置。例如,根据至少一个实施例,可在方框420处搜索的区域可以是带宽倒数(例如,25Hz)的半个宽度。

[0065] 配对选择

[0066] 对于内插情况,可将在缺失信号段的过去末端和未来末端处识别的显著正弦波配对。根据至少一个实施例,本发明的方法将峰值排序,并且然后从最高峰值开始找到适当的互补峰值。从最大峰值开始是重要失配的数目小(例如,如果有的话,失配一般仅在较小的

峰值上发生)的指示。

[0067] 图5图示出根据本文所述的一个或多个实施例的用于选择配对互补峰值的示例性过程500。在方框505处,过程500识别多个对象(例如,创建对象的列表或集合),其中,针对过去和未来频谱两者的每个峰值识别一个对象。现在目标是找到用于对象的匹配。根据至少一个实施例,每个已识别对象可具有四个标签:(1)峰值振幅;(2)峰值位于该处的频率;(3)指定峰值是属于未来频谱还是过去频谱的标签;以及(4)互补频谱中的匹配峰值的位置(如果不存在匹配的话,则为空或零)。

[0068] 图7图示出示例性表700,示例性表700包含用于在上文描述且在图5中图示出的示例性过程500的方框505处识别的每个对象的数据。根据至少一个实施例,表700可包括识别组成表700的每个条目的条目(或行)号705。表700中的每个条目705可以是不同对象710(例如,在示例性过程500的方框505处识别的对象),并且可以针对每个对象710包括对象的振幅(例如,峰值振幅)715、对象位于该处的频率720、指示对象是属于未来频谱还是过去频谱的标签725以及互补频谱中的匹配峰值在表700中的位置(例如,条目号705)(如果不存在匹配,则其可以留置为空或者包括零)。应理解的是在本文所述的一个或多个实施例中,除仅仅出于举例说明的目的而提供的上述示例性数据和格式之外或者作为其替代,表700可包括以各种其它格式布置的各种其它数据。

[0069] 返回参考图5的示例性过程500,在方框510处,可确定按对象的峰值的振幅对其进行分类时对象的索引。在方框515处,使用分类表的索引序列,从具有最大峰值的对象移动至具有最小峰值的对象,针对每个对象分离地执行实际配对。下面将参考图6来描述方框515处的关于对象配对的附加细节。在方框520处,缺失数据段(在方框515处形成)的未来末端和过去末端处的峰值可被用作跨该段的内插的基础。

[0070] 图6图示出用于将对象配对的示例性过程。根据本文所述的一个或多个实施例,示例性过程600(方框605—640)可以是上文描述且图5中图示出的示例性过程500中的方框515的一部分,其中,根据用于分类表的索引序列,从具有最大峰值的对象开始至具有最小峰值的对象,针对在方框505处创建的对象列表中的每个对象分离地执行配对。

[0071] 应理解的是,根据本公开的至少一个实施例,可针对示例性过程500的方框505处创建的对象列表中的每个对象(例如,对于过去和未来频谱两者的每个峰值的一个对象)以迭代方式执行示例性过程600(例如,过程600的方框605—640)。

[0072] 针对每个对象,可在方框605处做出关于正被评定的对象(仅仅为了明了起见有时在下文中称为“当前对象”)是否具有匹配对象的确定。如果在方框605处确定当前对象不具有匹配对象,则可在方框610处跳过当前对象。

[0073] 另一方面,如果在方框605处确定当前对象不具有匹配对象,则在方框615处,可对所有其它对象执行对匹配对象的搜索,以识别候选匹配对象。

[0074] 在方框620处,可排除作为候选匹配对象的已经具有匹配对象的候选匹配对象排除(例如,从候选匹配对象列表中去除)。

[0075] 在方框625处,可针对每个剩余候选匹配对象评估匹配准则。下面在等式(1)中更详细地提供了示例性匹配准则。

[0076] 在方框630处,可基于评估的匹配准则将候选匹配对象中的一个选为对于当前对象的匹配对象。例如,根据本文所述的一个实施例,在方框630处可将对于匹配准则而言(基

于例如在方框625处执行的评估与其它候选匹配对象的匹配准则值相比)具有最高(例如,最大)值的候选匹配对象选为对于当前对象的匹配对象。然而,应理解的是,情况可能是对于当前对象未找到匹配对象(例如,不存在匹配峰值)。例如,互补频谱可不具有留待选为匹配对象的对象,或者匹配准则可能不允许从互补频谱中选择剩余峰值。有利地,本公开的系统被设计成使得这种情形(即,不存在匹配对象)一般将仅针对较小的峰值而发生。

[0077] 虽然在图6中未示出,在方框630处选择匹配对象之后,可在当前对象的第四标签中识别(例如,写入、输入、插入等)匹配对象的列表位置,并且可在所选匹配对象的第四标签中同样地识别当前对象的位置(例如,在上文所述和图7中所示的示例性表700中)。

[0078] 如上所述,可针对示例性过程500的方框505处创建的对象列表(在上文描述并在图5中图示出)中的每个对象以迭代方式(例如,在所有对象范围内循环过程600)执行示例性过程600。

[0079] 根据本方法的至少一个实施例,将“附近频率”设置成100Hz。该准则与峰值乘以有利于附近峰值的距离相关加权成比例。也就是说,令 $j$ 识别候选对象, $a(j)$ 表示其峰值,以及 $f(j)$ 表示候选对象的频率。则该准则为:

$$[0080] \quad \eta = a(k) \cdot w(f(k), f(j)), \quad (1)$$

[0081] 其中, $f(k)$ 是当前对象的频率,并且 $w: \mathbb{R}_0^+ \times \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ 是将非负实频率映射到非负实距离惩罚值的函数。对 $w$ 的合理选择是取决于基准 $f_0$ 随距离的线性衰退:

$$[0082] \quad w(f(k), f(j)) = \max(f_0, f_1 - |f(k) - f(j)|). \quad (2)$$

[0083] 根据至少一个实施例, $f_1$ 是100Hz,并且 $f_0=1$ 。

[0084] 应注意的是在某些情况下并非所有由上述峰值识别算法识别的峰值都将获得匹配,因为在没有适当候选匹配对象的情况下可留下某些对象。根据本文所述的一个实施例,可通过创建具有相同频率且在该频率下具有零振幅或互补频谱的振幅的“虚拟”峰值来匹配此类对象(例如,在可以将信号解释为正弦波的简单和的情况下,则零振幅方法更加合理)。

[0085] 内插

[0086] 可以多种不同方式来执行正弦波的内插。因此,下文概述了不同内插方法的某些基本原理,并描述了根据本公开的实施例所利用的内插方法的细节。

[0087] 正弦波的参数是其振幅、其频率以及其相位。在下文中,术语“相位偏移”指代特定参考时刻的相位值。这些参数中的每一个是针对频谱的峰值指定的。然而,复杂化因素是频率是相位的导数。因此,通过执行振幅、频率以及相位偏移的线性内插,频率由于相位偏移并非恒定而遭受偏差。此频率偏差跨内插间隔是恒定的,并且等于内插间隔的倒数的一半或更小。

[0088] 相位偏移的线性内插意味着内插间隔的端点处的频率不等于在间隔的端点处估计的值。第一现有内插方法通过执行相位的三阶多项式内插来避免了此问题,其值和导数(频率)由两个端点处的测量值所确定。很明显,这一定导致比线性相位偏移内插更大但更短的频率偏离,因为总体相位旋进必须与线性相位偏移内插情况的相位旋进相同。这种第一内插方法的附加缺点是其导致了正弦函数的不规则采样,这阻碍了计算高效实施方式的构造。

[0089] 在实践中,对于语音上述第一内插方法被证明遭受混响性质,这可能归因于不期

望的频率偏离。对于语音,可以在不使用所观察的相位偏移而替代地使用谐波最小相位模型的第二现有内插方法中去除该混响性质。然而,此第二方法不适合于其中正弦波可能并非谐波的音乐。另外,此第二方法也未保留谐波信号的语音波形。

[0090] 通过将正弦波视为复指数的实部,第三内插方法自然地出现。例如,考虑  $\cos(ft + \phi) = \Re(\alpha \exp(ift))$ , 其中,  $f$  是频率,  $t$  是时间,  $\phi$  是相位偏移,  $\alpha$  是复系数,  $i = \sqrt{-1}$ , 以及运算符  $\Re$  提取实分量。则可以及时地执行频率  $f$  和复系数  $\alpha$  两者的线性内插。与上述第一和第二现有内插方法相反,复系数的内插导致了振幅在内插间隔上的变化。此第三方法还促进了复指数的规则采样,这导致了低的计算复杂性。

[0091] 对于其中正弦波集合是谐波且存在对准结果得到的波形的情形,复系数内插类似于用于语音的波形内插方法(例如,保留信号波形的波形内插尝试)。波形内插系统对于语音信号而言一般与三阶多项式内插方法相比混响较少。与第二内插方法(上文所述)相反,复系数内插保留语音信号的波形。

[0092] 根据本公开的一个或多个实施例,使用复系数内插方法。如果原始信号平滑地演进,并且重构数据的段的长度等于缺失数据的长度,则这种方法表现很好。

[0093] 可以通过利用此类信号本质上是谐波的事实而对发出的语音改善系统的内插性能。如果正弦频率接近于是谐波,则其并不是完美谐波的原因可能是测量值的结果。因此,在这种情况下,可以将频率修正成是谐波。例如,在实际实施方式中,如果在3000Hz以下的正弦波属于使用准则  $\xi = \frac{1}{N} \sum_{n=1}^N a_n (1 - \frac{f_0}{f_n} \text{round}(\frac{f_n}{f_0}))^2$  的发出语音段,则可以首先测试该正弦波,其中,  $f_0$  是候选节距,  $f_n$  是正弦波  $n$  的频率,  $a_n$  是正弦波  $n$  的幅度,  $N$  是在3000Hz下的正弦波的数目,以及函数  $\text{round}(\frac{f_n}{f_0})$  将  $\frac{f_n}{f_0}$  舍入至最近整数。首先,可针对使  $\xi$  最小化的  $f_0$  值执行搜索,并且然后可以使用用于该最佳  $\xi$  的阈值来判定信号是否是谐波。如果信号是谐波,则可将接近于是谐波的频率修改成是谐波(例如,  $f_0$  的倍数)。

[0094] 异步内插

[0095] 上述各种内插方法已经集中于在具有缺失数据段的长度的给定间隔上的内插。在根据本文所述的一个或多个实施例中可利用的异步内插中,以使内插中的振幅偏离最小化为目的来调整间隔的长度,因为此类偏移可作为混响而可听到。

[0096] 可将由索引  $n$  识别的正弦波的复系数的线性内插写为:

$$[0097] \quad \alpha_n(t) = \alpha_{n,-} + \alpha_{n,1}t, \quad (3)$$

[0098] 其中,  $t$  是时间,其中假设内插在  $t=0$  处开始,其中,  $\alpha_{n,-}$  是复系数的起始值,以及  $\alpha_{n,1}$  是影响变化速率的系数。另外,令  $\alpha_{n,+}$  表示用于所观察的未来数据的  $\alpha$  的估计值,以及  $\tau$  表示标称内插端点。则,对于线性内插,可将等式 (3) 重写为

$$[0099] \quad \alpha_n(t) = \alpha_{n,-} + \frac{\alpha_{n,+} - \alpha_{n,-}}{\tau} t, \quad (4)$$

[0100] 令  $f_{n,-}$  和  $f_{n,+}$  表示恢复间隙之前和之后的正弦频率。合理的目标是找到使下式最小化的内插间隔  $\tau$

$$\eta = \sum_{n=1}^N \left| \alpha_{n,+} - \frac{|\alpha_{n,+}|}{|\alpha_{n,-}|} \alpha_{n,-} \exp \left( i \int_0^{\tau} \left( f_{n,-} + \frac{f_{n,+} - f_{n,-}}{\tau} t \right) dt \right) \right|^2 \quad (5)$$

$$[0101] \quad = \sum_{n=1}^N \left| \alpha_{n,+} - \frac{|\alpha_{n,+}|}{|\alpha_{n,-}|} \alpha_{n,-} \exp \left( i \left( f_{n,-} \tau + \frac{1}{2} (f_{n,+} - f_{n,-}) \tau \right) \right) \right|^2 \quad (6)$$

$$= \sum_{n=1}^N |\alpha_{n,+}|^2 \left| \frac{\alpha_{n,+}}{|\alpha_{n,+}|} - \frac{\alpha_{n,-}}{|\alpha_{n,-}|} \exp \left( \frac{i(f_{n,-} + f_{n,+})\tau}{2} \right) \right|^2 \quad (7)$$

[0102] 可如下解释等式(7)。在每个时间点处,系数 $\alpha_n$ 和频率 $f_n$ 表征“瞬时”波形。然后,可以将上述结果(等式(7))理解为瞬时信号波形的对准度量。因数 $|\alpha_{n,+}|^2$ 表示正弦波n的重要性的加权。项 $\exp \left( \frac{i(f_{n,-} + f_{n,+})\tau}{2} \right)$ 表示间隔上的正弦波的相位的旋进,并且显示该旋进平均

来看以初始频率和最终频率的平均值前进。比值 $\frac{\alpha_{n,+}}{|\alpha_{n,+}|}$ 和 $\frac{\alpha_{n,-}}{|\alpha_{n,-}|}$ 表示起始和结束相位。同样地,总体目标与找到用于恢复间隔的结束时间 $\tau$ 一致,结束时间 $\tau$ 使得内插信号的相位与在未来末端处观察到的信号的相位很好地对准。

[0103] 鉴于异步内插的以上描述,根据本公开的一个或多个实施例,从内插间隔的范围中选择可执行的内插的水平以使由内插器产生的瞬时波形与根据已知信号段估计的瞬时波形的对准度量最大化。此外,根据本文所述的一个或多个其它实施例,可选择内插间隔的范围以将接收设备的抖动缓冲器填充至与预期分组丢失一致的水平。

[0104] 应注意的是根据本公开的至少一个实施例,可在不牺牲良好性能的情况下且以较低的计算复杂性省略等式(5)中的因数 $\frac{|\alpha_{n,+}|}{|\alpha_{n,-}|}$ 。

[0105] 以下是根据一个或多个实施例的本公开的异步内插过程的示例性性质的非穷举列表:

[0106] (1) 由于对准准则(等式(7))一般具有许多最小值,所以应执行在准则内的搜索。

[0107] (2) 本公开的异步内插方法可区别于直接的时域对准程序之处在于本文所述的方法包括频率的内插。

[0108] (3) 本公开的异步内插方法提供了任意长度的段上的内插;不再假设模型相对于所观察数据和缺失段长度的“合理”拟合。

[0109] (4) 本公开的内插方法可被用于通过插入虚拟缺失段且通过去除数据并使用内插来排除不连续性,减慢并在一定程度上加速音频段。

[0110] (5) 本文所述的内插方法使混响最小化,因为其使正弦波的振幅偏离最小化。重要的是在这方面注意到复系数内插方法在任何情况下不具有频率偏离。

[0111] 图8是根据本文所述的一个或多个实施例的被布置成用于基于正弦波的内插和外推来隐藏(例如,恢复)音频信号中的缺失段和/或不连续性的示例性计算机(800)的高层级框图。在非常基本的配置(801)中,计算设备(800)通常包括一个或多个处理器(810)和系统存储器(820)。存储器总线(830)可以被用于在处理器(810)与系统存储器(820)之间进行通信。

[0112] 根据期望的配置,处理器(810)可以是任何类型的,包括但不限于微处理器( $\mu$ P)、微控制器( $\mu$ C)、数字信号处理器(DSP)或其任何组合。处理器(810)可以包括高速缓存的一个或多个层级,诸如层级1高速缓存器(811)和层级2高速缓存器(812)、处理器核(813)以及

寄存器(814)。处理器核(813)可以包括算术逻辑单元(ALU)、浮点单元(FPU)、数字信号处理核(DSP核)或其任何组合。还可以将存储器控制器(816)与处理器(810)一起使用,或者在某些实施方式中,存储器控制器(815)可以是处理器(810)的内部部分。

[0113] 根据期望配置,系统存储器(820)可以是任何类型的,包括但不限于易失性存储器(诸如RAM)、非易失性存储器(诸如ROM、闪存等)或其任何组合。系统存储器(820)通常包括操作系统(821)、一个或多个应用(822)以及程序数据(824)。根据本文所述的一个或多个实施例,应用(822)可包括用于恢复音频信号中的一段连续缺失的数据样本的内插和外推算法(823)。程序数据(824)可包括存储指令,在该指令由一个或多个处理设备执行时该指令实现根据本文所述的一个或多个实施例的用于使用针对正弦波设计的内插和外推操作来隐藏音频信号中的缺失段和/或不连续性的方法。

[0114] 另外,根据至少一个实施例,程序数据(824)可包括音频信号数据(825),音频信号数据(825)可包括关于例如通过对信号进行窗口化而获得的音频信号的采样的数据。根据至少一个实施例,可使用在16kHz下对应于512个样本的32毫秒(ms)窗口。根据至少一个其它实施例,可使用Hann窗口对音频信号进行窗口化,并以零填补该音频信号以获得可在其上执行快速傅立叶变换(FFT)的2048个样本的序列。在某些实施例中,可以将应用(822)布置成使用程序数据(824)在操作系统(821)上进行操作。

[0115] 计算设备(800)可以具有附加特征或功能,以及用以促进基本配置(801)与任何所需设备和接口之间的通信的附加接口。

[0116] 系统存储器(820)是计算机存储介质的示例。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪存或其它存储技术、CD-ROM、数字多功能磁盘(DVD)或其它光盘储存器、磁带盒、磁带、磁盘存储器或其它磁存储器件或者可以用来存储期望信息且可以被计算设备800访问的任何其它介质。任何此类计算机存储介质可以是设备(800)的一部分。

[0117] 可以将计算设备(800)实现为小外形因数便携式(或移动)电子设备的一部分,诸如蜂窝电话、智能电话、个人数据助理(PDA)、个人媒体播放器设备、平板计算机(平板电脑)、无线网络手表设备、个人耳机设备、专用设备或包括任何上述功能的混合设备。还可以将计算设备(800)实现为包括膝上型计算机和非膝上型计算机配置两者的个人计算机。

[0118] 经由使用框图、流程图和/或示例前述详细描述已经阐述了设备和/或过程的各种实施例。在此类框图、流程图和/或示例包含一个或多个功能和/或操作的范围内,本领域的技术人员将理解的是此类框图、流程图或示例内的每个功能和/或操作可以单独地和/或共同地通过广泛的硬件、软件、固件或实际上其任何组合来实现。在一个实施例中可经由专用集成电路(ASIC)、现场可编程门阵列(FPGA)、数字信号处理器(DSP)或其它集成格式来实现本文所述的主题的多个部分。然而,本领域的技术人员将认识到本文公开的实施例的某些方面完全地或部分地可以以集成电路来等价地实现为在一个或多个计算机上运行的一个或多个计算机程序、在一个或多个处理器上运行的一个或多个程序、固件或者实际上其任何组合,并且就本公开而言为软件和/或固件而设计电路和/或编写代码将很好地在本领域的技术人员的技能内。

[0119] 另外本领域的技术人员将认识到本文所述的主题的机制能够以多种形式作为程序产品分发,并且无论用来实际上执行该分发的非临时信号承载介质的特定类型如何,本文所述的主题的说明性实施例都适用。非临时信号承载介质的示例包括但不限于以下各

项：诸如软盘、硬盘驱动、紧凑式磁盘 (CD)、数字视频磁盘 (DVD)、数字磁带、计算机存储器等的可记录类型介质；以及诸如数字和/或模拟通信介质 (例如，光纤电缆、波导、有线通信链路、无线通信链路等) 的传输型介质。

[0120] 关于本文中相当多的任何复数和/或单数术语的使用，本领域的技术人员可以在适合于上下文和/或应用时从复数转换成单数和/或从单数转换成复数。在本文中为了明了起见而明确地阐述了各种单数/复数变更。

[0121] 因此，已描述了本主题的特定实施例。其它实施例在以下权利要求的范围内。在某些情况下，在权利要求中叙述的动作可以按照不同的顺序执行且仍实现期望的结果。另外，在附图中描绘的过程不一定要求所示的特定顺序或连续顺序，以实现期望的结果。在某些实施方式中，多任务和并行处理可以是有利的。

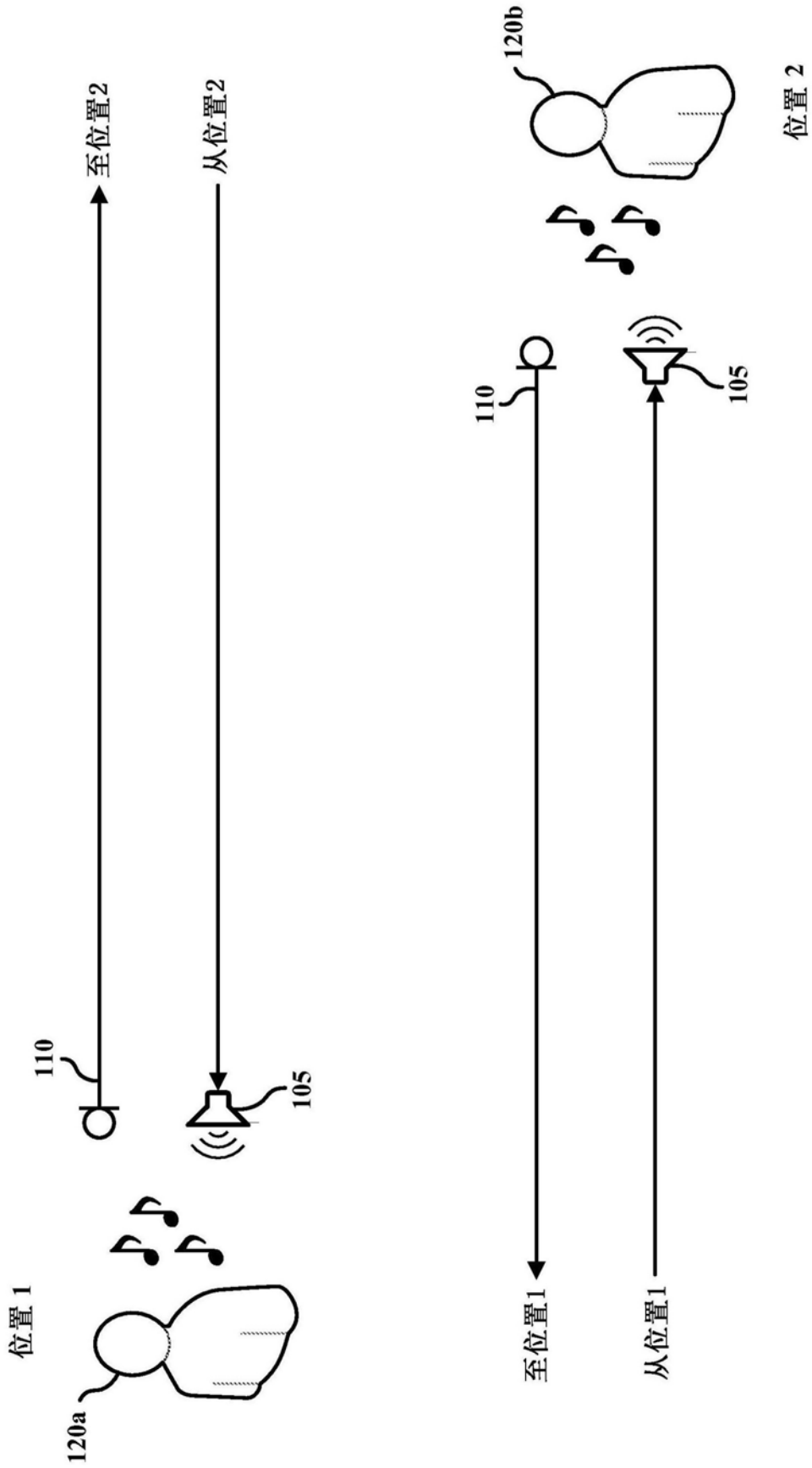


图1

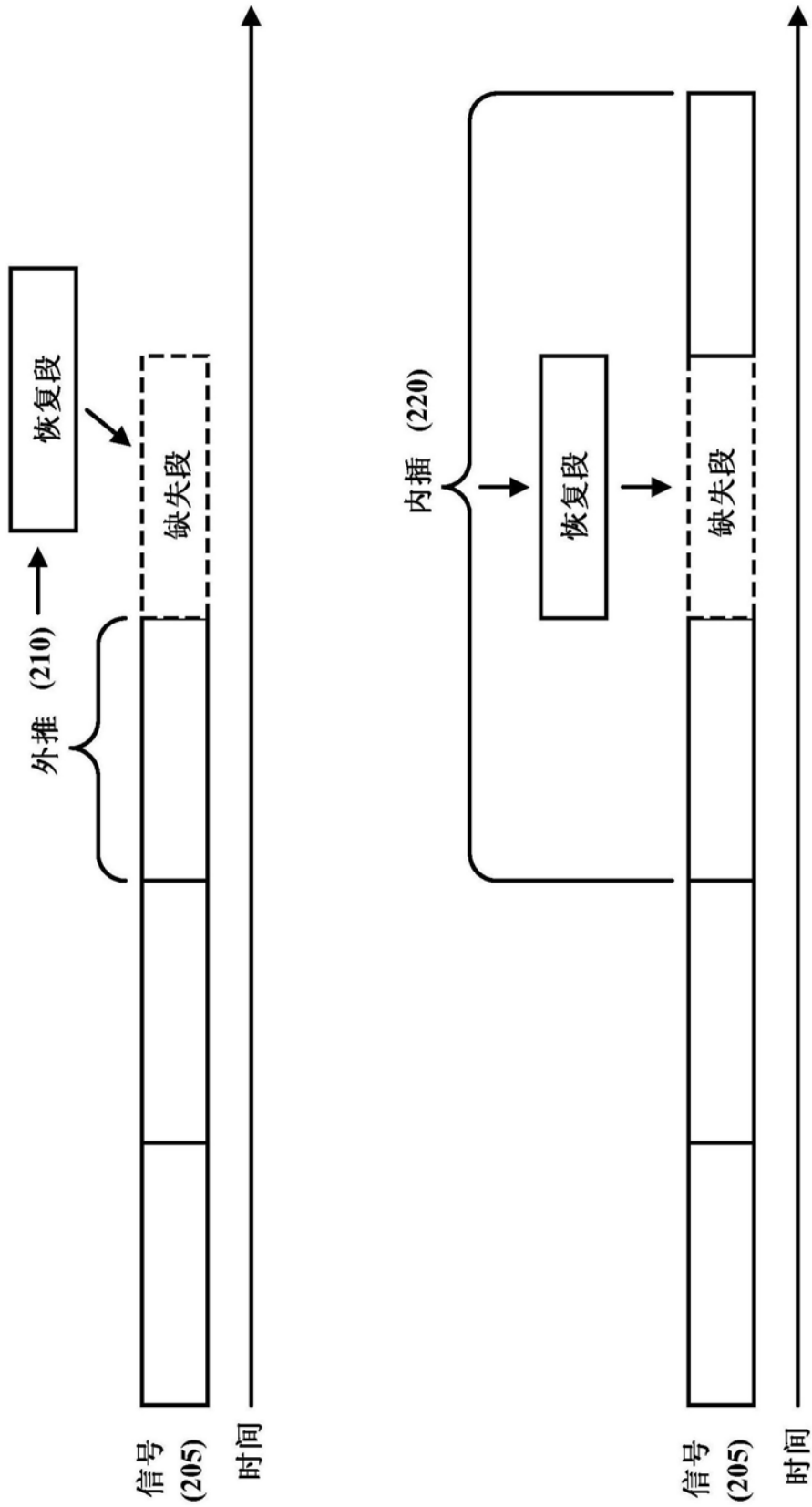


图2

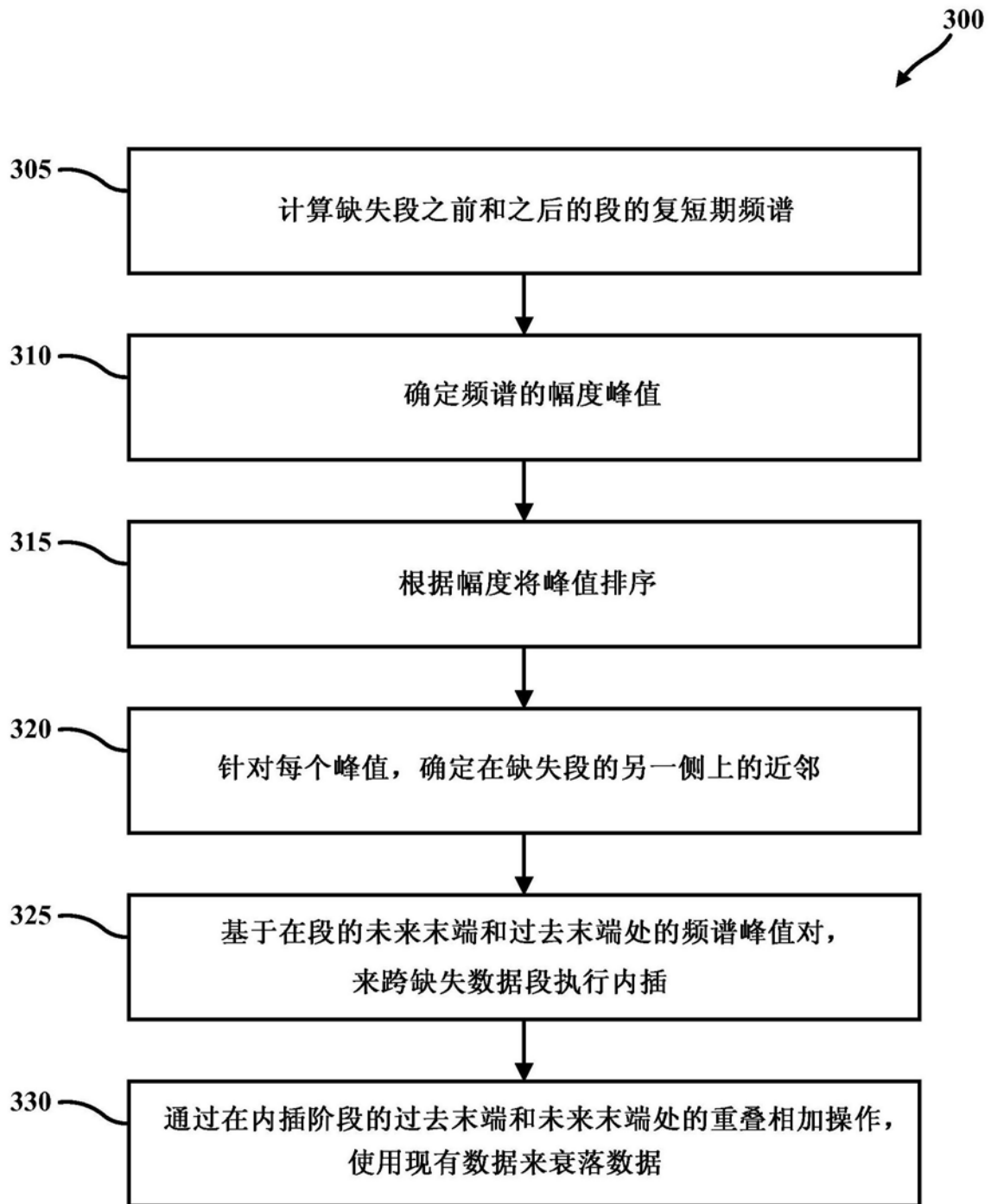


图3

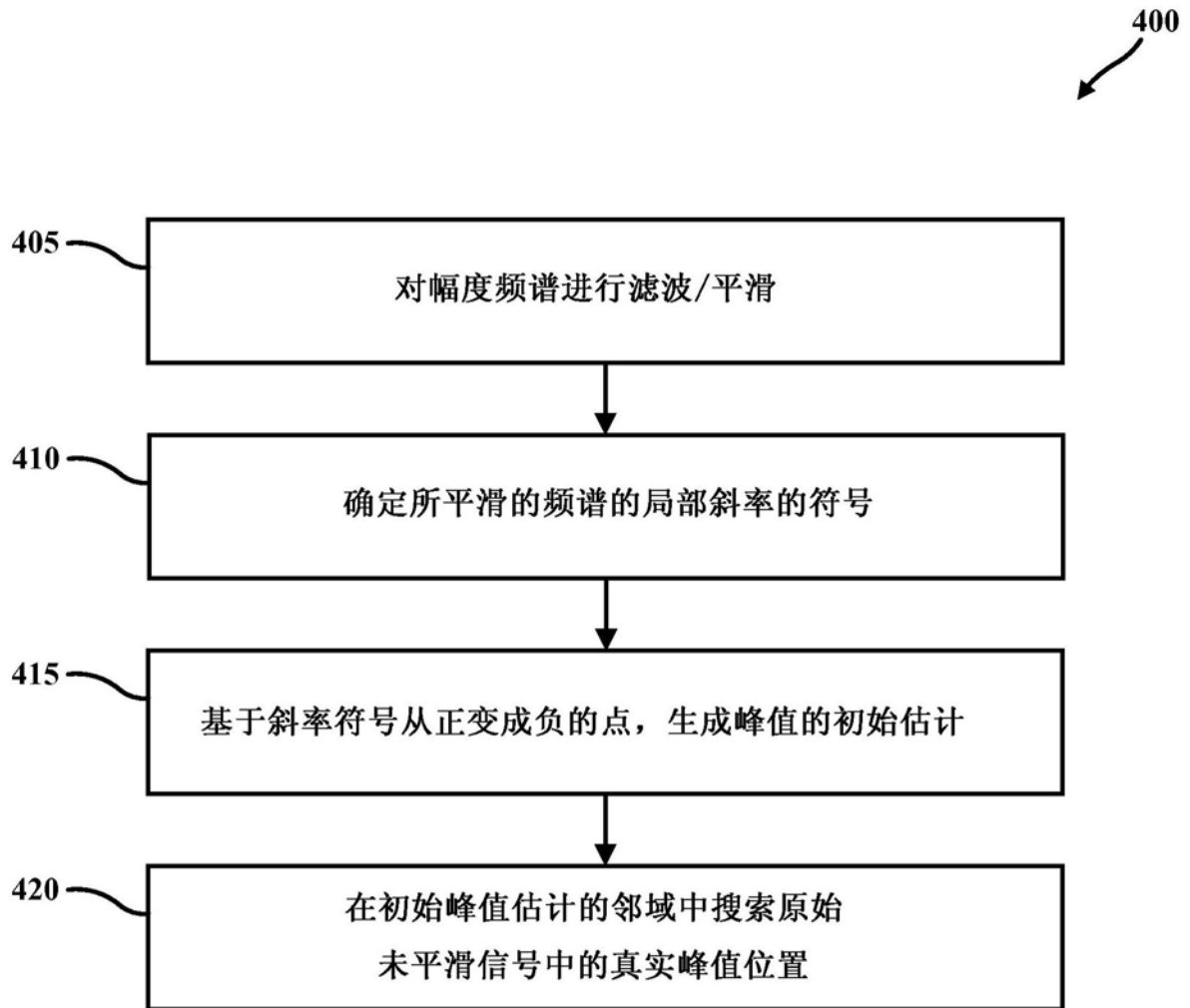


图4

500  
↙

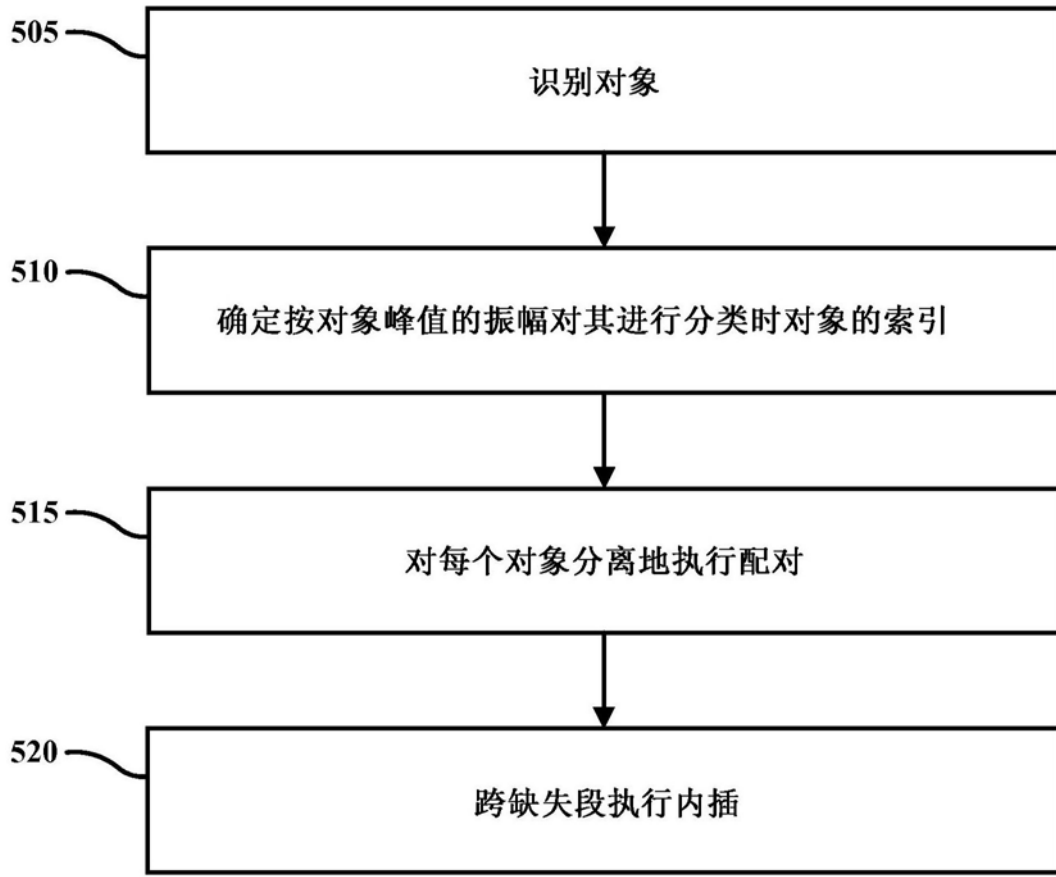


图5

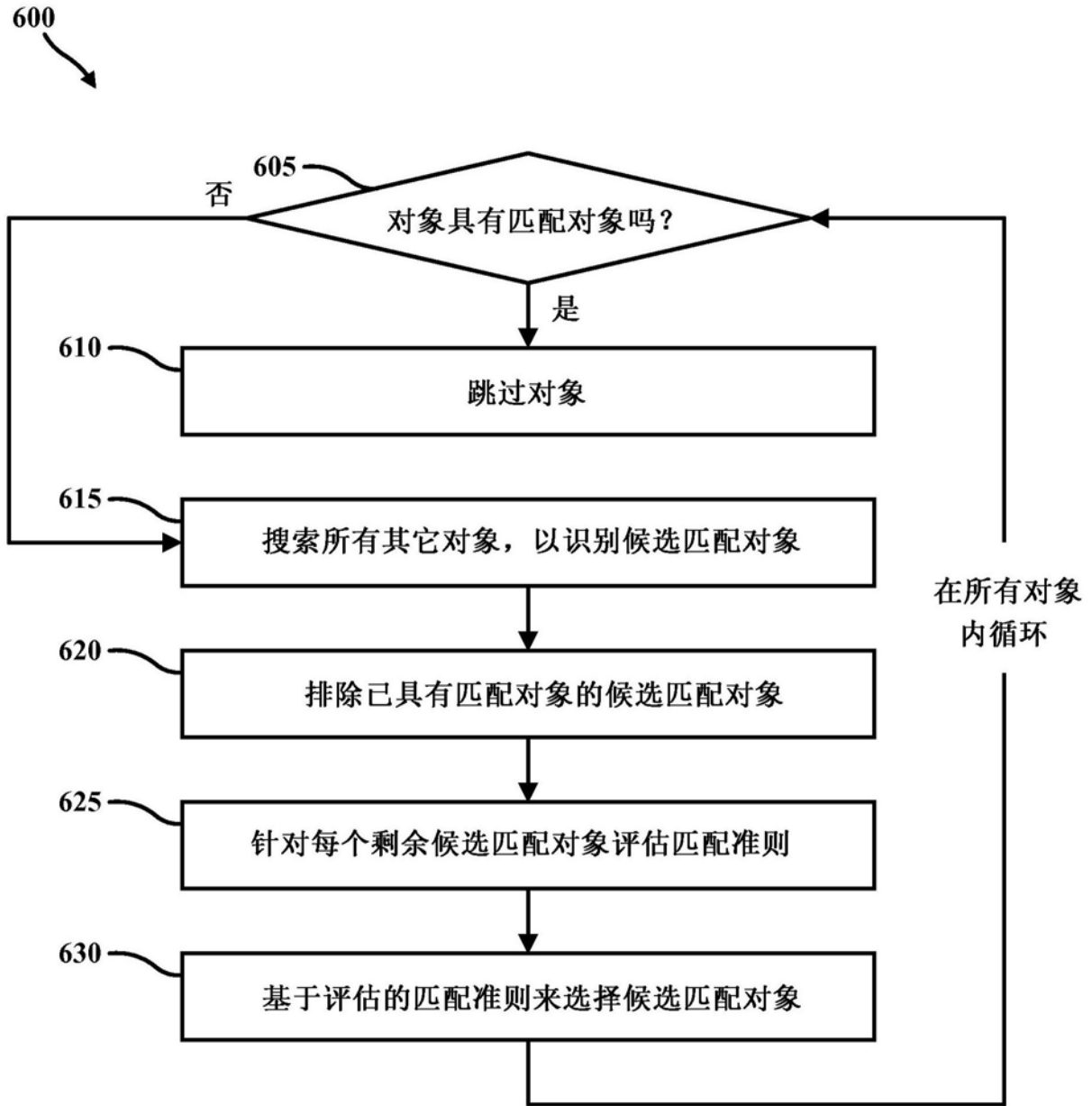


图6

700

	对象	振幅	频率位置	未来/过去频谱	匹配对象的 列表位置
1	对象 1	(振幅)	(频率)	未来	3
2	对象 2	(振幅)	(频率)	过去	5
3	对象 3	(振幅)	(频率)	过去	1
•	•	•	•	•	•
•	•	•	•	•	•
•	•	•	•	•	•
N	对象 N	(振幅)	(频率)	未来	X

图7

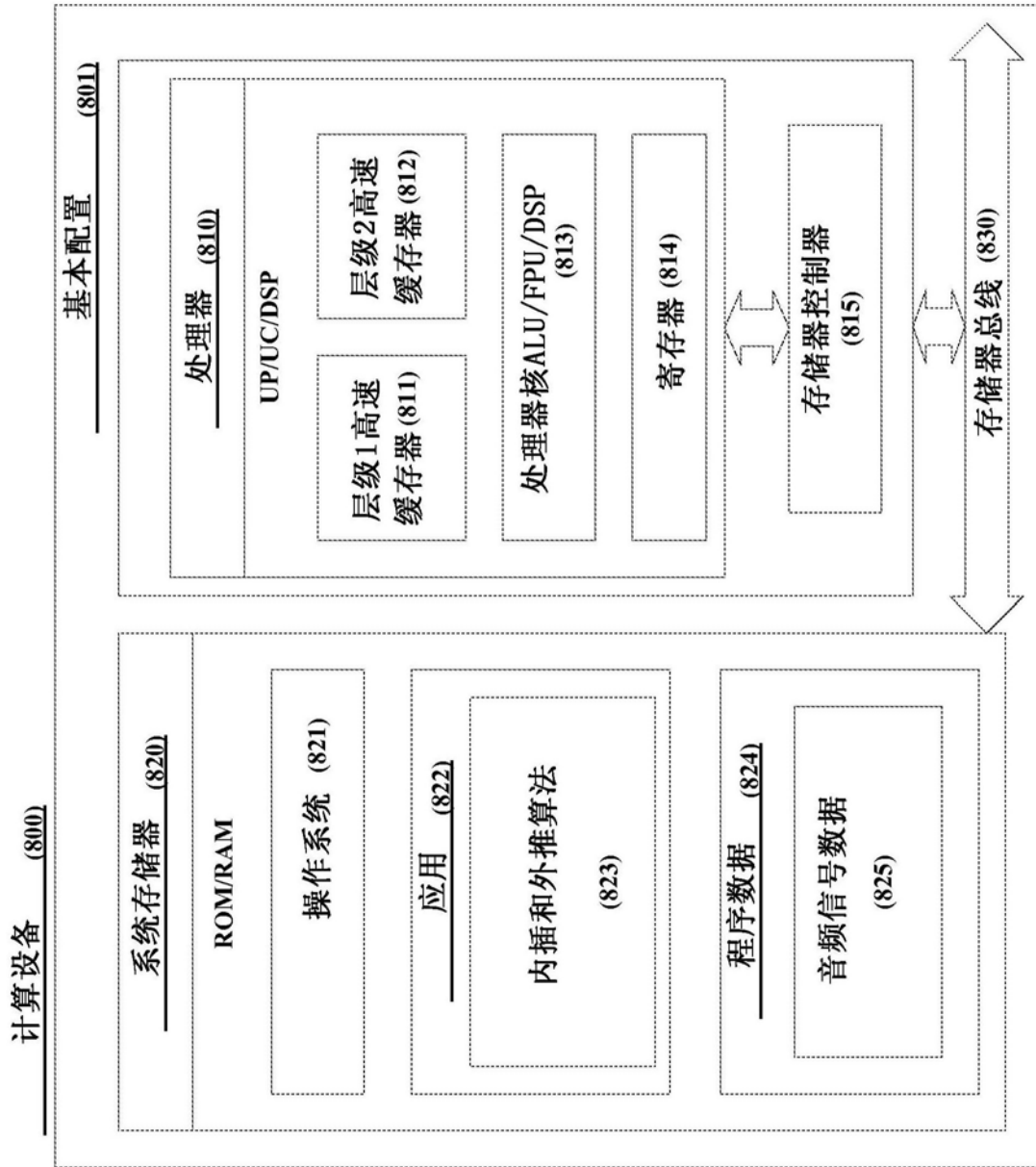


图8