



(12) 发明专利申请

(10) 申请公布号 CN 105099795 A

(43) 申请公布日 2015. 11. 25

(21) 申请号 201410151665. 2

(22) 申请日 2014. 04. 15

(71) 申请人 杜比实验室特许公司
地址 美国加利福尼亚州

(72) 发明人 霍家泉 孙学京 李凯

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227
代理人 杜诚 陈炜

(51) Int. Cl.

H04L 12/26(2006. 01)

H04L 29/06(2006. 01)

H04L 1/20(2006. 01)

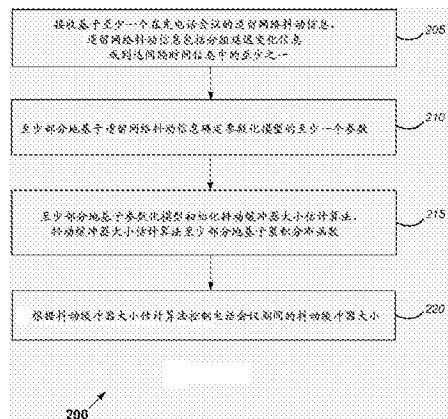
权利要求书3页 说明书13页 附图6页

(54) 发明名称

抖动缓冲器水平估计

(57) 摘要

一些实现涉及根据至少部分地基于累积分布函数(CDF)的抖动缓冲器大小估计算法控制电话会议期间的抖动缓冲器大小。CDF可以至少部分地基于网络抖动参数。可以根据参数化模型初始化CDF。参数化模型的至少一个参数可以至少部分地基于遗留网络抖动信息。



1. 一种方法,包括:

接收基于至少一个在先电话会议的遗留网络抖动信息,所述遗留网络抖动信息包括分组延迟变化(PDV)信息或到达间隔时间(IAT)信息中的至少之一;

至少部分地基于所述遗留网络抖动信息确定参数化模型的至少一个参数;

至少部分地基于所述参数化模型初始化抖动缓冲器大小估计算法,所述抖动缓冲器大小估计算法至少部分地基于网络抖动参数的累积分布函数(CDF);以及

根据所述抖动缓冲器大小估计算法控制电话会议期间的抖动缓冲器大小。

2. 根据权利要求1所述的方法,其中,控制所述抖动缓冲器大小涉及针对起音过程和衰减过程应用不对称的平滑参数,并且其中所述不对称的平滑参数是时变的。

3. 根据权利要求2所述的方法,还包括确定分组延迟变化(PDV),其中,针对起音过程和衰减过程应用不对称的平滑参数涉及:如果所述PDV大于当前抖动缓冲器大小,则应用起音平滑参数。

4. 根据权利要求2所述的方法,还包括确定分组延迟变化(PDV),其中,针对起音过程和衰减过程应用不对称的平滑参数涉及:如果所述PDV不大于当前抖动缓冲器大小,则应用衰减平滑参数。

5. 根据权利要求1至4中任一项所述的方法,其中,所述抖动缓冲器大小估计算法基于所述CDF并且基于第二分布函数。

6. 根据权利要求5所述的方法,其中,一个分布函数表示所述抖动缓冲器大小估计算法的确定性分量,另一个分布函数表示所述抖动缓冲器大小估计算法的非确定性分量。

7. 根据权利要求1至6中任一项所述的方法,其中,所述抖动缓冲器大小估计算法基于所述CDF并且基于参数化函数。

8. 根据权利要求7所述的方法,其中,所述参数化函数表示所述抖动缓冲器大小估计算法的确定性分量,所述CDF表示所述抖动缓冲器大小估计算法的非确定性分量。

9. 根据权利要求1至8中任一项所述的方法,其中,所述遗留网络抖动信息包括所述参数化模型的至少一个参数。

10. 根据权利要求1至9中任一项所述的方法,其中,所述遗留网络抖动信息包括与基于一个或多个在先电话会议的统计模型相对应的信息。

11. 根据权利要求1至10中任一项所述的方法,其中,所述遗留网络抖动信息包括与对来自至少一个在先电话会议的PDV数据或IAT数据中至少之一执行的曲线拟合过程相对应的信息。

12. 根据权利要求1至11中任一项所述的方法,其中,所述遗留网络抖动信息包括来自至少一个在先电话会议的PDV数据或IAT数据中至少之一的平均值。

13. 根据权利要求1至12中任一项所述的方法,其中,所述遗留网络抖动信息包括来自至少一个在先电话会议的原始PDV数据或原始IAT数据中的至少之一。

14. 根据权利要求1至13中任一项所述的方法,其中,所述遗留网络抖动信息基于一个或多个环境因素。

15. 根据权利要求14所述的方法,其中,所述环境因素包括网络类型、通信装置类型、目的地信息或往返行程延迟信息中的一个或多个。

16. 根据权利要求1至15中任一项所述的方法,还包括:

获取所述电话会议期间的当前网络抖动信息 ;以及
基于所述当前网络抖动信息更新遗留网络抖动信息的数据库。

17. 根据权利要求 1 至 16 中任一项所述的方法,其中,所述 CDF 至少部分地基于网络抖动参数。

18. 根据权利要求 17 所述的方法,其中,所述网络抖动参数为 PVD 或 IAT。

19. 根据权利要求 17 或 18 所述的方法,其中,根据所述抖动缓冲器大小估计算法控制所述抖动缓冲器大小涉及 :至少部分地基于当前网络抖动参数值确定是否递增所述 CDF。

20. 根据权利要求 17 至 19 中任一项所述的方法,其中,根据所述抖动缓冲器大小估计算法控制所述抖动缓冲器大小涉及 :根据所述 CDF 的函数和分组丢失阈值确定最小抖动缓冲器大小。

21. 一种设备,包括 :

接口系统 ;

存储器系统,其包括抖动缓冲器 ;以及

逻辑系统,其能够 :

经由所述接口系统接收基于至少一个在先电话会议的遗留网络抖动信息,所述遗留网络抖动信息包括分组延迟变化(PDV)信息或到达间隔时间(IAT)信息中的至少之一 ;

至少部分地基于所述遗留网络抖动信息确定参数化模型的至少一个参数 ;

至少部分地基于所述参数化模型初始化抖动缓冲器大小估计算法,所述抖动缓冲器大小估计算法至少部分地基于网络抖动参数的累积分布函数(CDF);以及

根据所述抖动缓冲器大小估计算法控制电话会议期间的抖动缓冲器大小。

22. 根据权利要求 21 所述的设备,其中,控制所述抖动缓冲器大小涉及 :针对起音过程和衰减过程应用不对称的平滑参数,并且其中所述不对称的平滑参数是时变的。

23. 根据权利要求 22 所述的设备,其中,所述逻辑系统能够确定分组延迟变化(PDV),其中针对起音过程和衰减过程应用不对称的平滑参数涉及 :如果所述 PDV 大于当前抖动缓冲器大小,则应用起音平滑参数。

24. 根据权利要求 23 所述的设备,其中,针对起音过程和衰减过程应用不对称的平滑参数涉及 :如果所述 PDV 不大于当前抖动缓冲器大小,则应用衰减平滑参数。

25. 根据权利要求 21 至 24 中任一项所述的设备,其中,所述抖动缓冲器大小估计算法基于所述 CDF 并且基于第二分布函数。

26. 根据权利要求 25 所述的方法,其中,一个分布函数表示所述抖动缓冲器大小估计算法的确定性分量,另一个分布函数表示所述抖动缓冲器大小估计算法的非确定性分量。

27. 根据权利要求 21 至 26 中任一项所述的设备,其中,所述抖动缓冲器大小估计算法基于所述 CDF 并且基于参数化函数。

28. 根据权利要求 27 所述的设备,其中,所述参数化函数表示所述抖动缓冲器大小估计算法的确定性分量,所述 CDF 表示所述抖动缓冲器大小估计算法的非确定性分量。

29. 根据权利要求 21 至 28 中任一项所述的设备,其中,所述接口系统包括网络接口、所述逻辑系统与所述存储器系统之间的接口、或外部装置接口中的至少之一。

30. 根据权利要求 21 至 29 中任一项所述的设备,其中,所述逻辑系统包括通用单芯片处理器或多芯片处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列

(FPGA) 或其他可编程逻辑器件、分立的门或晶体管逻辑、或分立的硬件部件中的至少之一。

31. 一种存储有软件的非暂态介质,所述软件包括用于控制至少一个设备进行以下操作的指令:

接收基于至少一个在先电话会议的遗留网络抖动信息,所述遗留网络抖动信息包括分组延迟变化(PDV)信息或到达间隔时间(IAT)信息中的至少之一;

至少部分地基于所述遗留网络抖动信息确定参数化模型的至少一个参数;

至少部分地基于所述参数化模型初始化抖动缓冲器大小估计算法,所述抖动缓冲器大小估计算法至少部分地基于网络抖动参数的累积分布函数(CDF);以及

根据所述抖动缓冲器大小估计算法控制电话会议期间的抖动缓冲器大小。

32. 根据权利要求 31 所述的非暂态介质,其中,控制所述抖动缓冲器大小涉及针对起音过程和衰减过程应用不对称的平滑参数,并且其中所述不对称的平滑参数是时变的。

33. 根据权利要求 32 所述的非暂态介质,其中,所述软件包括用于确定分组延迟变化(PDV)的指令,其中针对起音过程和衰减过程应用不对称的平滑参数涉及:如果所述 PDV 大于当前抖动缓冲器大小,则应用起音平滑参数。

34. 根据权利要求 33 中所述的非暂态介质,其中,针对起音过程和衰减过程应用不对称的平滑参数涉及:如果所述 PDV 不大于当前抖动缓冲器大小,则应用衰减平滑参数。

35. 根据权利要求 31 至 34 中任一项所述的非暂态介质,其中,所述抖动缓冲器大小估计算法基于所述 CDF 并且基于第二分布函数。

36. 根据权利要求 35 中所述的非暂态介质,其中,一个分布函数表示所述抖动缓冲器大小估计算法的确定性分量,另一个分布函数表示所述抖动缓冲器大小估计算法的非确定性分量。

37. 根据权利要求 31 至 36 中任一项所述的非暂态介质,其中,所述抖动缓冲器大小估计算法基于所述 CDF 并且基于参数化函数。

38. 根据权利要求 37 中所述的非暂态介质,其中,所述参数化函数表示所述抖动缓冲器大小估计算法的确定性分量,所述 CDF 表示所述抖动缓冲器大小估计算法的非确定性分量。

抖动缓冲器水平估计

技术领域

[0001] 本公开内容涉及音频信号处理。具体地,本公开内容涉及对远程通信的音频信号的处理,包括但不限于对电话会议或视频会议的音频信号的处理。

背景技术

[0002] 分组网络上的语音传输受到延迟变化,通常称为抖动(jitter)。例如,可以根据到达间隔时间(IAT)变化或分组延迟变化(PDV)来测量抖动。可以根据相邻分组的接收时间差来测量 IAT 变化。例如,可以参照与数据或“锚”分组接收时间的的时间间隔来测量 PDV。在基于因特网协议(IP)的网络中,固定延迟可以归因于由于材料和距离而导致的算法延迟、处理延迟和传播延迟,而可变延迟是由于 IP 网络流量的波动、因特网上不同的传输路径等引起的。

[0003] VoIP (网络电话)接收机通常依赖“抖动缓冲器”来对抗抖动的负面影响。通过引入接收音频数据分组的时刻与再现该分组的时刻之间的附加延迟,抖动缓冲器可以将到达分组的不均匀流转换成规则的分组流,使得延迟变化不会对终端用户造成可感知的音质下降。语音通信对于延迟非常敏感。例如,根据国际电信联盟(ITU)推荐 G. 114,对于普通对话,单向延迟应当保持低于 150ms,超过 400ms 被认为不可接受。因此,由抖动缓冲器添加的附加延迟需要足够小以避免造成可感知的音质下降。不幸的是,当分组由于网络延迟而晚于所期望的时间到达时,小的抖动缓冲器会导致更频繁的分组丢失。

发明内容

[0004] 根据本文中所描述的一些实现,方法可以涉及接收基于至少一个在先电话会议的遗留网络抖动信息。遗留网络抖动信息可以包括分组延迟变化(PDV)信息和 / 或到达间隔时间(IAT)信息。该方法可以涉及至少部分地基于遗留网络抖动信息确定参数化模型的至少一个参数。

[0005] 该方法可以涉及至少部分地基于参数化模型初始化抖动缓冲器大小估计算法。抖动缓冲器大小估计算法可以至少部分地基于网络抖动参数的累积分布函数(CDF)。该方法可以涉及根据抖动缓冲器大小估计算法控制电话会议期间的抖动缓冲器大小。在一些实现中,该方法可以涉及在电话会议期间获得当前网络抖动信息并且基于当前网络抖动信息更新遗留网络抖动信息的数据库。

[0006] 在一些示例中,控制抖动缓冲器大小可以涉及针对起音(attack)过程和衰减(decay)过程应用不对称的平滑参数。平滑参数可以是时变的。

[0007] 一些实现可以涉及确定网络抖动参数如分组延迟变化(PDV)。针对起音过程和衰减过程应用不对称的平滑参数可以涉及:如果 PDV 大于当前抖动缓冲器大小,则应用起音平滑参数。该方法可以涉及:如果 PDV 不大于当前抖动缓冲器大小,则应用衰减平滑参数。

[0008] 在一些实现中,抖动缓冲器大小估计算法可以基于 CDF 并且基于第二分布函数。例如,一个分布函数可以表示抖动缓冲器大小估计算法的确定性分量,另一分布函数可以

表示抖动缓冲器大小估计算法的非确定性分量。

[0009] 在替选实现中,抖动缓冲器大小估计算法可以基于 CDF 并且基于参数化函数。例如,参数化函数可以表示抖动缓冲器大小估计算法的确定性分量,CDF 可以表示抖动缓冲器大小估计算法的非确定性分量。

[0010] 根据一些实现,遗留网络抖动信息可以包括参数化模型的至少一个参数。在一些示例中,遗留网络抖动信息可以包括与基于一个或多个在先电话会议的统计模型相对应的信息。在一些实例中,遗留网络抖动信息可以包括与对来自至少一个在先电话会议的 PDV 数据和 / 或 IAT 数据执行的曲线拟合过程相对应的信息。可替代地或附加地,遗留网络抖动信息可以包括来自至少一个在先电话会议的 PDV 或 IAT 中至少之一的平均值。遗留网络抖动信息可以包括来自至少一个在先电话会议的原始 PDV 数据或原始 IAT 数据中至少之一。

[0011] 在一些实现中,遗留网络抖动信息可以基于一个或多个环境因素。环境因素例如可以包括网络类型、通信装置类型、目的地信息或往返行程延迟信息中的一个或多个。

[0012] 根据一些实现,CDF 可以至少部分地基于网络抖动参数。例如,网络抖动参数可以是 PDV 或 IAT。根据抖动缓冲器大小估计算法控制抖动缓冲器大小可以涉及至少部分地基于当前网络抖动参数值确定是否递增 CDF。根据抖动缓冲器大小估计算法控制抖动缓冲器大小可以涉及根据 CDF 的函数和分组丢失阈值确定最小的抖动缓冲器大小。

[0013] 根据一些实现,本文中所公开的这些方法和 / 或其他方法可以通过存储有软件的一个或多个非暂态介质来实现。软件可以包括用于控制一个或多个装置至少部分地执行这种方法的指令。

[0014] 本公开内容的至少一些方面可以经由设备来实现。例如,一个或多个装置能够至少部分地执行本文中所公开的方法。根据一些实现,设备可以包括接口系统、存储器系统以及逻辑系统。存储器系统可以被配置成包括抖动缓冲器。接口系统可以包括网络接口、逻辑系统与存储器系统之间的接口、和 / 或外部装置接口。逻辑系统可以包括通用单芯片或多芯片处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其他可编程逻辑器件、离散门或晶体管逻辑、或离散硬件部件中至少之一。

[0015] 根据一些实现,逻辑系统能够经由接口系统接收基于至少一个在先电话会议的遗留网络抖动信息。遗留网络抖动信息例如可以包括 PDV 信息和 / 或 IAT 信息。逻辑系统能够至少部分地基于遗留网络抖动信息确定参数化模型的至少一个参数。

[0016] 逻辑系统能够至少部分地基于参数化模型初始化抖动缓冲器大小估计算法。抖动缓冲器大小估计算法可以至少部分地基于网络抖动参数的累积分布函数(CDF)。逻辑系统能够根据抖动缓冲器大小估计算法控制电话会议期间的抖动缓冲器大小。

[0017] 在一些实现中,控制抖动缓冲器大小可以涉及针对起音过程和衰减过程应用不对称的平滑参数。平滑参数可以是时变的。逻辑系统能够确定网络抖动参数如分组延迟变化(PDV)。针对起音过程和衰减过程应用不对称的平滑参数可以涉及:如果 PDV 大于当前抖动缓冲器大小,则应用起音平滑参数。针对起音过程和衰减过程应用不对称的平滑参数可以涉及:如果 PDV 不大于当前抖动缓冲器大小,则应用衰减平滑参数。

[0018] 根据一些实现,抖动缓冲器大小估计算法可以基于 CDF 并且基于第二分布函数。例如,一个分布函数可以表示抖动缓冲器大小估计算法的确定性分量,另一个分布函数可以表示抖动缓冲器大小估计算法的非确定性分量。

[0019] 然而, 在一些实现中, 抖动缓冲器大小估计算法可以基于 CDF 并且基于参数化函数。例如, 参数化函数可以表示抖动缓冲器大小估计算法的确定性分量, CDF 可以表示抖动缓冲器大小估计算法的非确定性分量。

[0020] 在附图和下面的描述中阐述了本说明书中所描述的主旨的一个或多个实现的细节。根据该描述、附图和权利要求, 其他特征、方面和优点将变得明显。注意, 没有按比例绘制下面的附图的相对尺寸。

附图说明

[0021] 图 1A 是示意性示出可以应用本申请的实施方式的语音通信系统的示例的图;

[0022] 图 1B 是示意性示出可以实现本申请的方面的语音通信系统的另一示例的图;

[0023] 图 2 是示出本文所提供的一些抖动缓冲器控制方法的块的流程图;

[0024] 图 3 是示出 WiFi 网络中的分组延迟变化的示例的曲线图;

[0025] 图 4 是提供能够实现本公开内容的各个方面的设备的部件的示例的框图; 以及

[0026] 图 5 是提供音频处理设备的部件的示例的框图。

[0027] 各个图中的相似的附图标记和名称指示相似的元素。

具体实施方式

[0028] 下面的描述针对用于描述本公开内容的一些创新方面的某些实现, 以及可以实现这些创新的方面的上下文的示例。然而, 可以以多种不同的方式来应用本文中的教导。例如, 尽管就音频数据处理的具体示例描述了多种实现, 但是本文中的教导可广泛应用于其他已知的音频数据处理实现以及在未来可能引入的音频数据处理实现。

[0029] 在本文中同义地使用术语“块”和“分组”。因此, “音频块”或“音频数据的块”将具有与“音频分组”或“音频数据的分组”相同的含义。

[0030] 如本文中所使用的, 术语“缓冲器”可以指代用于暂时存储数据的物理存储器装置的区域, 或指代“指向”物理存储器中的位置的逻辑或虚拟数据缓冲器。“抖动缓冲器”将一般指代用于存储所接收的音频帧的逻辑或物理缓冲器。尽管抖动缓冲器通常可以用于在解码处理之前暂时存储编码音频数据, 但是依赖于具体实现, 抖动缓冲器可以存储各种形式的音频分组或音频帧。因此, 贯穿本说明书, 术语“抖动缓冲器”应当被理解为包括实际存储(或指向)音频帧的抖动缓冲器以及实际存储(或指向)各种形式的分组(块)的抖动缓冲器, 随后该各种形式的分组(块)在被播放或被反馈到部件以进一步处理之前被解码成音频帧。尽管通常在音频数据的再现或“回放”之前执行解码, 但可以不总是明确地结合缓冲处理讨论解码处理。因此, 本文中所使用的术语“帧”应当被广义地理解为包括已经从分组解码的帧、仍然编码在分组中的帧、本身包括一个或多个帧的分组、或者编码在分组中或已经从分组中解码的多于一个的帧。换言之, 在本申请的上下文中, 涉及帧的处理还可以被理解为涉及分组的处理或同时涉及包含在分组中的多于一个的帧的处理。

[0031] 在本申请的上下文中, 表达“同时”等的含义包括但不限于精确的字面上的含义, 并且应当被理解为“在预定粒度的相同的时间间隙/时间间隔内”。在本申请中, 例如, 预定粒度可以是两个相继发送的帧/分组之间的时间间隙(这样的时间间隙称为帧间隙)、或用于检查分组到达的网络探测速率、或处理时间粒度, 但不限于此。例如, 可以由帧时段/间

隙例如 20ms 对到达时间进行量化。替代地或附加地,可以将时间表示为对应于整数分组数量。类似地,在本申请的上下文中,当涉及具体时间点时,依赖于上下文,其也可以表示预定粒度的时间间隙。此外,当涉及针对某个帧的应当被理解为时间点的具体时间 t_i (i 为整数) 时,为了清楚起见,假定在发射机侧其表示该帧开始的时间点,或者在接收机侧其表示开始接收该帧时的时间点。

[0032] 如上面所提到的,抖动缓冲器可以通过暂时存储所接收的音频数据的分组(在本文中也称为“音频分组”)对抗网络不稳定性的负面影响,音频数据的分组可以对应于语音数据。在一些实现中,抖动缓冲器可以在音频分组被提供至解码器并且随后由通信终端的扬声器再现之前存储音频分组。

[0033] 图 1A 是示意性示出可以应用本申请的实施方式的语音通信系统的示例的图。如图 1A 中所示,对话参与者 A 操作通信终端 A,对话参与者 B 操作通信终端 B。类似本文中所示出和描述的其他通信终端,通信终端 A 和 B 可以包括例如电话、个人计算机、移动电子装置(例如,蜂窝电话、智能手机、平板电脑等)等。通信终端 A 和 B 可以包括诸如下面描述的那些(例如,参照图 4 和图 5)的部件。

[0034] 在语音通信会话如电话会议期间,对话参与者 A 和对话参与者 B 可以经由其通信终端 A 和 B 彼此谈话。在该示例中,通信终端 A 和 B 能够经由数据链路 103 通信。数据链路 103 可以实现为点到点连接,如通信网络等。

[0035] 在该示例中,通信终端 A 和 B 能够对由一个或更多个麦克风捕获的音频信号的音频块执行 VAD (语音活动检测)。如果在音频块中检测到存在语音,则可以由通信终端的逻辑系统对音频块进行相应的处理(例如,施加适合于语音数据的增益),并且可以经由数据链路 103 将音频块发送至另一对话参与者的通信终端。如果音频块中未检测到语音存在,则可以由通信终端的逻辑系统对音频块进行相应的处理(例如,施加适合于非语音数据的增益),并且经由数据链路 103 将音频块发送至另一对话参与者的通信终端。

[0036] 在该示例中,通信终端 A 和 B 能够向其他对话参与者的通信终端指示静默时间。如本文中所使用的,“静默时间”是在其期间对话参与者没有讲话的时间。在“静默时间”期间,对话参与者的通信终端可以检测非语音声音,如环境噪声。可以处理与这种声音相对应的音频数据,并且将该音频数据发送至一个或更多个其他通信终端。在一些实现中,可以通过发送静默时间元数据(例如,通过设置相应的标志或位)来指示静默时间,而在其他实现中,可以通过在与音频块相对应的时间周期期间什么也不发送来指示静默时间。在一些实现中,静默时间元数据可以对应于通信终端的“静音”控制的对话参与者的激活。

[0037] 在本实现中,通信终端 A 和 B 能够建立和控制抖动缓冲器,抖动缓冲器在图 1A 和图 1B 中表示为“JB”。在此,通信终端 A 和 B 能够接收例如编码音频数据作为音频块,并且将它们作为条目存储在抖动缓冲器中。条目可以对应于时间,例如,接收到音频块的时间。抖动缓冲器中的音频块可以被解码和/或以其他方式处理以用于由通信终端的一个或更多个扬声器再现。接收到静默时间元数据或什么也没接收到可以导致抖动缓冲器中的相应空条目。通信终端 A 和 B 能够如本文中所示和所描述的控制抖动缓冲器大小。

[0038] 图 1B 是示意性示出可以实现本申请的方面的语音通信系统的另一示例的图。在该示例中,可以在对话参与者 A、B 和 C 之间进行语音会议。

[0039] 如图 1B 所示,对话参与者 A 操作通信终端 A,对话参与者 B 操作通信终端 B,对话

参与者 C 操作通信终端 C。在语音会议会话期间,对话参与者 A、对话参与者 B 和通信参与者 C 可以分别通过其通信终端 A、B 和 C 彼此谈话。从会话参与者 A、B 和 C 的角度来看,图 1B 中示出的通信终端能够提供与图 1A 中示出的功能基本相同的功能。虽然图 1B 中示出了三个通信终端,但是其他实现可以涉及更多或更少通信终端。

[0040] 然而,在图 1B 中示出的示例中,通信终端 A、B 和 C 被配置成通过公共的数据链路 113 或分离的数据链路 113 与另一装置(在该示例中为服务器)通信。数据链路 113 可以实现为点到点连接或通信网络。例如,通信终端 A、B 和 C 能够如上面所描述的对由通信终端捕获的音频信号的音频块执行 VAD 和适当的处理。

[0041] 在该实现中,通信终端 A、B 和 C 能够向服务器指示静默时间。在一些实现中,可以通过发送静默时间元数据(例如,通过设置相应的标志或位)来指示静默时间,而在其他实现中,可以通过在与音频块相对应的时段期间什么也不发送来指示静默时间。通信终端 A、B 和 C 能够使用所发送的音频分组包括“时间戳”或类似时间元数据,指示音频分组的传输时间。

[0042] 在该实现中,服务器也能够建立和控制抖动缓冲器。在图 1B 中示出的示例中,服务器已经建立了与通信终端 A、B 和 C 中的每个相对应的抖动缓冲器 JB_A 、 JB_B 和 JB_C 。例如,服务器能够如本文中所公开的控制抖动缓冲器大小。在该实现中,服务器能够接收由通信终端 A、B 和 C 发送的音频块,并且能够将它们存储为抖动缓冲器 JB_A 、 JB_B 和 JB_C 中的与音频块的时间相对应的条目。例如,服务器能够将音频块存储为抖动缓冲器 JB_A 、 JB_B 和 JB_C 中的与音频块的时间戳相对应的条目。接收静默时间元数据(或什么也不接收)可以导致抖动缓冲器中的相应空条目。

[0043] 在该示例中,服务器也能够将来自抖动缓冲器 JB_A 、 JB_B 和 JB_C 中的每个的对应于相同时间的音频块混合成混合音频块。混合音频块的副本可以被发送至通信终端 A、B 和 C 中的每个。服务器可以使混合音频块中的每个包括一种或更多种时间戳。

[0044] 在该示例中,通信终端 A、B 和 C 能够建立和控制抖动缓冲器 JB 。通信终端 A、B 和 C 可以接收来自服务器的混合音频块,并且可以将它们存储为与混合音频块的时间相对应的条目。例如,抖动缓冲器条目可以对应于接收到音频块的时间。在每个通信终端中,抖动缓冲器中的音频块可以由通信终端的扬声器系统解码和再现。通信终端 A、B 和 C 能够如本文中所公开的控制抖动缓冲器大小。

[0045] 确定适当的抖动缓冲器大小可能是充满挑战的。太小的抖动缓冲器尤其在由例如增加的网络活动造成的延迟尖峰的时间期间可以造成不可接受的音频分组数量下降。然而,太长的抖动缓冲器可以导致可感知的延迟。在本文中所描述的一些实现中,可以根据至少部分地基于涉及至少一个网络抖动参数的累积分布函数(CDF)的抖动缓冲器大小估计算法控制抖动缓冲器大小。这样的网络抖动参数例如可以包括分组延迟变化(PDV)信息和/或到达间隔时间(IAT)信息。

[0046] 如上面指出的,在本文中术语“到达间隔时间”用于表示相邻分组之间的接收时间差。如本文中所使用的,参照参考分组或“锚”分组来测量“分组延迟变化”。例如,假定在时间 $R(0)$ 接收到第一分组,并且分组之间的预期时间间隔(可以称为“分组间隙”)为 20ms,那么接下来的分组的“预期”接收时间为 $R(0)+20ms$ 、 $R(0)+40ms$ 、 $R(0)+60ms$ 、...等。可以将抖动测量为接收分组的实际时间与接收该分组的预期时间之间的差异,例如,如下:

$$[0047] \quad d_t = \left\lfloor \frac{R_t - R_0 - tL}{L} \right\rfloor \quad (\text{等式 1})$$

[0048] 在等式 1 中, 下标 t 表示分组索引, d_t 表示与第 t 个分组相关联的 PDV, R_t 表示第 t 个分组的到达时间, L 表示分组之间的预期时间间隔。

[0049] 可以如下表达 PDV 的 CDF :

$$[0050] \quad c(\tau) := P(d_t < \tau) \quad (\text{等式 2})$$

[0051] 在等式 2 中, $P(d_t < \tau)$ 表示 d_t 小于变量 τ 的任何给定值的概率。

[0052] 本文中所公开的一些实现涉及根据 CDF 的函数以及分组丢失阈值来确定最小抖动缓冲器大小, 例如, 如下 :

$$[0053] \quad B^* = \operatorname{argmin}_{\tau} \{ \tau : c(\tau) \geq 1 - C \} \quad (\text{等式 3})$$

[0054] 在等式 3 中, B^* 表示最佳抖动缓冲器大小, C 表示分组丢失阈值。例如, C 可以被设置成能够容忍的最大分组丢失。在该示例中, C 是满足 $0 < C < 1$ 的常数。例如, 在一些实现中, C 可以小于 0.1, 例如 0.03、0.02、0.01、0.005 等。

[0055] 然而, 实际上, 将实时估计 CDF。在一些这种实现中, 根据抖动缓冲器大小估计算法控制抖动缓冲器大小可以涉及如果当前网络抖动参数值超过阈值则递增 CDF, 例如, 如下 :

$$[0056] \quad \hat{c}_t(\tau) = \begin{cases} \alpha \hat{c}_{t-1}(\tau) + (1 - \alpha)\tau \geq d_t \\ \alpha \hat{c}_{t-1}(\tau) & \text{否则} \end{cases} \quad (\text{等式 4})$$

[0057] 在等式 4 中, $\hat{c}_t(\tau)$ 表示估计的 DCF, α 表示平滑参数, 在该示例中, α 在 0 到 1 ($[0, 1]$) 的范围内。在一些实现中, α 可以至少在某些时候与接近 0 的程度相比更接近 1。例如, 在一些实现中, α 可以是 0.9 或更大, 例如在 0.95、0.96、0.97、0.98、0.99、0.991 等范围内。然而, 如下面更详细讨论的, 在一些实现中, α 可以是时变的。本文中所公开的一些实现涉及对于“起音”处理和“衰减”处理应用不对称的平滑参数。下面提供这种实现的示例。在等式 4 中, d_t 表示网络抖动参数的当前值, 在该示例中 d_t 为 PDV, $\hat{c}_{t-1}(\tau)$ 表示在时间 $t-1$ 处 τ 的估计的 CDF。根据具体的实现, 对应于 d_t 的值可以是例如时间值或分组数量值 (例如, 分组的数量, 对应于分组之间的预期时间间隔的倍数)。

[0058] 因此, 在一些实现中, 可以如下估计最佳抖动缓冲器大小 :

$$[0059] \quad \hat{B}^* = \operatorname{argmin}_{\tau} \{ \tau : \hat{c}_t(\tau) \geq 1 - C \} \quad (\text{等式 5})$$

[0060] 在等式 5 中, \hat{B}^* 表示最佳抖动缓冲大小的估计。

[0061] 有效的 CDF 将满足下面的约束 :

$$\hat{c}_t(\tau) \leq 1 \quad \forall \tau$$

[0062]

$$\hat{c}_t(\tau_0) \leq \hat{c}_t(\tau_1) \quad \forall \tau_0 < \tau_1$$

[0063] 可以证明, 当根据等式 3 更新 CDF 时保持这些属性。使 a 、 A 、 B 为范围 $[0, 1]$ 内的实数, 使得 $A \leq B$, 并且使得 $Q(\cdot)$ 为量化操作符。因为 $Q(\cdot)$ 是单调的, 因此有 :

$$\begin{aligned}
 & A \leq B \\
 [0064] \quad & \xrightarrow{\Delta} \alpha A \leq \alpha B \\
 & \xrightarrow{\Delta} Q(\alpha A) \leq Q(\alpha B)
 \end{aligned}$$

[0065] 进一步假定 $(1-\alpha)$ 是预先计算的常数。可以继续上面的推理以推断：

$$\begin{aligned}
 & A \leq B \\
 [0066] \quad & \xrightarrow{\Delta} \alpha A \leq \alpha B \\
 & \xrightarrow{\Delta} Q(\alpha A) \leq Q(\alpha B) \\
 & \xrightarrow{\Delta} Q(\alpha A) \leq Q(\alpha B) + (1-\alpha) \\
 & \xrightarrow{\Delta} Q(Q(\alpha A)) \leq Q(Q(\alpha B) + (1-\alpha))
 \end{aligned}$$

[0067] 以及

$$\begin{aligned}
 & A \leq B \\
 [0068] \quad & \xrightarrow{\Delta} \alpha A \leq \alpha B \\
 & \xrightarrow{\Delta} Q(\alpha A) \leq Q(\alpha B) \\
 & \xrightarrow{\Delta} Q(\alpha A) + (1-\alpha) \leq Q(\alpha B) + (1-\alpha) \\
 & \xrightarrow{\Delta} Q(Q(\alpha A) + (1-\alpha)) \leq Q(Q(\alpha B) + (1-\alpha))
 \end{aligned}$$

[0069] 因此,如果初始化时 CDF 满足该不等式,当根据等式 3 更新 CDF 时,将满足不等式 $c_t(B) \leq c_t(B+1)$ 。如果进一步的合理假设为 $Q(Q(\alpha \times 1) + (1-\alpha)) = 1$,那么保证根据等式 3 更新的 CDF 是有限精度算法中的有效的 CDF。

[0070] 在一些实现中,可以根据参数化模型如指数函数、指数分布或帕累托(Pareto)分布来初始化 CDF。例如,可以基于涉及指数函数的概率密度函数(PDF)来初始化 CDF,例如：

$$[0071] \quad f(x|\alpha) = \begin{cases} \alpha^x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (\text{等式 6})$$

[0072] 在等式 6 中, α 表示平滑参数, x 表示网络抖动参数如 IAT 或 PDV。可以将基于等式 4 的 CDF 表达为：

$$[0073] \quad F(x|\alpha) = \int_{-\infty}^x f(x|\alpha) dx$$

[0074] 或者,可以基于具有指数分布形状的 PDF 初始化 CDF,例如：

$$[0075] \quad f(x|\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (\text{等式 7})$$

[0076] 在等式 7 中, λ 表示速率参数。在一些示例中, $\lambda = 1/x_{\text{平均}}$ 。可以将基于等式 5 的 CDF 表达为：

$$[0077] \quad F(x|\lambda) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

[0078] 或者,可以基于具有帕累托分布形状的 PDF 初始化 CDF,例如：

$$[0079] \quad f(x|\alpha, k) = \frac{\alpha k^\alpha}{x^{\alpha+1}}; k \leq x < \infty; \alpha, k > 0 \text{ (等式 8)}$$

[0080] 在等式 8 中, k 表示变量 x 的最小值, α 表示形状参数, α 有时也被称为“尾指数”。可以将基于等式 6 的 CDF 表达为:

$$[0081] \quad F(x|\alpha, k) = 1 - \left(\frac{k}{x}\right)^\alpha; k \leq x < \infty; \alpha, k > 0$$

[0082] 本文中所公开的一些实现涉及确定用于至少部分地基于遗留网络抖动信息初始化 CDF 的参数化模型的至少一个参数。遗留网络抖动信息可以基于操作环境的在先知识, 例如用于电话会议和 / 或在电话会议期间移动终端在其上通信的网络的通信终端的在先知识。该知识越准确, CDF 越好地收敛于基本的 IAT 和 / 或 PDV 分布。该收敛可以提供改进的抖动缓冲器性能。

[0083] 图 2 是示出本文中所提供的一些抖动缓冲器控制方法的块的流程图。方法 200 例如可以(至少部分地)由服务器或被配置用于与通信终端如上面参照图 1B 描述的终端通信的另一这种装置执行。然而, 可以(至少部分地)由通信终端执行本文中所提供的一些方法。例如, 可以根据存储在非暂态介质上的软件来执行这些方法。至于本文中所描述的其他方法, 方法 200 的块不一定以所表示的顺序执行。而且, 方法 200 的一些实现(以及本文中所公开的其他方法)可以包括比所表示的或所描述的块更多或更少的块

[0084] 在该示例中, 方法 200 开始于块 205, 块 205 涉及接收基于至少一个在先电话会议的遗留网络抖动信息。遗留网络抖动信息例如可以包括分组延迟变化 (PDV) 信息或到达间隔时间 (IAT) 信息中的至少之一。遗留网络抖动信息可以基于一个或更多个环境因素, 如网络类型、通信装置类型(例如, 通信终端类型)、目的地信息和 / 或往返行程延迟信息。

[0085] 根据一些实现, 遗留网络抖动信息可以选自数据库。该数据库可以本地存储或远程存储。例如, 数据库是经由网络(例如, 经由服务器, 如图 1B 中示出的服务器)可访问的。遗留网络抖动信息例如可以对应于一个或更多个环境因素。例如, 数据库可以包括基于在多个网络上进行的在先电话会议的遗留网络抖动信息、通信装置类型等。根据一些这种示例, 当初始化电话会议时, 可以根据在当前电话会议期间将使用的网络、通信装置类型等查询数据库。可以在块 205 中接收与当前电话会议的环境因素的最接近匹配相对应的遗留网络抖动信息。

[0086] 在图 2 中示出的示例中, 块 210 涉及至少部分地基于遗留网络抖动信息确定参数化模型的至少一个参数。在一些实现中, 遗留网络抖动信息可以包括来自至少一个在先电话会议的原始网络抖动参数值, 如原始 PDV 数据或原始 IAT 数据。基于这种原始数据, 可以确定原始 PDV 数据或原始 IAT 数据的平均值。根据该平均值, 可以将指数分布的速率参数 λ 的估计确定为 $\lambda = 1/x_{\text{平均}}$, 其中, $x_{\text{平均}}$ 表示网络抖动参数如 PDV 或 IAT 的平均值。

[0087] 或者, 遗留网络抖动信息可以基于原始网络抖动参数值如原始 PDV 数据或原始 IAT 数据, 但可以以更高效的方式存储和 / 或提供给后续处理。例如, 遗留网络抖动信息可以包括来自至少一个在先电话会议的网络抖动参数如 PDV 和 / 或 IAT 的平均值。根据该平均值, 可以确定指数分布的速率参数 λ 的估计。替代地或附加地, 遗留网络抖动信息可以包括与对来自至少一个在先电话会议的网络抖动参数如 PDV 数据和 / 或 IAT 数据执行的曲

线拟合过程相对应的信息。

[0088] 在一些实现中,遗留网络抖动信息可以涉及参数化模型的至少一个参数。在这种实现中,已经执行了上述计算,并且已经存储了相应值用于后续检索。在一些示例中,遗留网络抖动信息可以包括与基于在一个或多个在先电话会议期间检测到的网络抖动数据的统计模型相对应的信息。

[0089] 在此,块 215 涉及至少部分地基于参数化模型初始化抖动缓冲器大小估计算法。在一些实现中,参数化模型可以是指数函数、指数分布、帕累托分布或另外的参数化模型。在该示例中,抖动缓冲器大小估计算法至少部分地基于网络抖动参数的累积分布函数(CDF)。然而,如下面更详细地描述的,在一些实现中,抖动缓冲器大小估计算法可以至少部分地基于 CDF 并且基于第二函数。第二函数例如可以是参数化函数或者第二分部函数如第二 CDF。网络抖动参数或多个参数例如可以是 PDV 和 / 或 IAT。

[0090] 在该实现中,块 220 涉及在电话会议期间根据抖动缓冲器大小估计算法控制抖动缓冲器大小。在一些实现中,根据抖动缓冲器大小估计算法控制抖动缓冲器大小可以涉及根据当前网络抖动参数值如瞬间 PDV 来更新 CDF 估计。一些这种实现可以涉及至少部分地基于当前网络抖动参数值确定是否递增 CDF。在一些示例中,根据抖动缓冲器大小估计算法控制抖动缓冲器大小可以涉及根据 CDF 的函数和分组丢失阈值确定抖动缓冲器大小。例如,参照等式 3 至等式 5,在前面的讨论中提供了两者的示例。

[0091] 在一些实现中,在电话会议期间获得的当前网络抖动信息可以用于更新遗留网络抖动信息的数据库。例如,可以存储原始 PDV 或 IAT 数据。在一些实现中,可以确定和存储 PDV 或 IAT 数据的平均值。在一些示例中,可以对来自电话会议的 PDV 或 IAT 数据执行曲线拟合过程。在一些实现中,可以根据在电话会议期间获得的当前网络抖动信息确定参数化模型的参数。

[0092] 可以使用一种或更多种前述类型的网络抖动信息更新遗留网络抖动信息的数据库。这种信息可以与对应于电话会议的一个或多个环境因素如网络类型、通信装置类型、目的地信息或往返行程延迟信息相关联。

[0093] 在一些实现中,控制抖动缓冲器大小可以涉及针对起音过程和衰减过程应用不对称的平滑参数。在一些这种实现中,平滑参数可以是时变的平滑参数。如本文中所使用的,术语“起音”可以对应于对延迟尖峰(例如,通过增加抖动缓冲器大小)的响应,“衰减”可以对应于非起音过程,如在起音过程之后返回较低的抖动缓冲器大小。依赖于具体实现,涉及不对称的平滑参数和 / 或时变的平滑参数的实现可以结合涉及遗留网络抖动信息的实现来使用,或者可以不结合涉及遗留网络抖动信息的实现来使用。例如,在本文中所描述的一些实现中,可以根据至少部分地基于涉及至少一个网络抖动参数的累积分布函数的抖动缓冲器大小估计算法来控制抖动缓冲器大小。抖动缓冲器大小估计算法可以涉及应用不对称的平滑参数和 / 或时变的平滑参数。然而,在一些这种实现中,可以不基于遗留网络抖动信息初始化抖动缓冲器大小估计算法。例如,针对起音过程和衰减过程应用不对称的平滑参数可以涉及如果 PDV 大于当前抖动缓冲器大小则应用起音平滑参数。在一些实现中,针对起音过程和衰减过程应用不对称的平滑参数可以涉及如果 PDV 不大于当前抖动缓冲器大小则应用衰减平滑参数。

[0094] 这种实现能够使得抖动缓冲器更响应于网络抖动的瞬间尖峰,同时保证相对低的

迟的分组丢失率。例如,在一些实现中,可以如下确定起音平滑参数或衰减平滑参数:

$$[0095] \quad \alpha = \begin{cases} \alpha_{\text{attack}}, & \text{如果 } PDV > B^* \\ \alpha_{\text{decay}}, & \text{否则} \end{cases} \quad (\text{等式 } 9)$$

[0096] 在等式 9 中, α 表示平滑参数,其可以是在本文所阐述的其他等式(如等式 4)中所使用的平滑参数。在等式 9 中, α_{attack} 表示起音平滑参数, α_{decay} 表示衰减平滑参数, PDV 表示当前分组延迟值, B^* 表示当前抖动缓冲器大小。根据等式 9, 默认平滑参数为 α_{decay} 。仅在当前分组延迟值超过当前抖动缓冲器大小时,将平滑参数设置成 α_{attack} 。一些这种实现涉及响应于网络抖动的尖峰递增起音平滑参数,例如,如下:

$$[0097] \quad \alpha_{\text{attack}} = \min(\alpha_{\text{max}}, \alpha_{\text{attack}} + \alpha_{\text{attack_step}}) \quad (\text{等式 } 10)$$

[0098] 在等式 10 中, α_{attack} 表示起音平滑参数, α_{max} 表示最大平滑参数值, $\alpha_{\text{attack_step}}$ 表示起音平滑参数值的递增。在一些实现中,可以如下确定 α_{max} 和 $\alpha_{\text{attack_step}}$:

$$[0099] \quad \alpha_{\text{max}} = 1 - \frac{1}{T_{\text{max}} / T_{\text{packet}}} \quad (\text{等式 } 11)$$

$$[0100] \quad \alpha_{\text{attack_step}} = \frac{\alpha_{\text{max}} - \alpha_{\text{attack_min}}}{T_{\text{max}} / T_{\text{packet}}} \quad (\text{等式 } 12)$$

[0101] 在等式 11 和等式 12 中, T_{max} 表示平滑的最大时间常数,在一些实现中,其可以在几十秒的范围内,例如 10 秒、15 秒、20 秒、25 秒、30 秒等。在等式 11 和等式 12 中, T_{packet} 表示分组的到达之间的预期时间间隔, $\alpha_{\text{attack_min}}$ 表示 α_{attack} 的最小值。在一些实现中,可以如下确定 $\alpha_{\text{attack_min}}$:

$$[0102] \quad \alpha_{\text{attack_min}} = 1 - \frac{1}{T_{\text{attack_min}} / T_{\text{packet}}} \quad (\text{等式 } 13)$$

[0103] 在等式 13 中, $T_{\text{attack_min}}$ 表示“起音平滑”的最小时间间隔。换言之, $T_{\text{attack_min}}$ 表示在其期间可以将平滑参数设置成起音平滑参数 α_{attack} 的最小的时间间隔。在一些实现中, $T_{\text{attack_min}}$ 可以在几秒的范围内,例如 1 秒、2 秒、3 秒、4 秒、5 秒、6 秒等。通过每个时间间隔期间递增起音平滑参数——在每个时间间隔期间当前分组延迟值超过当前起音缓冲器大小(例如,受到最大平滑参数值和 / 或最大时间间隔)——平滑变得越来越积极(更平滑),并且抖动缓冲器变得相对较少响应于瞬间网络抖动。

[0104] 根据一些这种实现,可以如下确定衰减平滑参数 α_{decay} :

$$[0105] \quad \alpha_{\text{decay}} = \min(\alpha_{\text{max}}, \alpha_{\text{decay}} + \alpha_{\text{decay_step}}) \quad (\text{等式 } 14)$$

[0106] 在等式 14 中, $\alpha_{\text{decay_step}}$ 表示 α_{decay} 的递增。在一些实现中,可以如下确定 $\alpha_{\text{decay_step}}$:

$$[0107] \quad \alpha_{\text{decay_step}} = \frac{\alpha_{\text{max}} - \alpha_{\text{decay_min}}}{T_{\text{max}} / T_{\text{packet}}} \quad (\text{等式 } 15)$$

[0108] 在等式 15 中, $\alpha_{\text{decay_min}}$ 表示平滑参数 α_{decay} 的最小值。在一些实现中, $\alpha_{\text{decay_min}}$ 可以为 0.9 或更大, 例如, 0.90、0.91、0.92、0.93、0.94、0.95、0.96、0.97、0.98、0.99、0.991 等。在一些实现中, 可以如下确定 $\alpha_{\text{decay_min}}$:

$$[0109] \quad \alpha_{\text{decay_min}} = 1 - \frac{1}{T_{\text{decay_min}} / T_{\text{packet}}} \quad (\text{等式 16})$$

[0110] 在等式 16 中, $T_{\text{decay_min}}$ 表示“衰减平滑”的最小时间间隔。换言之, $T_{\text{decay_min}}$ 表示在其期间平滑参数被设置成衰减平滑参数 α_{decay} 的最小时间间隔。在一些实现中, $T_{\text{decay_min}}$ 可以在几十秒或几秒的范围内, 例如, 0.1s、0.2s、0.3s、0.4s、0.5s、0.6s、0.7s、0.8s、0.9s、1.0s、1.1s、1.2s、1.3s、1.4s、1.5s、1.6s、1.7s、1.8s、1.9s、2.0s、2.1s、2.2s、2.3s、2.4s、2.5s 等。

[0111] 然而, 前述等式和参数值仅是不对称的时变的平滑参数的示例。在其他实现中, 例如, 最小起音和 / 或衰减平滑参数可以不是用于应用平滑参数的时间常数的函数。例如, 最小起音和 / 或衰减平滑参数可以是预先确定的恒定值。类似地, 最大平滑参数值可以是预先确定的常数。

[0112] 在一些实现中, 抖动缓冲器大小估计算法可以基于 CDF 并且基于另外的函数。这种实现可以是有利的, 以便更接近地近似于已经对于特定网络观察到的网络抖动行为。

[0113] 图 3 是示出 WiFi 网络中的分组延迟变化的示例的曲线图。在图 3 中, y 轴表示分组延迟变化, x 轴表示样本索引。样本索引例如可以对应于在特定时刻如电话会议开始的时刻之后所接收的分组的数量。

[0114] 在图 3 中可以观察到 PDV 的至少两个一般特征。一个特征是确定性分量 305, 其也可以称为周期性或准周期性分量。在该示例中, 确定性分量 305 包括高达近似 0.7 秒的延迟尖峰, 该延迟尖峰出现在近似 3000 个样本的样本索引间隔处。图 3 中的 PDV 曲线图的另一个特征是非确定性分量 310, 其也可以称为随机或准随机分量。在该示例中, 非确定性分量 310 包括具有比确定性分量 305 的幅度显著较低的幅度的延迟变化。然而, 非确定性分量 310 的延迟尖峰比确定性分量 305 的延迟尖峰更频繁得多地出现。

[0115] 图 3 中的 PDV 曲线图可以被视为包括确定性分量和非确定性分量两者的函数或时间序列。如上面所指出的, 在一些实现中, 抖动缓冲器大小估计算法可以基于 CDF 并且基于另外的函数。可以通过例如相加或通过另外的数学运算如相乘或卷积来组合这两个函数。

[0116] 根据一些实现, 抖动缓冲器大小估计算法可以基于 CDF 并且基于第二分布函数。一个分布函数可以表示抖动缓冲器大小估计算法的确定性分量, 另一分布函数可以表示抖动缓冲器大小估计算法的非确定性分量。在一些实现中, 第二分布函数可以是第二 CDF。

[0117] 或者, 抖动缓冲器大小估计算法可以基于 CDF 以及基于参数化函数如周期性或准周期性函数。参数化函数可以表示抖动缓冲器大小估计算法的确定性分量, CDF 可以表示抖动缓冲器大小估计算法的非确定性分量。

[0118] 图 4 是提供能够实现本公开内容的各个方面的设备的部件示例的框图。设备 400 例如可以是通信终端、服务器等(或可以是其一部分)。在一些示例中, 该设备可以以另外的装置的部件来实现。例如, 在一些实现中, 设备 400 可以是线路卡。

[0119] 在该示例中, 设备 400 包括接口系统 405、存储器系统 410 和逻辑系统 415。逻辑

系统 415 和 / 或存储器系统 410 能够在存储器系统 410 中建立一个或多个抖动缓冲器。接口系统 405 可以包括网络接口、逻辑系统与存储器系统之间的接口和 / 或外部装置接口 (如通用串行总线(USB)接口)。逻辑系统 415 例如可以包括通用单芯片或多芯片处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA) 或其他可编程逻辑器件、离散门或晶体管逻辑、和 / 或离散硬件部件。

[0120] 在该示例中,逻辑系统 415 能够经由接口系统 405 接收基于至少一个在先电话会议的遗留网络抖动信息。遗留网络抖动信息可以包括分组延迟变化(PDV)信息和 / 或到达间隔时间(IAT)信息。逻辑系统 415 能够至少部分地基于遗留网络抖动信息确定参数化模型的至少一个参数。逻辑系统 415 能够至少部分地基于参数化模型初始化抖动缓冲器大小估计算法。抖动缓冲器大小估计算法可以至少部分地基于网络抖动参数的累积分布函数(CDF)。逻辑系统 415 能够在电话会议期间根据抖动缓冲器大小估计算法控制抖动缓冲器大小。

[0121] 在一些实现中,控制抖动缓冲器大小可以涉及针对起音过程和衰减过程应用不对称的平滑参数。在一些示例中,不对称的平滑参数可以是时变的。逻辑系统能够确定分组延迟变化(PDV) 或其他网络抖动参数。在一些实现中,针对起音过程和衰减过程应用不对称的平滑参数可以涉及:如果 PDV (或其他网络抖动参数)大于当前抖动缓冲器大小则应用起音平滑参数。在一些示例中,针对起音过程和衰减过程应用不对称的平滑参数可以涉及:如果 PDV 不大于当前抖动缓冲器大小则应用衰减平滑参数。

[0122] 图 5 是提供音频处理设备的部件示例的框图。在该示例中,装置 500 包括接收系统 405。接口系统 405 可以包括网络接口如无线网络接口。替代地或附加地,接口系统 405 可以包括通用串行总线(USB) 接口或另外的这种接口。

[0123] 装置 500 包括逻辑系统 415。逻辑系统 415 可以包括处理器如通用单芯片或多芯片处理器。逻辑系统 415 可以包括数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA) 或其他可编程逻辑器件、离散门或晶体管逻辑、或离散硬件部件、或其组合。逻辑系统 415 可以被配置成控制装置 500 的其他部件。例如,逻辑系统 415 能够控制存储器系统 410 中的一个或多个抖动缓冲器的大小。尽管图 5 中未示出装置 500 的部件之间的接口,逻辑系统 415 可以被配置有用于与其他部件通信的接口。其他部件可以适当地被配置成或可以不被配置成彼此通信。

[0124] 逻辑系统 415 可以被配置成执行音频数据分析和抖动缓冲器控制功能,包括但不限于本文中所描述的功能。在一些这种实现中,逻辑系统 415 可以被配置成(至少部分地)根据一个或多个非暂态介质中存储的软件进行操作。非暂态介质可以包括与逻辑系统 415 相关联的存储器如随机存储存储器(RAM) 和 / 或只读存储器(ROM)。非暂态介质可以包括存储器系统 410 的存储器。存储器系统 410 可以包括一个或多个适当类型的非暂态存储介质,如闪存、硬盘驱动器等。

[0125] 逻辑系统 415 可以被配置成经由接口系统 405 接收多帧编码音频数据,并且对编码音频数据进行解码。替代地或附加地,逻辑系统 415 可以被配置成经由存储器系统 410 与逻辑系统 415 之间的接口接收多帧编码音频数据。逻辑系统 415 可以被配置成根据解码音频数据控制扬声器(或多个扬声器) 520。

[0126] 依赖于装置 500 的表现形式,显示系统 530 可以包括一种或更多种适当类型的显

示器。例如,显示系统 530 可以包括液晶显示器、等离子显示器、双稳态显示器等。

[0127] 用户输入系统 535 可以包括被配置成接收来自用户的输入的一个或更多个装置。在一些实现中,用户输入系统 535 可以包括重叠显示系统 530 的显示的触摸屏。用户输入系统 535 可以包括鼠标、跟踪球、姿势检测系统、操纵杆、在显示系统 530 上显示的一个或更多个 GUI 和 / 或菜单、按钮、键盘、开关等。在一些实现中,用户输入系统 535 可以包括麦克风 525 :用户可以经由麦克风 525 针对装置 500 提供语音命令。逻辑系统可以被配置用于语音识别以及用于根据这种语音命令控制装置 500 的至少一些操作。

[0128] 电力系统 540 可以包括一个或更多个适当的能量存储装置如镍镉电池或锂离子电池。电力系统 540 可以被配置成从电气插座接收电力。

[0129] 本公开内容中所描述的实现的的各种修改对本领域普通技术人员来说是明显的。在不偏离本公开内容的精神或范围的情况下,本文中所定义的总体原理也可以应用于其他实现。

[0130] 此外,所描述的实施方式可以以各种硬件、软件、固件等来实现。例如,本申请的各个方面可以被实施为系统、装置(例如蜂窝电话、便携式媒体播放器、个人计算机、服务器、电视机顶盒、数字录像机、或其他媒体播放器)、方法或计算机程序产品。因此,本申请的各方面可以采取以下形式:硬件实施方式、软件实施方式(包括固件、驻留软件、微代码等)或组合软件方面与硬件方面的实施方式。在本文中这种实施方式可以称为“电路”、“模块”或“系统”。此外,本申请的各个方面可以采取体现为一个或更多个非暂态介质的计算机程序产品的形式,该非暂态介质上体现有计算机可读程序代码。这种非暂态介质例如可以包括硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EEPROM 或闪存)、便携式紧凑磁盘只读存储器(CD-ROM)、光存储装置、磁存储装置、或前述各项的任何适当组合。因此,本公开内容的教示不意在限制图中示出的和 / 或本文所描述的实现,而是具有广泛的可应用性。从而,权利要求不意在限制本文中示出的实现,而是与符合本公开内容、本文中所公开的原理和新颖的特征的最宽范围一致。

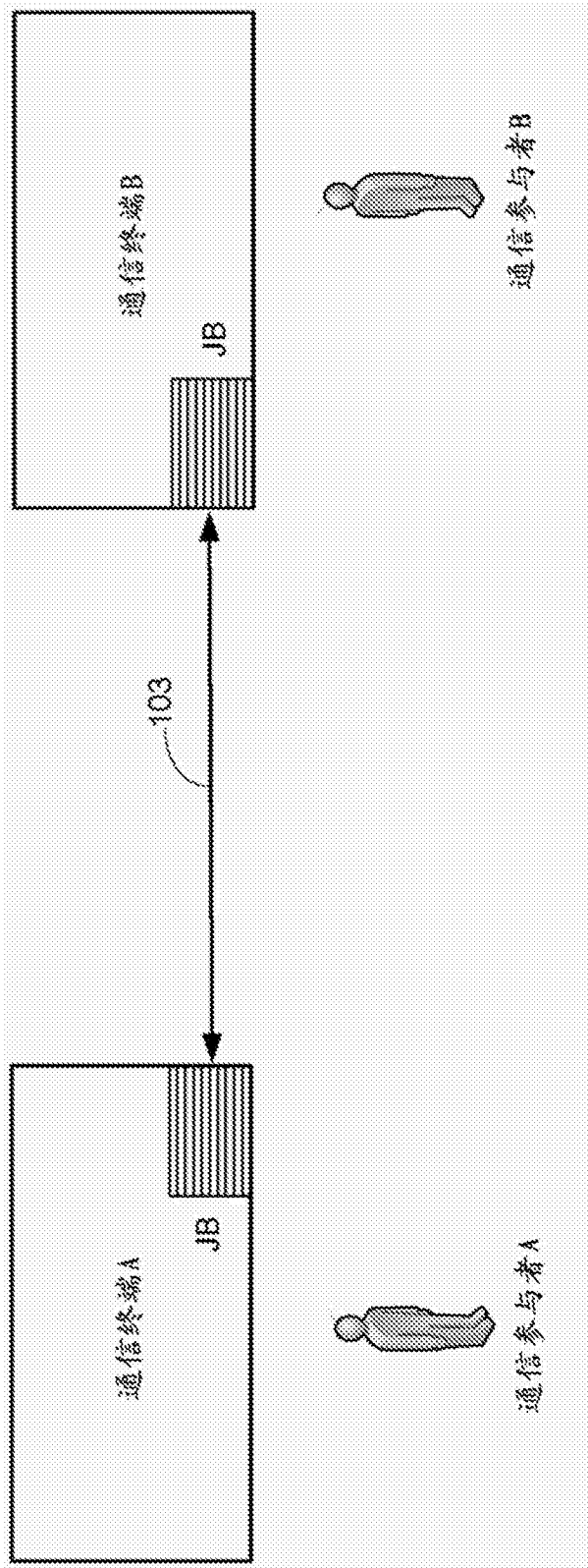


图 1A

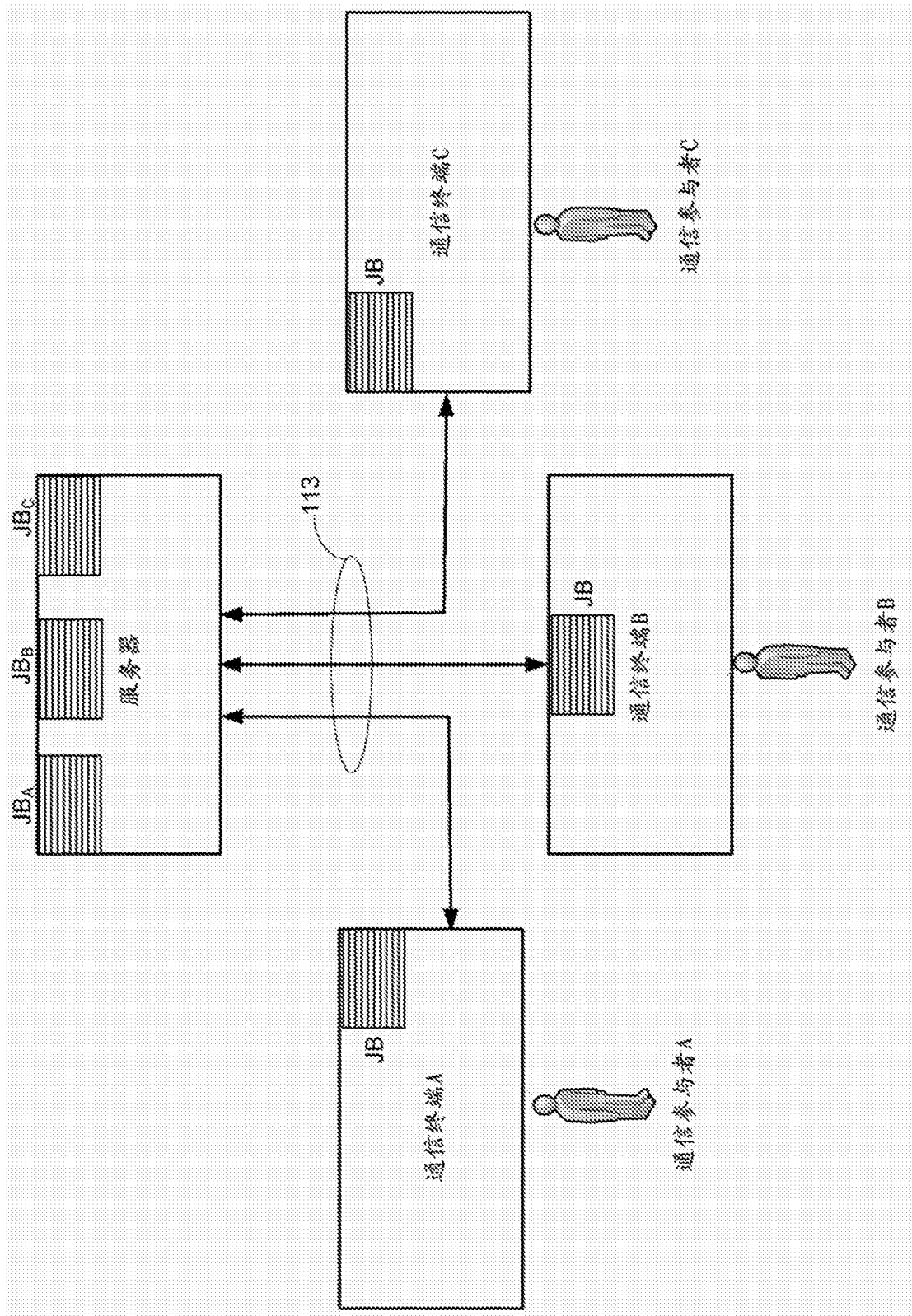


图 1B

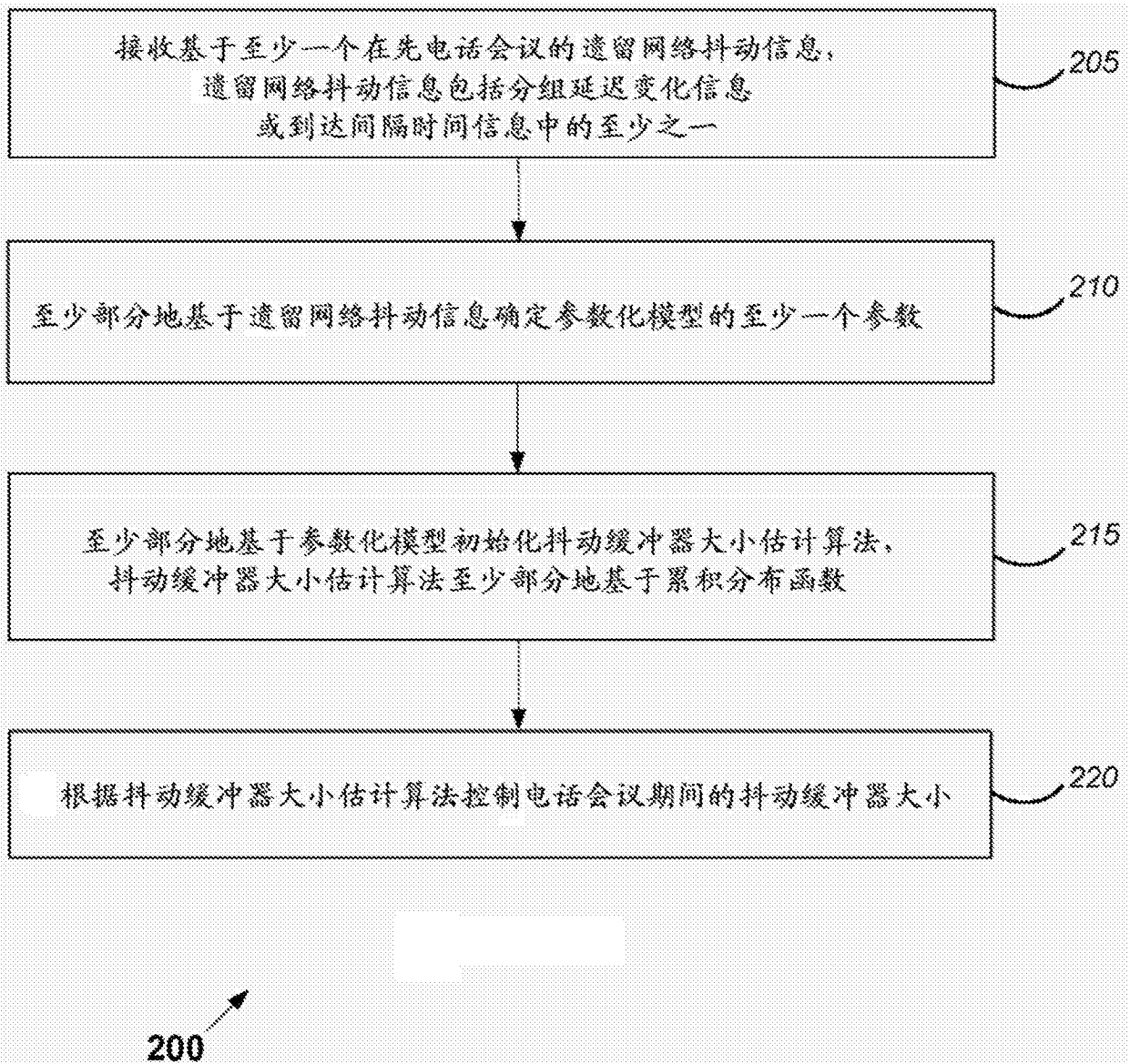


图 2

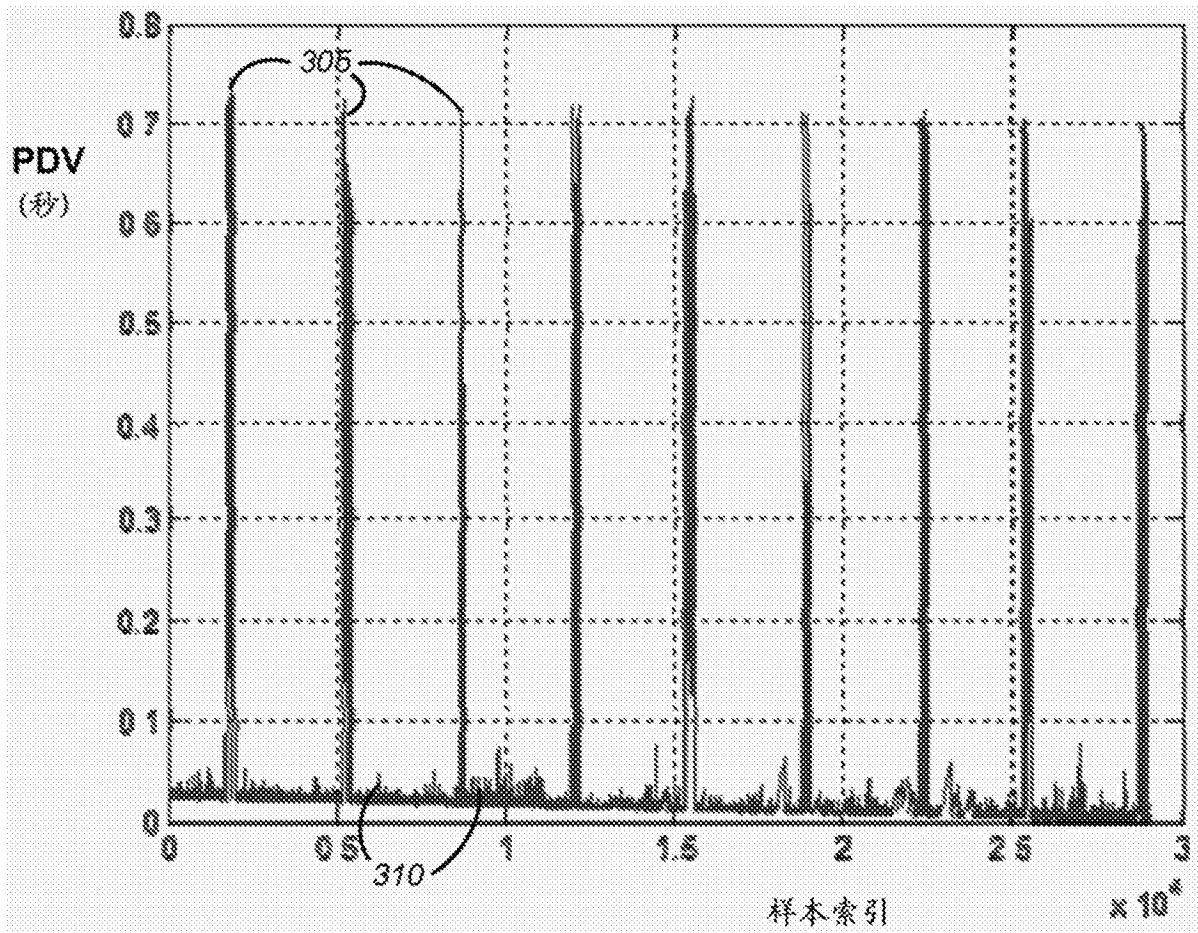


图 3

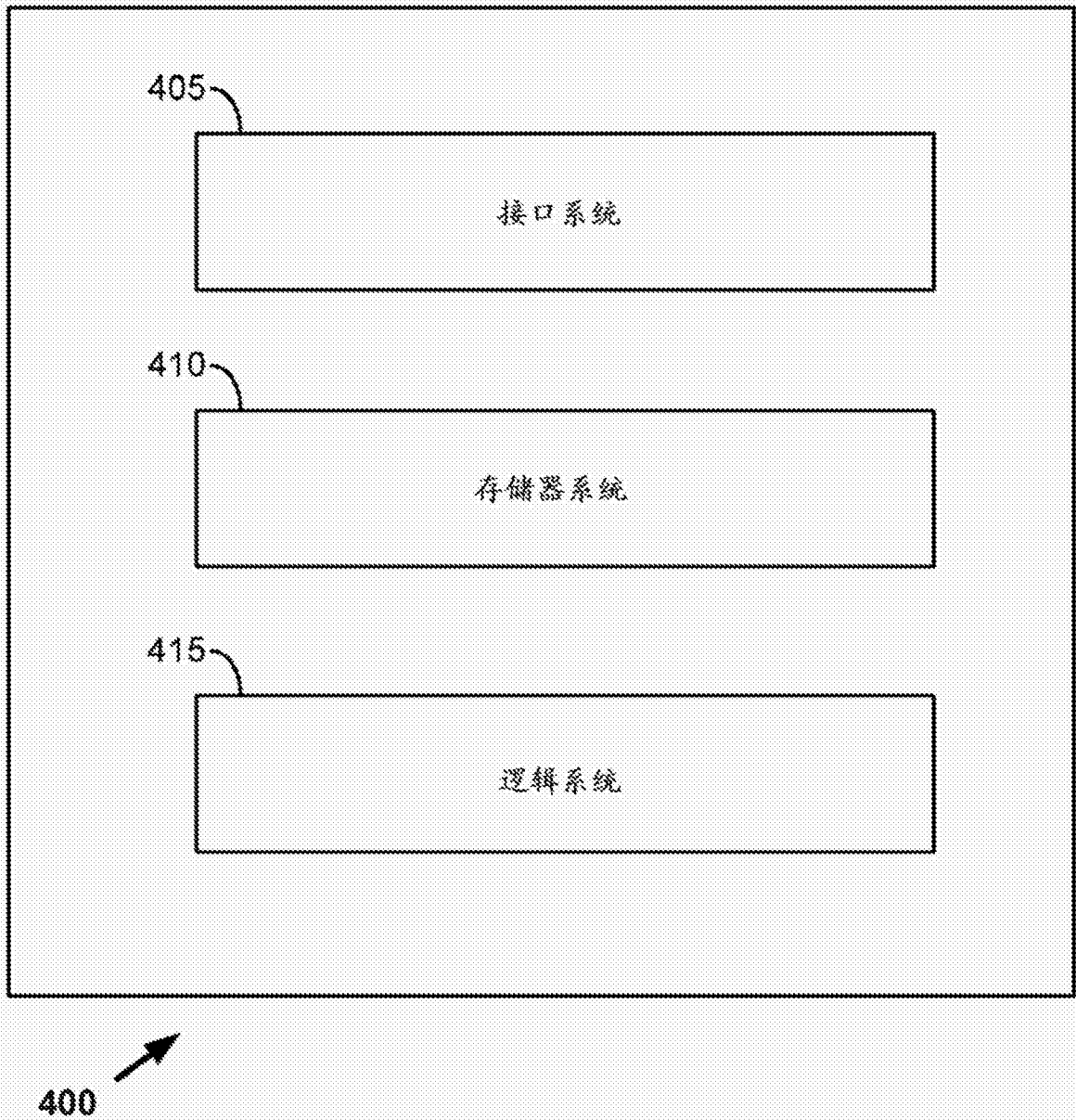


图 4

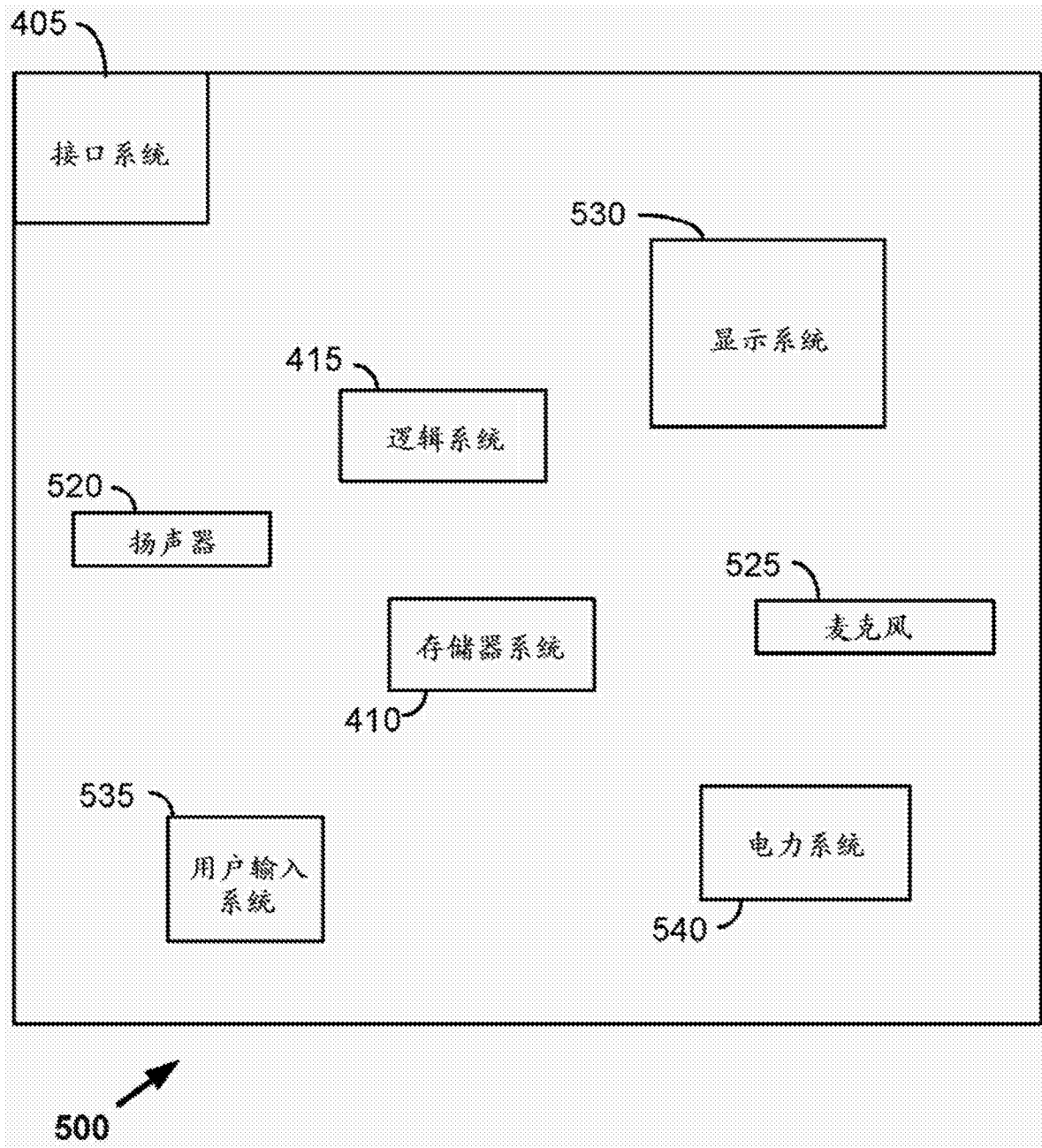


图 5