



(12)发明专利

(10)授权公告号 CN 103039051 B

(45)授权公告日 2017.04.19

(21)申请号 201180027592.0

(73)专利权人 斯凯普爱尔兰科技控股公司

(22)申请日 2011.06.03

地址 爱尔兰都柏林

(65)同一申请的已公布的文献号

(72)发明人 A.皮金 L.武克 N.阿布凯罗夫
V.特拉

申请公布号 CN 103039051 A

(74)专利代理机构 永新专利商标代理有限公司
72002

(43)申请公布日 2013.04.10

代理人 王英

(30)优先权数据

61/351711 2010.06.04 US

(51)Int.Cl.

H04L 29/06(2006.01)

(85)PCT国际申请进入国家阶段日

2012.12.04

(56)对比文件

US 6771594 B1,2004.08.03,

CN 1520694 A,2004.08.11,

US 2005021809 A1,2005.01.27,

(86)PCT国际申请的申请数据

PCT/US2011/039132 2011.06.03

审查员 谢正程

(87)PCT国际申请的公布数据

W02011/153475 EN 2011.12.08

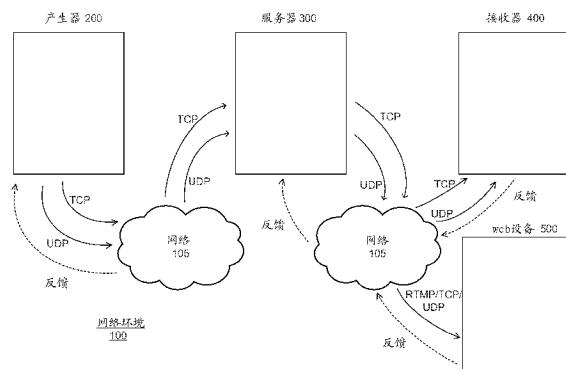
权利要求书4页 说明书16页 附图9页

(54)发明名称

服务器辅助的视频会话

(57)摘要

一种方法、计算机程序产品和系统提供了移动计算设备之间的实时双向音频-视频会话。低延时是通过以下操作来实现的：将来自给定音频-视频会话的音频和视频数据流进行分割，使用两个不同的传输协议在网络上发送分离的流，并在另一端处重新同步这些分离的流。基于指示网络或其他移动计算设备的可用带宽的反馈数据来调谐每个流的传输。服务器卸下否则将由移动计算设备处置的处理需求。双向会话可以由基于web的用户进行外部观察。系统通过移动计算设备端点和基于web的端点的全异集合并通过不同无线运营商网络基础结构而工作。



1. 一种由第一移动计算设备执行的计算机实现方法,用于在所述第一移动计算设备与第二移动计算设备之间创建双向音频-视频会话,所述方法包括:

记录音频数据流;

对所述音频数据流进行编码,所述音频数据流包括音频分组的流,每个音频分组包括音频时间戳;

接收视频数据流;

对所述视频数据流进行编码,所述视频数据流包括视频分组的流,每个视频分组包括视频时间戳,所述视频时间戳与对应的音频时间戳和同所述视频分组同时记录的音频分组相匹配;将用于所述第一移动计算设备和所述第二移动计算设备之间的双向音频-视频会话的音频数据流和视频数据流的处理卸下到集中式服务器,该卸下通过以下方式:

将音频数据流和视频数据流分割成分离的数据流以传送到集中式服务器;

通过第一传输协议,将所述音频数据流发送至集中式服务器;以及

通过第二传输协议,将所述视频数据流分离于所述音频数据流发送至所述集中式服务器,所述集中式服务器被配置为基于所述第二移动计算设备的特性来对所述音频数据流和所述视频数据流再编码,并将再编码的音频数据流和再编码的视频数据流二者传递到所述第二移动计算设备,用于基于相应的音频和视频时间戳进行同步。

2. 根据权利要求1所述的方法,进一步包括使得所述音频数据流和所述视频数据流的拷贝被存档。

3. 根据权利要求1所述的方法,其中,所述第二传输协议是可靠协议。

4. 根据权利要求2所述的方法,其中,所述音频数据流和所述视频数据流的拷贝以由所述第一移动计算设备编码的格式或以由集中式服务器再编码的格式存档。

5. 根据权利要求3所述的方法,其中,所述第二传输协议是传输控制协议。

6. 根据权利要求3所述的方法,其中,所述第一移动计算设备是web设备。

7. 一种由集中式服务器执行的计算机实现方法,用于在第一移动计算设备与第二移动计算设备之间创建双向音频-视频会话,所述方法包括:

在所述集中式服务器处,代表所述第一移动计算设备和所述第二移动计算设备执行操作,以处置双向音频-视频会话的实时数据流的处理任务,包括:

通过第一传输协议,从所述第一移动计算设备接收音频数据流,所述音频数据流是利用第一音频编解码器来编码的;

通过第二传输协议,从所述第一移动计算设备接收分离地传送的视频数据流,所述视频数据流是利用第一视频编解码器来编码的,所述第二传输协议不同于所述第一传输协议;

从所述第二移动计算设备接收编解码器数据,所述编解码器数据包括安装在所述第二移动计算设备上的编解码器的列表,所述编解码器的列表包括第二音频编解码器和第二视频编解码器;

确定所述编解码器的列表是否包括所述第一音频编解码器;

响应于确定所述编解码器的列表不包括所述第一音频编解码器,使用所述第二音频编解码器来对音频流进行代码转换;

确定所述编解码器的列表是否包括所述第一视频编解码器;

响应于确定所述编解码器的列表不包括所述第一视频编解码器，使用所述第二视频编解码器来对视频流进行代码转换；

由所述集中式服务器的比特率适配模块至少部分地基于所述第二移动计算设备的比特率限制来确定是否从所接收的视频数据流中丢弃一个或多个帧，从而使得所述第二移动计算设备的处理要求相对于不丢弃所述一个或多个帧降低；

通过第一网络协议，将所述音频数据流发送至所述第二移动计算设备；以及

响应于确定是否从所接收的视频数据流丢弃所述一个或多个帧，通过第二网络协议，将所述视频数据流分离于所述音频数据流发送至所述第二移动计算设备。

8. 根据权利要求7所述的方法，还包括：

从接收器接收反馈数据，所述反馈数据包括处理带宽、网络带宽和延时数据；

响应于接收到所述反馈数据，基于所述处理带宽、所述网络带宽和所述延时数据，调整所述视频数据流的比特率。

9. 根据权利要求7所述的方法，其中，所述第一移动计算设备和所述第二移动计算设备中的至少一个是web设备。

10. 一种由第二移动计算设备执行的计算机实现方法，用于在第一移动计算设备与所述第二移动计算设备之间创建双向音频-视频会话，所述方法包括：

通过第一传输协议，从集中式服务器接收音频数据流，所述音频数据流包括音频分组的流，每个分组包括音频时间戳；

通过第二传输协议，接收从集中式服务器分离地传送的视频数据流，所述视频数据流包括视频分组的流，每个分组包括视频时间戳，所述音频数据流和所述视频数据流在从所述第一移动计算设备到所述第二移动计算设备的整个通信中被维持为分割的流，并且由所述集中式服务器处理以卸下双向音频-视频会话的处理操作以降低延迟，所述集中式服务器被配置为基于所述第二移动计算设备的特性来对所述音频数据流和所述视频数据流再编码；

在缓冲器中对所述音频数据流和所述视频数据流进行缓冲；

将所述音频数据流与所述视频数据流进行同步，所述同步包括如果在所述缓冲器中存在匹配的视频时间戳，则将每个音频时间戳和音频分组与视频时间戳和视频分组相匹配；

通过音频子系统来输出同步后的音频数据流，所述第二移动计算设备的音频子系统被配置来对同步的音频数据流解码；以及

与输出同步后的音频数据流同时，通过视频子系统来输出同步后的视频数据流，所述第二移动计算设备的视频子系统被配置来对同步的视频数据流解码。

11. 根据权利要求10所述的方法，还包括：

监视所述缓冲器；

确定所述缓冲器是否已超过预定大小；以及

响应于确定所述缓冲器已超过所述预定大小，将反馈数据发送至集中式服务器，所述反馈数据包括处理带宽和延时数据。

12. 根据权利要求10所述的方法，其中，所述第二移动计算设备是web设备。

13. 一种第一移动计算设备处的装置，用于在所述第一移动计算设备与第二移动计算设备之间创建双向音频-视频会话，所述装置包括：

音频子系统,用于记录音频数据流以及用于对所述音频数据流进行编码,所述音频数据流包括音频分组的流;

视频子系统,用于接收视频数据流以及用于对所述视频数据流进行编码,所述视频数据流包括视频分组的流;

将所述双向音频-视频会话的音频数据流和视频数据流维持为分离的数据流并将分离的数据流的处理卸下到集中式服务器用于降低延迟的组件,包括:

恢复模块,用于将时间戳添加至所述音频分组和所述视频分组,每个音频分组包括音频时间戳,以及每个视频分组包括与对应的音频时间戳和同所述视频分组同时记录的音频分组相匹配的视频时间戳;

音频输出,用于通过第一传输协议将所述音频数据流作为所述分离的数据流中的一个发送至集中式服务器;以及

视频输出,用于通过第二传输协议将所述视频数据流分离于所述音频数据流发送至所述集中式服务器,所述集中式服务器被配置为基于所述第二移动计算设备的特性来对所述音频数据流或所述视频数据流进行再编码,将所述音频数据流和所述视频数据流二者处理为分离的数据流,并且将所述分离的数据流传递至所述第二移动计算设备以用于基于相应的音频和视频时间戳进行同步。

14. 根据权利要求13所述的装置,其中,所述第一传输协议是数据报协议。

15. 根据权利要求13所述的装置,其中,所述第二传输协议是可靠协议。

16. 根据权利要求14所述的装置,其中,所述第一传输协议是用户数据报协议。

17. 根据权利要求15所述的装置,其中,所述第二传输协议是传输控制协议。

18. 一种集中式服务器处的装置,用于在第一移动计算设备与第二移动计算设备之间创建双向视频-音频会话,所述装置包括:

代表所述第一移动计算设备和所述第二移动计算设备执行双向音频-视频会话的操作的装置,包括:

用于通过第一传输协议从所述第一移动计算设备接收音频数据流的装置,所述音频数据流是利用第一音频编解码器来编码的;

用于通过第二传输协议从所述第一移动计算设备接收视频数据流的装置,所述视频数据流是利用第一视频编解码器来编码的;

用于从所述第二移动计算设备接收编解码器数据的装置,所述编解码器数据包括安装在所述第二移动计算设备上的编解码器的列表,所述编解码器的列表包括第二音频编解码器和第二视频编解码器;

用于确定所述编解码器的列表是否包括所述第一音频编解码器的装置;

用于响应于确定所述编解码器的列表不包括所述第一音频编解码器,使用所述第二音频编解码器来对音频流进行代码转换的装置;

用于确定所述编解码器的列表是否包括所述第一视频编解码器的装置,响应于确定所述编解码器的列表不包括所述第一视频编解码器;

用于使用所述第二视频编解码器来对视频流进行代码转换的装置;

用于至少部分地基于所述第二移动计算设备的比特率限制来确定是否从所接收的视频数据流中丢弃一个或多个帧,从而使得所述第二移动计算设备的处理要求相对于不丢弃

所述一个或多个帧降低的装置；

用于通过第一网络协议将所述音频数据流发送至所述第二移动计算设备的装置；以及
用于响应于确定是否从所接收的视频数据流丢弃所述一个或多个帧，通过第二网络协议将所述视频数据流分离于所述音频数据流发送至所述第二移动计算设备的装置。

19. 根据权利要求18所述的装置，还包括：

用于从所述第二移动计算设备接收反馈数据的装置，所述反馈数据包括处理带宽、网络带宽和延时数据；

用于响应于接收到所述反馈数据，基于所述处理带宽、所述网络带宽和所述延时数据，调整所述视频数据流的比特率的装置。

20. 一种第二移动计算设备处的装置，用于在第一移动计算设备与所述第二移动计算设备之间创建双向音频-视频会话，所述装置包括：

音频输入，用于通过第一传输协议，接收音频数据流，所述音频数据流包括音频分组的流，每个分组包括音频时间戳；

视频输入，用于通过第二传输协议，接收从集中式服务器分离地传送的视频数据流，所述视频数据流包括视频分组的流，每个分组包括视频时间戳，所述音频数据流和所述视频数据流被接收为分割的流，其在从所述第一移动计算设备到所述第二移动计算设备的整个通信中被维持为分割的流，并且由所述集中式服务器处理以卸下双向音频-视频会话的处理操作以降低延迟，所述集中式服务器被配置为基于所述第二移动计算设备的特性来对所述音频数据流和所述视频数据流再编码；

恢复模块，用于在缓冲器中对所述音频数据流和所述视频数据流进行缓冲并将所述音频数据流与所述视频数据流进行同步，所述同步包括如果在所述缓冲器中存在匹配的视频时间戳，则将每个音频时间戳和音频分组与视频时间戳和视频分组相匹配；

音频子系统，用于输出同步后的音频数据流，所述第二移动计算设备的音频子系统被配置来对同步的音频数据流解码；以及

视频子系统，用于与输出同步后的音频数据流同时地输出同步后的视频数据流，所述第二移动计算设备的视频子系统被配置来对同步的视频数据流解码。

21. 根据权利要求20所述的装置，还包括：

所述恢复模块监视所述缓冲器并确定所述缓冲器是否已超过预定大小；以及响应于确定所述缓冲器已超过所述预定大小，将反馈数据发送至集中式服务器，所述反馈数据包括处理带宽和延时数据。

服务器辅助的视频会话

[0001] 相关申请

[0002] 本申请要求于2010年6月4日提交的美国临时申请No. 61/351,711的权益。上述申请的全部教导以引用的方式并入于此。

技术领域

[0003] 本公开总体涉及移动计算设备网络上的数据传输的领域,更具体地涉及在移动计算设备网络上对音频和视频数据的实时传输。

背景技术

[0004] 具有记录、传输和显示音频和视频数据的能力的移动计算设备是公知的。这种计算设备还包括用于经由通信网络与世界进行通信的网络接口。从移动计算设备流传输音频和视频是已知的,但是难以在多个网络上实现支持异构设备的大规模服务。存在与从移动设备流传输视频和音频相关联的三个问题。网络延时、设备电池寿命和设备处理限制均使得降低质量并增加音频和视频实时流的延迟。

[0005] 关于网络延时,全异的移动和web端点导致移动设备网络具有波动的带宽。该问题预期持续,原因是:具有视频记录能力的移动和web设备端点的数目增加;以及无线运营商网络扩充至具有有限网络基础结构的国家中。此外,甚至在一个无线运营商网络内,对每个移动计算设备来说可用的带宽也作为时间以及移动计算设备位置的函数而波动。因此,无线接入的固有特性在于其为前后不一致的。

[0006] 关于设备处理和电池限制,众所周知,移动计算设备具有有限的电池寿命和有限的处理能力。在移动设备上对音频和视频数据的记录和查看是处理器和电池密集型的。在网络上对该数据进行流传输需要附加的处理和电池功率。此外,移动计算设备为了改进传输质量或使传输稳定而进行的处理越多,则电池寿命越短并且移动计算设备所发送的流的延时越高。

[0007] 目前的解决方案不足以实现移动计算设备中的音频和视频流传输的延时的实时水平。

发明内容

[0008] 本发明提供了对移动计算设备之间的音频-视频会话的实时双向流传输。在一个方面,本发明针对在移动计算设备之间进行流传输时延时的实时“会话”水平优化了流传输。本发明还提供了用于最小化为了提供实时流传输所需的处理和电池功率的量的方式。“会话”或实时延时表示在对不间断会话来说足够的视频和音频质量的情况下约500毫秒或更小的端到端视频和音频延时,使得双向会话的两侧均可以实时交互。

[0009] 会话延时是部分地通过以下操作来实现的:将来自给定音频-视频会话的音频和视频数据流进行分割,使用两个不同的传输协议在网络上发送分离的流,并在另一端处重新同步这些分离的流以提供双向会话体验。用于不同流的协议是鉴于从用户体验的观点来

看可接受的每个流的质量和连续性的水平来选择的。本发明针对音频传输利用“数据报”(例如,没有对传输的肯定应答)传输协议并针对视频传输利用“可靠的”(或基于肯定应答的)传输协议,以最小化流的传输的延时。在一个实施例中,针对音频流使用UDP,并针对视频流使用TCP。针对音频流选择UDP以最小化传输的延时;针对视频流选择TCP以确保适当地传输视频数据的特定分组。此外,基于反馈数据来调谐每个流的传输参数,以基于网络或其他移动计算设备的可用带宽来调整流的比特率。

[0010] 本发明利用集中式服务器系统来卸下否则将由移动计算设备处置的处理需求。这进而降低了移动计算设备电池消耗。在集中式服务器系统上利用与移动计算设备相比显著更多的带宽和处理能力执行处理活动降低了这些任务中的延时量。相比于不具有借以依靠的基于服务器的处理源的任何所提出的基于对等的视频和音频聊天解决方案,这种基于服务器的解决方案将是特别有利的。

[0011] 本发明能够通过移动计算设备端点和基于web的端点的全异集合以及通过不同无线运营商网络基础结构来实现移动计算设备之间的实时双向音频-视频会话。此外,本发明允许从基于web的用户对双向会话进行外部观察。即,web端点的任何数目的用户可以查看两个远程定位的移动设备之间的音频-视频会话。

[0012] 本发明具有各种用于移动设备和集中式服务器系统的各种方法以及移动设备装置、集中式服务器系统和计算机程序产品中的实施例。

[0013] 根据本发明的一方面,提供了一种由第一移动计算设备执行的计算机实现方法,用于在所述第一移动计算设备与第二移动计算设备之间创建双向音频-视频会话,所述方法包括:记录音频数据流;对所述音频数据流进行编码,所述音频数据流包括音频分组的流,每个音频分组包括音频时间戳;接收视频数据流;对所述视频数据流进行编码,所述视频数据流包括视频分组的流,每个视频分组包括视频时间戳,所述视频时间戳与对应的音频时间戳和同所述视频分组同时记录的音频分组相匹配;通过第一传输协议,将所述音频数据流发送至集中式服务器;以及通过第二传输协议,将所述视频数据流发送至所述集中式服务器。

[0014] 根据本发明的另一方面,提供了一种由集中式服务器执行的计算机实现方法,用于在第一移动计算设备与第二移动计算设备之间创建双向音频-视频会话,所述方法包括:通过第一传输协议,从所述第一移动计算设备接收音频数据流,所述音频数据流是利用第一音频编解码器来编码的;通过第二传输协议,从所述第一移动计算设备接收视频数据流,所述视频数据流是利用第一视频编解码器来编码的;从所述第二移动计算设备接收编解码器数据,所述编解码器数据包括安装在所述第二移动计算设备上的编解码器的列表,所述编解码器的列表包括第二音频编解码器和第二视频编解码器;确定所述编解码器的列表是否包括所述第一音频编解码器;响应于确定所述编解码器的列表不包括所述第一音频编解码器,使用所述第二音频编解码器来对音频流进行代码转换;确定所述编解码器的列表是否包括所述第一视频编解码器;响应于确定所述编解码器的列表不包括所述第一视频编解码器,使用所述第二视频编解码器来对视频流进行代码转换;通过第一网络协议,将所述音频数据流发送至所述第二移动计算设备;以及通过第二网络协议,将所述视频数据流发送至所述第二移动计算设备。

[0015] 根据本发明的另一方面,提供了一种由集中式服务器执行的计算机实现方法,用

于在第一移动计算设备与第二移动计算设备之间创建双向音频-视频会话，所述方法包括：通过第一传输协议，从所述第一移动计算设备接收音频数据流，所述音频数据流包括分组的流和音频比特率；通过第二传输协议，从所述第一移动计算设备接收视频数据流，所述视频数据流包括分组的流和视频比特率；从所述第二移动计算设备接收反馈数据，所述反馈数据包括网络带宽和处理带宽；确定所述音频比特率与所述视频比特率之和是否超过所述网络带宽或所述处理带宽；响应于确定所述音频比特率与所述视频比特率之和超过所述网络带宽或所述处理带宽，将视频数据流的比特率降至低于所述网络带宽和所述处理带宽；通过第一网络协议，将所述音频数据流发送至所述第二移动计算设备；以及通过第二网络协议，将所述视频数据流发送至移动计算设备。

[0016] 根据本发明的另一方面，提供了一种由第二移动计算设备执行的计算机实现方法，用于在第一移动计算设备与所述第二移动计算设备之间创建双向音频-视频会话，所述方法包括：通过第一传输协议，接收音频数据流，所述音频数据流包括音频分组的流，每个分组包括音频时间戳；通过第二传输协议，接收视频数据流，所述视频数据流包括视频分组的流，每个分组包括视频时间戳；在缓冲器中对所述音频数据流和所述视频数据流进行缓冲；将所述音频数据流与所述视频数据流进行同步，所述同步包括如果在所述缓冲器中存在匹配的视频时间戳，则将每个音频时间戳和音频分组与视频时间戳和视频分组相匹配；通过音频子系统来输出同步后的音频数据流；以及与输出同步后的音频数据流同时，通过视频子系统来输出同步后的视频数据流。

[0017] 根据本发明的另一方面，提供了一种第一移动计算设备处的装置，用于在所述第一移动计算设备与第二移动计算设备之间创建双向音频-视频会话，所述装置包括：音频子系统，用于记录音频数据流以及用于对所述音频数据流进行编码，所述音频数据流包括音频分组的流；视频子系统，用于接收视频数据流以及用于对所述视频数据流进行编码，所述视频数据流包括视频分组的流；恢复模块，用于将时间戳添加至所述音频分组和所述视频分组，每个音频分组包括音频时间戳，以及每个视频分组包括与对应的音频时间戳和同所述视频分组同时记录的音频分组相匹配的视频时间戳；音频输出，用于通过第一传输协议将所述音频数据流发送至集中式服务器；以及视频输出，用于通过第二传输协议将所述视频数据流发送至所述集中式服务器。

[0018] 根据本发明的另一方面，提供了一种集中式服务器处的装置，用于在第一移动计算设备与第二移动计算设备之间创建双向视频-音频会话，所述装置包括：用于通过第一传输协议从所述第一移动计算设备接收音频数据流的装置，所述音频数据流是利用第一音频编解码器来编码的；用于通过第二传输协议从所述第一移动计算设备接收视频数据流的装置，所述视频数据流是利用第一视频编解码器来编码的；用于从所述第二移动计算设备接收编解码器数据的装置，所述编解码器数据包括安装在所述第二移动计算设备上的编解码器的列表，所述编解码器的列表包括第二音频编解码器和第二视频编解码器；用于确定所述编解码器的列表是否包括所述第一音频编解码器的装置；用于响应于确定所述编解码器的列表不包括所述第一音频编解码器，使用所述第二音频编解码器来对音频流进行代码转换的装置；用于确定所述编解码器的列表是否包括所述第一视频编解码器的装置，响应于确定所述编解码器的列表不包括所述第一视频编解码器；用于使用所述第二视频编解码器来对视频流进行代码转换的装置；用于通过第一网络协议将所述音频数据流发送至所述第

二移动计算设备的装置；以及用于通过第二网络协议将所述视频数据流发送至所述第二移动计算设备的装置。

[0019] 根据本发明的另一方面，提供了一种集中式服务器处的装置，用于在第一移动计算设备与第二移动计算设备之间创建双向音频-视频会话，所述装置包括：用于通过第一传输协议从所述第一移动计算设备接收音频数据流的装置，所述音频数据流包括分组的流和音频比特率；用于通过第二传输协议从所述第一移动计算设备接收视频数据流的装置，所述视频数据流包括分组的流和视频比特率；用于从所述第二移动计算设备接收反馈数据的装置，所述反馈数据包括网络带宽和处理带宽；用于确定所述音频比特率与所述视频比特率之和是否超过所述网络带宽或所述处理带宽的装置；用于响应于确定所述音频比特率与所述视频比特率之和超过所述网络带宽或所述处理带宽，将视频数据流的比特率降至低于所述网络带宽和所述处理带宽的装置；用于通过第一网络协议将所述音频数据流发送至所述第二移动计算设备的装置；以及用于通过第二网络协议将所述视频数据流发送至移动计算设备的装置。

[0020] 根据本发明的另一方面，提供了一种第二移动计算设备处的装置，用于在第一移动计算设备与所述第二移动计算设备之间创建双向音频-视频会话，所述装置包括：音频输入，用于通过第一传输协议，接收音频数据流，所述音频数据流包括音频分组的流，每个分组包括音频时间戳；视频输入，用于通过第二传输协议，接收视频数据流，所述视频数据流包括视频分组的流，每个分组包括视频时间戳；恢复模块，用于在缓冲器中对所述音频数据流和所述视频数据流进行缓冲并将所述音频数据流与所述视频数据流进行同步，所述同步包括如果在所述缓冲器中存在匹配的视频时间戳，则将每个音频时间戳和音频分组与视频时间戳和视频分组相匹配；音频子系统，用于输出同步后的音频数据流；以及视频子系统，用于与输出同步后的音频数据流同时地输出同步后的视频数据流。

[0021] 说明书中描述的特征和优势不是所有都包括进来，特别地，鉴于附图、说明书和权利要求，许多附加特征和优势将对本领域技术人员来说显而易见。此外，应当注意，说明书中使用的语言主要是出于可读性和指导性目的而选择的，并可能未被选择以划定或限定所公开的主题。

附图说明

[0022] 图1示意了双向视频会话的网络环境和一侧的一个实施例。

[0023] 图2示意了被配置为关于产生器功能支持双向视频会话的移动计算设备的一个实施例。

[0024] 图3示意了被配置为对产生器与接收器之间的双向视频会话进行修改和中继的服务器设备的一个实施例。

[0025] 图4示意了被配置为关于接收器功能支持双向视频会话的移动计算设备的一个实施例。

[0026] 图5示意了被配置为与在两个移动计算设备之间进行的双向会话进行交互的web设备的一个实施例。

[0027] 图6a示意了包括在第一移动计算设备的屏幕上示出的内容的示例双向会话的一个实施例。

[0028] 图6b示意了包括在第二移动计算设备的屏幕上示出的内容的示例双向会话的一个实施例。

[0029] 图7示意了被配置为支持双向视频会话的设备的硬件基础结构的一个实施例。

[0030] 图8示意了发起双向视频会话的过程的一个实施例。

[0031] 图9示意了用于利用基于文本的参与功能来观察双向视频会话的web设备的一个实施例。

[0032] 附图仅出于示意的目的示出了本发明的各个实施例。本领域技术人员将容易地从以下讨论中认识到,在不脱离这里描述的本发明的原理的前提下,可以采用这里示意的结构和方法的替换实施例。

具体实施方式

[0033] 系统级概述

[0034] 图1示意了被配置为支持双向音频-视频(或者更简单地,简称“视频”)会话的网络环境100的一个实施例。双向视频会话包括两个移动计算设备,这两个移动计算设备相互实时共享来自这两个设备的音频和视频数据。因此,第一移动计算设备将记录和发送将被发送和接收以在第二移动计算设备上查看和收听的音频和视频数据。同时,第二移动计算设备也将记录和发送将被发送和接收以在第一移动计算设备上查看和收听的音频和视频数据。双向会话具有音频和视频数据的两个集合的往返行程之间的足够低的延时,使得第一移动计算设备的用户可以体验到对用户所记录的内容作出反应的第二移动计算设备处的朋友的反应。

[0035] 因此,网络环境100中的每个移动计算设备同时充当音频和视频数据的产生器200和接收器400。每个移动计算设备被称为以全双工方式操作。图1示意了双向视频会话的一半。在图1中,充当产生器200角色的第一移动计算设备将其已记录的音频和视频数据发送至充当接收器400角色的第二移动计算设备。然而,这仅为整个双向会话的一半,这是由于同时,第二移动计算设备还充当产生器200,并且第一移动计算设备充当在第二移动计算设备处记录的数据的接收器。因此,双向会话功能对这两个设备来说是相同的。为了清楚,图1仅示意了双向会话的一半。

[0036] 产生器200通过网络105将数据发送至接收器400。产生器200和接收器400通过多个传输层网络协议进行通信。产生器200和接收器400使用用户数据报协议(UDP)来交换音频数据,并使用传输控制协议(TCP)来交换视频数据。产生器200被配置为发送UDP和TCP数据,并且接收器被配置为接收UDP和TCP数据。

[0037] 网络环境100还包括服务器系统300,以便促进以最低延时在最高可能质量下传输音频和视频数据,并最小化远程定位的移动计算设备上的处理器和电池压力。服务器系统300被配置为分别与接收器400和产生器200发送和接收UDP和TCP数据。即,服务器系统300在给定的接收器400和产生器200对之间的信号路径中进行操作,从每个移动设备接收两个流并将这些流中继至另一移动设备。服务器系统300还被配置为从接收器400接收反馈数据,该反馈数据用于调整从服务器系统300发送至接收器400的音频和视频流。

[0038] 可选地,网络环境100可以另外包括一个或多个web设备500,其中,这种设备的用户可以通过监视或参与两个移动设备之间的给定的视频会话,共享双向会话的体验。可替

换地,移动计算设备被设置为仅充当产生器200,以便用作单向广播定向的web设备500。在这些情况下,服务器系统300另外被配置为将音频和视频数据发送至web设备500。服务器系统300可以以如接收器设备400一样的UDP或TCP数据的形式将数据发送至web设备500。此外,服务器系统300可以在实时消息收发协议(RTMP)中发送音频和/或视频数据,以在web设备500的浏览器内的FLASH软件类型播放器中消耗。

[0039] 消耗来自单向或双向会话的流的web设备500另外被配置为通过提供音频或文本评论视频音频或文本录入接口来参与体验。可以与双向会话的参与者、单向广播的流传输者或者与体验该会话的任何其他web设备500共享该评论。

[0040] 产生器

[0041] 图2示意了在双向视频会话中充当产生器200的移动计算设备的一个实施例。产生器200被配置为:使用内置音频和视频记录组件来录音频和视频数据,对音频和视频数据执行最小处理量以将音频和视频记录封装为分组流,并在网络105上将音频和视频数据作为分组数据发送至服务器系统300。产生器200系统包括音频子系统205、视频子系统210、恢复模块215、网络层模块220、音频输出225和视频输出230。产生器200在音频子系统205处记录音频数据。产生器200在视频子系统210处记录视频数据。在恢复模块215中处理该数据。然后,通过音频225和视频230输出,将该数据发送至服务器系统300。

[0042] 音频子系统205被配置为与硬件输入接口和设备尽可能接近地捕获所记录的音频。音频子系统205被配置为与移动计算设备本机的多个组件进行交互。当发起会话时,音频子系统205与移动计算设备的操作系统进行交互,以发起对音频的记录。操作系统指示移动计算设备的麦克风设备记录在移动计算设备外部的环境中的音频。数字信号处理器将原始记录的音频信号从模拟格式转换为数字格式。操作系统结合音频编解码器进行操作,以将数字音频信号转换为具有补充编解码器的音频播放器(诸如,WINDOWS MEDIA PLAYER)可在计算设备上播放的编码比特流。

[0043] 一般地,音频子系统205在原始音频信号已被转换为符合特定编解码器的数字流之后立即接收所记录的音频信号。通过尽可能早地接收音频,音频子系统205通过与移动计算设备的操作系统相关联的处理软件,最小化与音频信号的非必要处理相关联的延时。此外,消除非必要处理减小了移动计算设备的处理器上的双向会话的压力,从而还降低了与维持双向会话流相关联的电池需求。

[0044] 音频子系统205借以接收所记录的音频的机制可以根据移动计算设备的类型而不同。例如,如果移动计算设备符合来自Apple, Inc的设备(诸如,iPHONE™)中规定的音频记录架构,则音频子系统从作为每个Apple, Inc设备的操作系统的一部分而可接入地包括的Audio Queue Services接口接收编码的录音音频。Audio Queue Services接口允许应用直接接入音频子系统,以直接从硬件自身接收音频,从而允许音频子系统205控制如何对音频进行转换和编码。一般地,音频队列是产生器400用以记录或播放音频的软件对象。音频队列进行以下工作:连接至移动设备的音频硬件;在记录期间管理设备的存储器;在需要时将编解码器用于压缩的音频格式;以及调解记录或回放。

[0045] 如果移动计算设备是WINDOWS MOBILE™计算设备,则音频子系统205通过WAVEIO接口来捕获音频。如果设备正在运行ANDROID™平台,则音频交互子系统使用AUDIORECORD本机代码系统或ANDROID处理器间(IPC)通信系统来获得音频。

[0046] 类似地,视频子系统被配置为与硬件输入电平尽可能接近地捕获所记录的视频。视频子系统被配置为与移动计算设备本机的多个组件进行交互。当发起会话时,视频子系统与移动计算设备的操作系统进行交互,以发起对视频的记录。操作系统指示视频摄像设备在移动计算设备外部的环境中记录视频。数字信号处理器将原始记录的视频信号从模拟格式转换为数字格式。操作系统结合视频编解码器进行操作,以将数字视频信号转换为视频播放器(诸如,WINDOWS MEDIA PLAYER)可在计算设备上播放的编码比特流。

[0047] 一般地,视频子系统210在原始视频信号已被转换为符合特定编解码器的数字流之后立即接收所记录的视频信号。通过尽可能早地接收视频,视频子系统210还去除了与关联于移动计算设备的操作系统的视频信号软件的非必要处理相关联的延时。同样,消除非必要处理减小了移动计算设备的处理器上的双向会话的压力,从而还降低了与维持双向会话流相关联的电池需求。

[0048] 恢复模块215同时接收编码后的音频和视频流二者,并准备这些音频和视频流以在网络105上发送至服务器系统300。当恢复模块215从音频205和视频210子系统接收到音频和视频数据流时,作为其同时记录的结果,在时间上将这些流彼此同步。有意地将音频和视频流保持分离。在本文中的一些点处,为了清楚,可以将其称为单个流,然而,始终分离地维持这两个流。这与在网络上经由单个传输协议将音频视频流一起传输的那些在先系统相反。

[0049] 恢复模块215在音频和视频流被发出之前对音频和视频流的每个帧(或分组)加时间戳。指派时间戳,以便在时间上同步音频和视频数据流。对应的音频和视频分组具有匹配的时间戳。因此,当回放具有匹配时间戳的音频和视频时,音频将与视频同步。将时间戳指派给音频和视频数据的每个分组,使得在数据的任何分组被丢弃或者以其他方式未被接收器400接收到的情况下,接收器将能够重新同步适当接收到的分组。

[0050] 还在丢失的视频数据流分组的恢复中使用时间戳。各个视频帧较大,并且因此,恢复模块210将这些视频帧分割为多于一个分组(或者进行分组化)。使用TCP发送视频数据是为了请求重传在产生器200、服务器300和接收器400之间的传输中丢失的视频分组。TCP请求所有丢失数据,与不请求任何丢失数据的其他传输协议(诸如RTP)相反。服务器300或接收器400针对丢失数据的请求可以是基于来自相同帧的其他所接收到的视频数据流分组来进行的。如果丢失了来自帧的分组,则使用来自适当接收到的其他分组的时间戳来请求丢失的分组。如果接收器400正在请求该数据,则在服务器300具有该数据的情况下从服务器300重新发送丢失的数据,否则,服务器300从产生器200重新请求分组以重新发送。这提高了恢复模块410在网络带宽(或信道容量)较小时恢复丢失的视频帧的能力。在音频的情况下,即使丢失了更早的音频分组,也可以对音频分组进行解码,这与如上讨论的视频帧的情况不同。

[0051] 恢复模块215还被配置为从网络层接收网络带宽信息。网络带宽信息包括同移动计算设备与服务器系统300之间的网络连接的强度或质量有关的信息。在一个实施例中。网络带宽信息包括移动计算设备所具有的可用于在网络上发送信息的千比特每秒(kbps)的总带宽。网络带宽信息可以包括音频带宽部分、视频带宽部分和反馈带宽,分别表示多少带宽可用于发送和接收音频数据、视频数据和反馈数据。可替换地,网络带宽信息可以包括对多少带宽可用于将数据发送至服务器系统300以及多少带宽可用于从服务器系统300接收

数据的度量。在双向会话中,分配可用于移动计算设备的带宽中的一些,以便除了仅发送数据外,还从其他移动计算设备接收数据。这是由于以下事实:在双向会话中,移动计算设备既充当产生器又充当接收器。

[0052] 恢复模块215使用网络带宽信息来确定将音频和视频数据发送至服务器系统300的比特率。在一个实施例中,恢复模块与音频子系统205和视频子系统210进行通信,以配置它们将对音频和视频进行记录和编码的比特率。在另一实施例中,音频205和视频210子系统基于移动计算设备的能力,在固定比特率下记录音频和视频。在这种情况下,恢复模块215基于可用带宽,动态地转换音频和视频流的比特率。由于该实施例需要将由产生器200执行的附加处理,因此优选地与具有更快的处理器和/或延长的电池寿命的移动设备一起使用该实施例。

[0053] 音频输出模块225或者直接从音频子系统205或者从恢复模块215接收所记录和编码的音频数据流。音频输出模块225在网络105上将音频率传输至服务器系统300。音频输出模块225通过利用网络层220来传输数据。

[0054] 如上所述,音频输出模块225使用UDP来传输音频流。UDP是出于多种原因而使用的。UDP具有极低的数据传输延时。这使其对传输其中非常低延时是重要的音频流来说是理想的。UDP通过不需要来自正接收UDP传输的目标设备(在这种情况下,其为服务器系统300)的肯定应答(ACK),部分地实现了其低延时。UDP是不需要所传输的分组的接收的肯定应答的一类传输协议(通称为“数据报”协议)之一。然而,不具有ACK的后果是:将在传输中丢失一些分组。在UDP下,从不重传这些分组技术。音频数据的这种丢失被视为可接受以交换由UDP提供的更好延时。在一个实施例中,丢失至多20毫秒的音频数据被视为可接受,这不显著干扰用户的会话体验,因为即使一些音频数据丢失,用户仍将理解语音传输。

[0055] 视频输出模块230或者直接从视频子系统210或者从恢复模块215接收所记录和编码的视频数据流。视频输出模块在网络105上将视频流传输至服务器系统300。视频输出模块通过利用网络层220来传输数据。

[0056] 如上所述,视频输出模块230使用TCP来传输视频流。相对于UDP,使用TCP是由于TCP利用ACK信号来确保TCP数据的每个帧被网络另一侧的接收者接收。TCP是需要所传输的分组的接收的肯定应答的一类传输协议(通称为“可靠”协议)之一。期望每个帧已到达接收者的肯定应答,这是由于许多视频编码机制区分被发送至接收者的视频数据的帧的类型。这些编码技术涉及使用在不参考任何其他帧的情况下编码的帧内编码的帧(典型地称为“I帧”)以及从一个或多个I帧预测编码的预测帧和可选地其他预测帧而区别地编码的视频数据。(预测帧典型地称为“P帧”。为了简明的目的,“P帧”将用于指代任何类型的预测帧,包括前向预测以及双向预测的帧。)I帧需要更多数据,这是由于I帧包含单个视频帧的所有数据并被相对不频繁地发送。然后,基于与I帧和相邻P帧的区别来对P帧进行编码。这允许利用比I帧少得多的数据来对P帧进行编码。

[0057] 因此,如果在传输期间丢失了I帧的一个或多个分组,则直到下一I帧之前的所有后续P帧可能不可被移动设备解码,导致视频播放器中丢弃的帧以及回放质量的可见下降,诸如回放中的显著滞后或停顿。相比之下,如果丢失了P帧的分组,则这种下降很可能不是视觉上显著的。因此,视频输出模块使用TCP来确保视频数据流的连贯性,使得所有帧被服务器系统300接收。

[0058] 服务器

[0059] 图3示意了促进双向会话的充当服务器系统300(或其部分)的服务器类计算机装置的一个实施例。服务器系统300被配置为:从产生器200接收音频和视频数据流,对数据流执行各种操控,并将数据流传输至接收器400和/或一个或多个web设备500。服务器系统300包括用于输入和输出数据流的网络层320、比特率适配模块310、媒体中继模块305和代码转换模块315。服务器系统320包括其他组件,诸如处理器、存储器、存储设备(例如硬盘)、网络和安全性机制(例如路由器、防火墙、高速缓存等),为了清楚而未示出这些组件。尽管为了清楚仅示出了单个服务器计算机,但是实际上,服务器系统300将包括一起耦合在网络中的多个服务器计算机。比特率适配模块310被配置为在音频和视频数据流从产生器200到达时从网络层320接收音频和视频数据流。比特率适配模块还从网络层320接收具有反馈数据的形式的与接收器400有关的信息。比特率适配模块310使用所有该数据来确定应当如何将音频和视频数据流转发至接收器400。比特率适配模块310的功能是:以必要的最小量的处理和延时,采用接收器理解的格式和比特率将音频和视频数据流中继至接收器400。

[0060] 为此,比特率适配模块310被配置为在音频和视频数据流经过服务器系统300时操控这些音频和视频数据流,以便基于接收器400处的移动计算设备的能力以及服务器系统300与接收器400之间的网络带宽,给接收器400提供接收器可以处理的音频和视频的最佳可能版本。

[0061] 比特率适配模块310还被配置为在从产生器200接收到音频和视频数据流时对音频和视频数据流进行存档拷贝,以存储在服务器系统300上。比特率适配模块310可以在必要时将该任务委派给媒体中继模块305或代码转换模块315,以确保对流的存档不会比必要时更多地影响流的延时。

[0062] 如果充足的网络带宽可用于接收器400以接收两个流的完全带宽传输,并且移动计算设备具有足以在服务器系统300接收到未修改的音频和视频数据流时足以处置未修改的音频和似乎品数据流的处理能力,则比特率适配模块310将音频和视频数据流路由通过媒体中继模块305。在一个实施例中,可以通过从接收器400接收反馈数据来确定网络带宽。可替换地或者与上述实施例相结合,比特率适配模块310或媒体中继模块305可以开始将数据流发送至接收器400,并测量接收数据流的速率以计算对接收器来说可用的网络带宽。

[0063] 媒体中继模块305被配置为充当通道(pass-through),以尽可能快地传输音频和视频数据流,以便降低双向会话的总体延时。媒体中继模块305还对音频和视频数据流进行存档拷贝,以存储在服务器系统300上。

[0064] 比特率适配模块310还被配置为在网络105上将音频和视频数据流传递至接收器400之前修改音频和视频数据流。这发生在对接收器400来说可用的网络带宽和/或接收器400处的移动计算设备的处理能力(或带宽)不足以处置在服务器处从产生器200接收到的数据流的比特率时。在一个实施例中,比特率适配模块310通过将音频数据流比特率与视频数据流比特率之和同网络带宽和处理带宽二者进行比较,确定不充足带宽可用。如果求和后的比特率超过这些带宽中的任一个,则比特率适配模块310降低数据流中的任一个或这两个数据流的比特率,以满足带宽限制。

[0065] 在一个实施例中,比特率适配模块310被配置为丢弃来自视频数据流的帧,以便满足网络105或接收器400移动计算设备的比特率限制。在另一实施例中,比特率适配模块310

被配置为将视频数据流向下转换至更低的屏幕分辨率和/或视频质量,以便满足上述限制。在其他实施例中,比特率适配模块310被配置为对音频数据流进行下采样或者在比特率上减小音频数据流,以便满足上述限制。比特率适配模块310也可以使用这些适配的任何组合。

[0066] 在视频流的视频帧速率响应于有限的接收器400带宽而降低的情况下,用户的体验将受到影响。所期望的是,如果网络或移动设备能力较差,则流的观察者将预期视频质量较差。然而,为了确保视频流不失去显著连贯性,仅可以丢弃通过TCP发送的视频的特定帧。如果丢弃了I帧或其部分,则视频的质量显著降低(潜在地包括解码中的一些延迟和回放中的一个或多个帧的丢失),直到接收到下一个I帧为止。当决定哪些帧要丢弃时,比特率适配模块310将仅丢弃P帧。比特率适配模块310决定哪些帧要丢弃的能力是使用TCP传输视频数据流而不是另一传输协议的另一个原因。

[0067] 相比于调整音频和视频流以在接收器处消耗的其他方法,比特率适配模块310丢弃视频流数据的帧的实施例具有优势。与通过将音频或视频流转换为更高质量格式来调整音频或视频流的比特率相比,丢弃帧需要由服务器系统300进行的非常少的处理。丢弃帧可以是即时执行的,并可以更快速地调整至在更低分辨率下再处理的改变的吞吐量条件。此外,在服务器系统300上丢弃帧而不是在被接收器400接收到时丢弃帧降低了移动计算设备上的处理需求,这在低端移动计算设备的情况下是有益的。此外,减少由移动计算设备执行的处理进而降低了编码的电池需求,因此改进了移动计算设备的总体性能。

[0068] 在一个实施例中,如果接收器400可以接受的时间平均的比特率小于从产生器200接收到的视频数据流比特率,则服务器系统300仅降低发送至接收器400的视频比特率。在另一实施例中,如果接收器400可以接受的时间平均的比特率小于从产生器200接收到的视频数据流比特率,则服务器系统300降低发送至接收器400的视频比特率,并且还将反馈数据(如上所述)发送至产生器200。在该实施例中,反馈数据用于降低从产生器200发送至服务器300的视频数据流的比特率。在这些实施例中的任一个中,基于由服务器系统300测量的接收器400的消耗模式,以规则间隔发送反馈数据。

[0069] 正常,音频数据流比特率足够低以使得不需要丢弃音频数据,尤其是考虑到由于将UDP用作传送协议而可能丢失一些音频。然而,在一个实施例中,比特率适配模块310还被配置为丢弃从产生器200接收到的音频数据的分组,以便降低发送至接收器400的音频比特率。然而,将仅在已经在接收器400处初始接收到/初始化音频数据流之后丢弃音频数据。

[0070] 根据数据流的接收者的硬件和/或软件配置,在一些情况下,将音频和视频数据流转换为接收者可以读取的格式是有益的。例如,用于在产生器200处对视频进行编码的编解码器可以不是安装在接收器400的移动计算设备上的编解码器。在另一示例中,web设备500可以被配置为接收ADOBE FLASH类型的视频数据流,而产生器200以H.264或H.263格式对视频进行原始编码。

[0071] 在比特率适配模块310确定了需要对数据流中的任一个重新编码的情况下,比特率适配模块310会将数据流传递出去至代码转换模块315。代码转换模块315被配置为即时对音频或视频数据流重新编码,并通过网络层310将重新编码的流传输至接收器400或任何web设备500。代码转换功能被设计为对延时进行最小化并维持音频和视频流的质量。还可以交给代码转换模块315以下任务:以新或旧的格式存档音频或视频流的拷贝。

[0072] 代码转换是将典型地把延时添加至音频和视频数据流的附加处理。可以通过降低从产生器200发送至服务器300的数据的比特率来对这种所添加的延时进行偏移。产生器200的恢复模块215、接收器400的恢复模块415和服务器的比特率适配模块310均被配置为计算每个音频和视频数据流的每个输入信道的比特率。将这些比特率统计与音频和视频数据的分组一起在产生器200、服务器300和接收器400之间发送。当比特率统计作为反馈数据与指示正在执行代码转换的信息一起被反馈至产生器200时，这向产生器200给出了同产生器与服务器之间以及服务器与接收器之间的网络带宽(或信道容量)有关的信息。然后，产生器200可以使音频和视频流的比特率适于避免服务器300中的否则将在双向视频会话中引入延迟的额外代码转换。

[0073] 接收器

[0074] 图4示意了充当接收器400的移动计算设备的一个实施例。接收器400被配置为从服务器系统300接收分离的音频和视频数据流，将它们同步在一起，并针对移动计算设备的用户回放它们。接收器400系统包括音频子系统405、视频子系统410、恢复模块415、网络层模块420、音频输出425和视频输入430。

[0075] 接收器400一般与产生器200反向地进行操作。音频输入425被配置为通过网络105从服务器系统300接收UDP音频数据流。类似地，视频输入430被配置为通过网络105从服务器系统300接收TCP视频数据流。将音频和视频数据流传递至恢复模块415。音频输入425一般将在视频输入430接收到视频数据的第一分组时之前接收到音频数据的第一分组。这是将UDP用于传送音频相对于用于视频的TCP的后果(但不是需求)。

[0076] 恢复模块415被配置为将由接收器400接收到的音频和视频数据流进行同步。当在产生器200处记录音频和视频数据流时，对音频和视频数据流进行初始同步。然而，这两个流将很可能在不同时刻到达接收器400。这可能是由于经由不同传输协议来分离地发送这些流的事实，原因在于可能已出现网络带宽问题，或者由于代码转换或比特率适配。更具体地，如果已经丢弃视频帧，则音频和视频将立即失步。恢复模块415使用音频数据流作为起始点来将音频和视频数据流进行同步。恢复模块415使用音频数据流中包括的时间戳信息来将音频数据与视频数据进行同步。

[0077] 恢复模块415另外被配置为维持自适应抖动缓冲器，以有助于确保同步的音频和视频回放。抖动缓冲器针对音频和视频数据流二者均维持至少250 ms的回放缓冲。这有助于确保音频和视频数据流的回放不会由于带宽或移动设备处理限制而被输入音频和视频数据流的瞬时丢弃所中断。在一个实施例中，恢复模块415查看视频分组，并在视频分组失步的情况下丢弃这些视频分组。在另一实施例中，模块415另外包括纠错算法用于纠正音频和视频流以改进回放质量。

[0078] 音频子系统405和视频子系统410被恢复模块415配置为针对移动计算设备的用户输出重新同步后的音频和视频数据流。音频和视频子系统利用移动计算设备本机的底层硬件和软件，如以上针对产生器200中的其等价物所述。因此，音频和视频系统利用移动计算设备的操作系统、数字信号处理器、视频显示器和扬声器组件以针对移动计算设备的用户实时回放音频和视频数据流。

[0079] 音频子系统405被配置为基于所安装的音频编解码器来对音频数据流进行解码，并将其输出至扬声器。音频子系统使用移动计算设备的数字信号处理器以将解码后的音频

数据转换为模拟时变电压信号,该模拟时变电压信号被馈送至移动计算设备的扬声器。视频子系统410使用所安装的视频编解码器和移动计算设备的操作系统来对视频流数据进行解码。将解码后的视频数据馈送至移动计算设备的视频显示屏。

[0080] 根据由接收器400接收到的音频和视频数据流的比特率以及接收器400移动计算设备处理输入数据的能力,恢复模块415的缓冲器可以利用等待给用户播放出来的数据进行填充。恢复模块415被配置为在网络105上通过网络层420将反馈数据发送至服务器系统300,以给服务器提供与接收器400的处理输入数据流的能力有关的数据。反馈数据包括同接收器与服务器系统300之间的连接有关的延时数据,该延时数据可以具有时间戳数据的形式。反馈数据另外包括指示接收器400处的移动计算设备可以处理的数据量的处理带宽。在一个实施例中,反馈数据具有对接收器来说可用的带宽量的形式、具有接收器400可以处置的若干千比特每秒的形式,二者均基于延时和处理能力。反馈数据另外可以包括编解码器数据,该编解码器数据包括与安装在接收器400上的音频和视频编解码器有关的信息,接收器400向服务器系统300通知代码转换是否是必要的。

[0081] Web设备

[0082] 图5示意了web设备500的一个实施例,web设备500被配置为两个移动计算设备之间的双向会话的第三方观察者,或者充当被广播以被第三方消耗的单向会话的第三方。Web设备500有意在与运行传统web浏览器的顶部软件平台有互联网连接的任何类型的硬件平台之上运行。Web设备500包括音频视频模块520、文本模块530、恢复模块515、闪式模块510和浏览器505。

[0083] 一般地,web设备500以与接收器400相同的方式工作。接收器的所有方面也由web设备500执行。例如,恢复模块515被配置为以与接收器400的恢复模块415相同的方式缓冲、同步从服务器系统300接收到的音频和视频流并提供关于这些音频和视频流的反馈数据。

[0084] Web设备500在几个方面不同。尽管web设备500可以被配置为如接收器400那样接收TCP视频数据流,但是web设备500还可以被配置为也接收实时消息收发协议(RTMP)视频数据流。Web设备500还被配置为经由实时流传输协议(RTSP)、Apple的专有HTTP实况流传输以及针对Google GTALK用户的可扩展消息收发和存在协议(XMPP)来接收流。

[0085] Web设备500还被配置为通过安装在移动计算设备的操作系统上的浏览器505来回放同步后的音频和视频数据流。恢复模块515另外被配置为与web设备500上的ADOBE FLASH 510安装进行通信,以促进通过浏览器505来回放音频和视频流。

[0086] Web设备500未被配置为充当双向会话的一个端点。因此,web设备500的底层硬件并不还充当产生器200,与在双向会话中既充当产生器200又充当接收器400的移动计算设备不同。

[0087] Web设备500包括文本模块530,文本模块530允许使用web设备500的观察者与双向或单向会话的各方进行交互或者通过文本消息的交换来与其他web设备500进行交互。Web设备500的用户通过浏览器505(具体地,通过闪式播放器510)来输入文本消息。音频视频模块520另外允许web设备500输出音频数据流,使得使用web设备500的观察者可以在单向或双向会话上提供音频反馈。

[0088] 网络层优化和电池优化

[0089] 为了有助于实现会话延时,使用数据流操控技术。该技术涉及对从产生器200和服

务器系统300的网络层220和330发出的网络业务的管理。本发明被设计为在所有不同类型的移动计算设备上工作。附加层的复杂度在于：存在大约380个不同的无线网络运营商给目前销售的大量不同的移动计算设备提供服务。无线网络运营商的示例包括AT&T、SPRINT、VERIZON和NTT。不同的无线运营商具有用于向其所支持的移动计算设备发送数据并从其所支持的移动计算设备接收数据的不同网络基础结构。

[0090] 具体地，不同运营商基于其网络基础结构的能力，具有不同的最大传送单元(MTU)。最大传送单元表示运营商可以一次传送的最大分组大小。如果移动计算设备试图发送比MTU更大的分组，则运营商在通过网络来传输该分组之前分解该分组。这是以用于执行转换的一些处理和延时时间为代价而进行的。尽管可以在不考虑移动计算设备正操作于其上的运营商的MTU的情况下实现产生器和服务器系统320，但是在考虑该运营商的MTU时更容易实现会话延时。

[0091] 在一个实施例中，产生器200和服务器系统300的网络层220和300被配置为检测数据的源和目的地之间的网络105的MTU。例如，产生器200与服务器系统300之间的MTU可能是1024个字节，而服务器与接收器之间的MTU可能是512个字节。通过在通过网络105来传输要在网络层220、320中发送的字节之前对其进行封装，对MTU的调整降低了与对数据进行转换以满足无线运营商的MTU需求相关联的延时。因此，本发明能够更好地实现单向或双向会话中的会话延时。

[0092] 在移动计算设备上操作单向或双向会话利用了触摸屏、网络设备、摄像机、软件音频-视频编码器(协处理器)、以及移动计算设备的CPU，所有这些组件均占用电池功率。在大多数移动计算设备操作系统中，“待机模式”通过在设备上全面地利用通知服务来节约电池。产生器200和接收器400使用待机模式来最小化电池使用。具体地，产生器200和接收器400接收器“保活”来自服务器300的分组以周期性地唤醒。服务器300通过对保活的分组的传输时间进行随机化来与移动计算设备一起工作，以降低电池消耗。另一选项是使用定时器，然而由于定时器占用了移动计算设备的CPU上的处理时间，因此这些定时器更好地仅用在具有高端处理器和电池寿命的移动计算设备中。

[0093] 使用情况情形

[0094] 双向会话的过程根据几个因素而稍有变化。这些因素包括双向会话中涉及的移动计算设备的处理能力、移动计算设备之间的网络带宽、以及是否需要针对移动计算设备或观察该会话的任何web设备执行代码转换。根据这些因素，可以出现多种使用情况。

[0095] 第一种情况是：两个移动计算设备均具有高端硬件，使得移动计算设备可以处置为对音频和视频进行输入、流传输和输出所需的处理。这里，产生器200产生音频和视频并通过网络105将其流传输至服务器系统300。服务器开始与接收器400移动计算设备的初始化(或“握手”)过程。服务器系统300从接收器400接收与安装在接收器上的用于音频和视频解码的编解码器有关的反馈数据、以及与接收器的处理能力有关的信息和网络带宽信息。

[0096] 在该第一种情况下，接收器400的反馈数据指示接收器能够处置产生器200已发送的所有视频和音频流数据。此外，在这种情况下，假定代码转换不是必要的，使得接收器可以接受现在的样子的流。服务器系统300将流的控制从比特率适配模块310传递至媒体中继模块305。媒体中继模块305以最小可能延迟将流向前传递至接收器400。一旦接收器接收到该流，进一步反馈就不必要，除非网络带宽下降。如果网络带宽下降，并且接收器400上的恢复

模块415中的缓冲器填满，则将反馈数据发送至服务器系统300，指示比特率适配模块310应当从媒体中继模块405接管，以降低被发送至接收器的流的比特率。

[0097] 第二种情况是：充当产生器200的移动计算设备具有高端硬件，并且充当接收器400的移动计算设备具有低端硬件。在这种情况下，从接收器400至服务器系统300的初始化过程反馈数据将指示接收器400无法处置从产生器20发送的全部数据流。比特率适配模块310将处置该流，并将降低视频流的比特率以便满足接收器400的处理限制。例如，如果来自产生器200的流中的视频的比特率是120kbps，并且接收器的反馈数据指示接收器400仅可以处置80kbps，则比特率适配模块310将丢弃视频帧，直到满足必需的比特率80kbps为止。

[0098] 第三种情况是：充当接收器400的移动计算设备安装有与充当产生器200的移动计算设备不同的音频或视频编解码器。在这种情况下，从接收器400至服务器系统300的初始化过程反馈数据将指示什么音频和视频编解码器安装在接收器400上。比特率适配模块310将会把需要代码转换的音频或视频流（或这两者）传递至代码转换模块315。代码转换模块315将在将这些流向前传递至接收器400之前在需要时进行代码转换。

[0099] 由于代码转换涉及额外处理，因此将流传递至代码转换模块315存在一些增加的延时。在必要时，代码转换模块310可以将音频或视频流传递回到比特率适配模块315，以降低输出流的比特率以便补偿该附加延时。

[0100] 双向会话示例

[0101] 图6示意了两个移动计算设备之间的示样双向会话的一个实施例。在图6a和图6b中，麦克风开启/关闭/锁定615指示了移动计算设备的麦克风目前是否正在录音频。同样在图6a和6b中，开始/停止指示器620指示移动计算设备是否正在活动地捕获视频和音频。这两个指示器是分离的，使得可以在期望时在没有音频的情况下捕获视频。

[0102] 图6a示意了在与图6b所示的第二移动计算设备的双向会话中既充当产生器又充当接收器的第一移动计算设备。图6a中的移动计算设备正在显示屏的相对侧使用摄像机以捕获移动计算设备的用户面前的外部环境。将由摄像机捕获的环境显示在第一移动计算设备的屏幕600a上屏幕605a的小部分中。该摄像机将视频流直接馈送至屏幕605a的小角落处，使得用户知道正在捕获并向第二移动计算设备600b发送的内容。充当产生器，第一移动计算设备将音频和视频流发送至服务器，然后该音频和视频流被实时传输至第二移动计算设备。

[0103] 充当接收器，第二移动计算设备从服务器接收音频和视频流。第二移动计算的屏幕600b显示接收到的视频605b。在一个实施例中，接收到的流填充了屏幕的大多数部分，以向用户给出正由第一移动计算设备记录的内容的良好画面。第二移动计算设备上的面向前的摄像机响应于查看来自第一移动计算设备的流，实时记录第二移动计算设备的用户的图像。所记录的图像被显示在屏幕610a的小角落处，并且是来自面向前的摄影机的直接馈送。因此，在这两个移动计算设备中，均保留屏幕的小角落，以向用户给出对所记录的内容的预览，使得用户可以将移动计算设备调整为捕获正确的内容，以将其发送至其他移动计算设备。

[0104] 同时充当产生器，第二移动计算设备将第二用户对第一用户的内容的反应的流发送至服务器，并从该服务器发送至第一移动计算设备。同时还充当接收器，第一移动计算设备将另一用户的反应的流显示在移动计算设备的屏幕610b的中央处。

[0105] 附加考虑

[0106] 在该描述中,术语“模块”指代用于提供指定功能的计算逻辑。模块可以以硬件、固件和/或软件实现。在这里描述的模块被实现为软件的情况下,模块可以被实现为独立的程序,但也可以通过其他装置而实现,例如作为更大程序的一部分、作为多个分离程序、或者作为一个或多个静态或动态链接库。将理解,这里描述的指定模块表示本发明的一个实施例,并且其他实施例可以包括其他模块。此外,其他实施例可以缺少这里描述的模块和/或以不同方式将所描述的功能分布在这些模块当中。此外,可以将归于多于一个模块的功能并入到单个模块中。在模块由软件实现的实施例中,模块永久性地(即,非瞬变地)存储在移动设备或服务器系统的计算机可读存储设备上,被加载至存储器中,并被一个或多个处理器执行。

[0107] 移动计算设备、服务器设备和web设备包括这里描述的操作所必需的硬件元件,包括一个或多个处理器、高速存储器、硬盘存储器和备份、网络接口和协议、用于输入录入的输入设备以及用于对数据进行显示、打印或其他呈现的输出设备。移动计算设备和web设备另外包括用于音频和视频捕获、流传输和记录的音频和视频输入。移动计算设备另外包括用于将所记录的音频和视频从模拟至数字信号来回转换的数字信号处理器。移动计算设备还包括用于记录音频的麦克风、用于播放所记录的音频的扬声器、以及用于记录视频的面向前或面向后的摄像机。

[0108] 移动计算设备另外包括操作系统。操作系统包括用于向和从计算机可读格式转换原始音频和视频输入的音频和视频编解码器。示例音频编解码器包括G711、AAC、Vorbis。示例视频编解码器包括QUICKTIME、MPEG H.264、H.263和VP8。本发明也可以适用于其他编解码器。

[0109] 以上描述的一些部分描述了信息上的运算的算法和符号表示方面的实施例。数据处理领域技术人员通常使用这些算法描述和表示,以向本领域其他技术人员有效地传达其工作的实质。尽管这些运算是功能上、计算上或逻辑上描述的,但是这些运算被理解为由处理器所执行的计算机程序、等效电路、微代码等实现。此外,将运算的这些布置称为模块也已被证明有时方便,而不失一般性。所描述的运算及其关联模块可以以软件、固件、硬件或其任何组合体现。

[0110] 如这里所使用的对“一个实施例”或“一实施例”的任何引用意味着:结合实施例而描述的特定元素、特征、结构或特性包括在至少一个实施例中。短语“在一个实施例中”在说明书各处的出现不必均指代相同实施例。

[0111] 如这里所使用的术语“包括”、“包含”、“具有”及其分词或其任何其他变型意在覆盖非穷举的包括。例如,包括元素列表的过程、方法、物品或装置不必仅限于那些元素,而是可以包括未明确列出或者对这种过程、方法、物品或装置固有的其他元素。此外,除非明确地相反声明,“或”指代包含或而非指代异或。例如,以下任一项满足条件A或B:A为真(或存在)且B为假(或不存在);A为假(或不存在)且B为真(或存在);以及A和B均为真(或存在)。

[0112] 此外,“一”或“一个”的使用用于描述这里实施例的元件和组件。这么做仅为了方便,并给出本发明的一般含义。该描述应当被解读为包括一个或至少一个,并且单数也包括复数,除非明显意在不包括多个。

[0113] 在阅读本公开后,本领域技术人员将意识到,系统和服务器过程还有附加的替换

结构和功能设计通过这里公开的原理辅助了双向视频会话。因此,尽管示意并描述了特定实施例和应用,但是要理解,所公开的实施例不限于这里公开的精确构造和组件。在不脱离所附权利要求中限定的精神和范围的前提下,可以在这里公开的方法和装置的布置、操作和细节上进行对本领域技术人员来说将显而易见的各种修改、改变和变更。

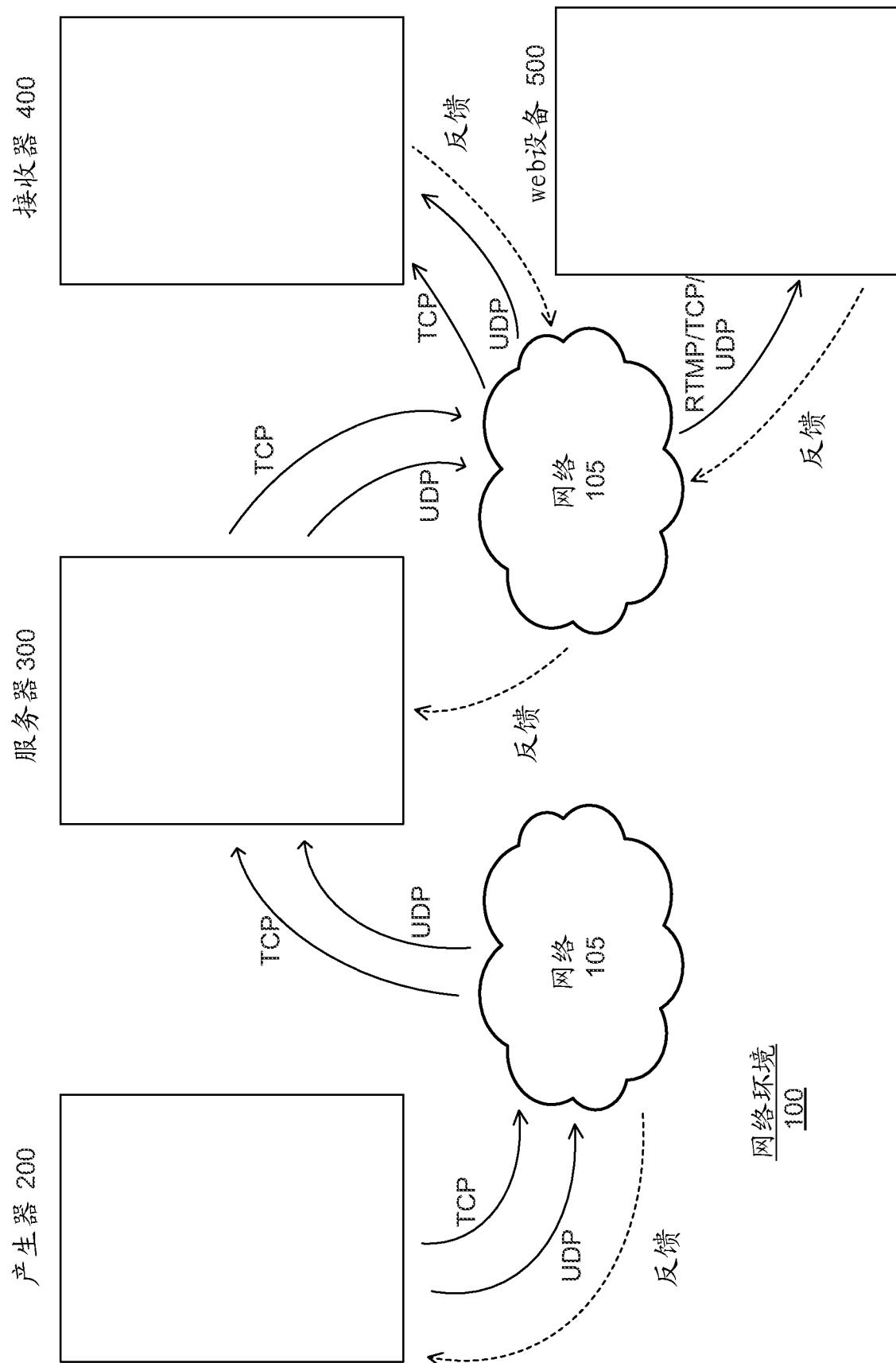


图 1

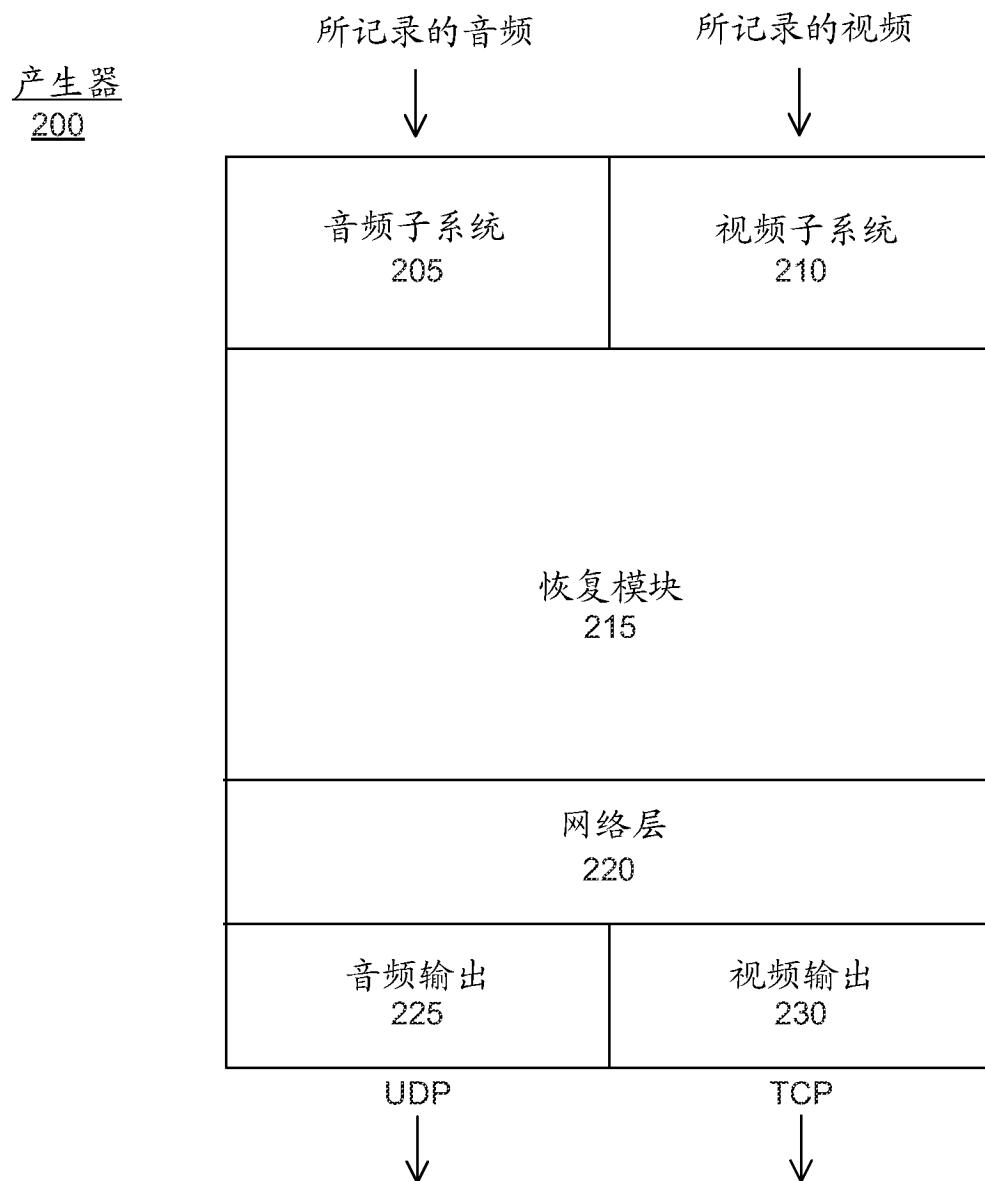


图 2

服务器
300

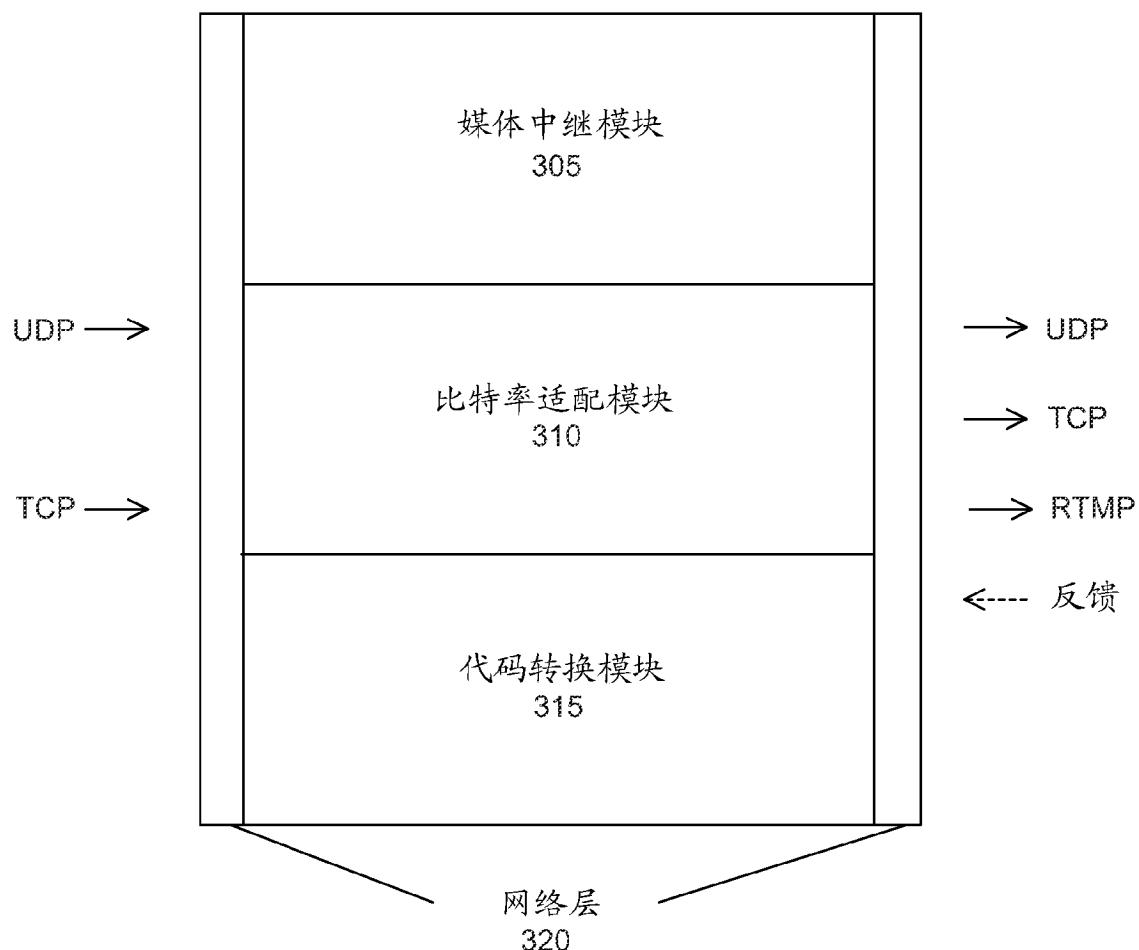


图 3



图 4



图 5

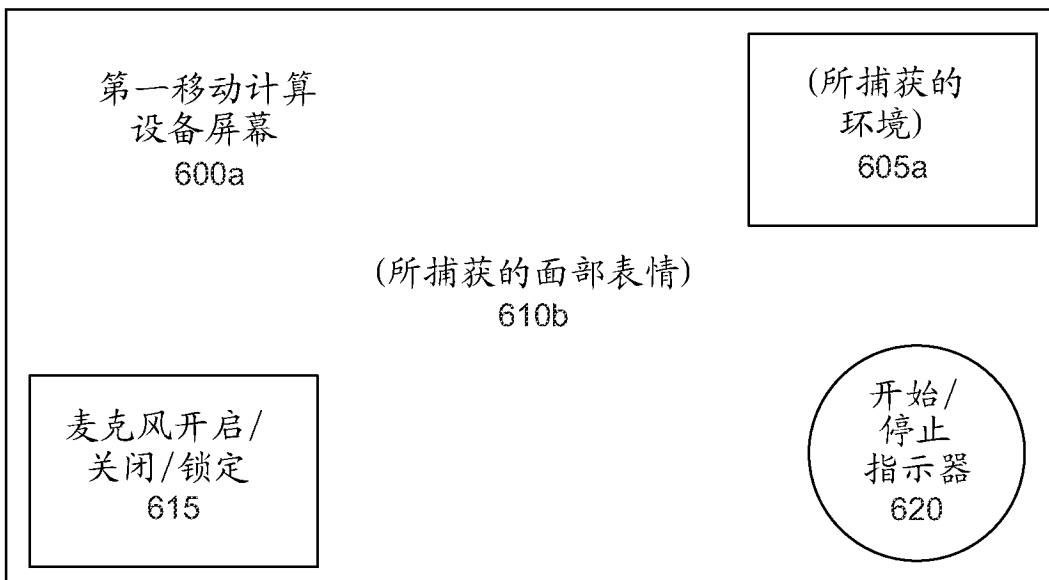


图 6a

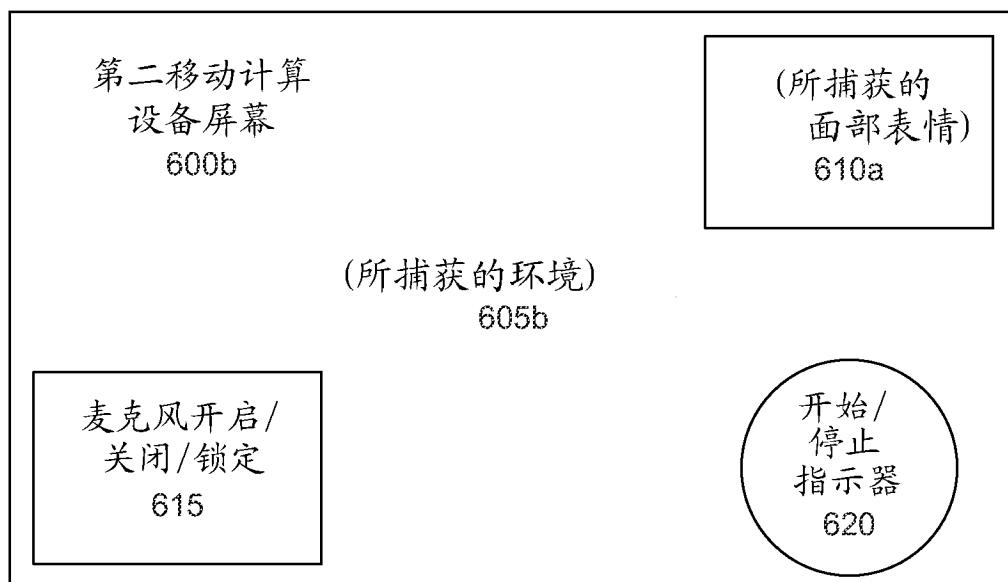


图 6b

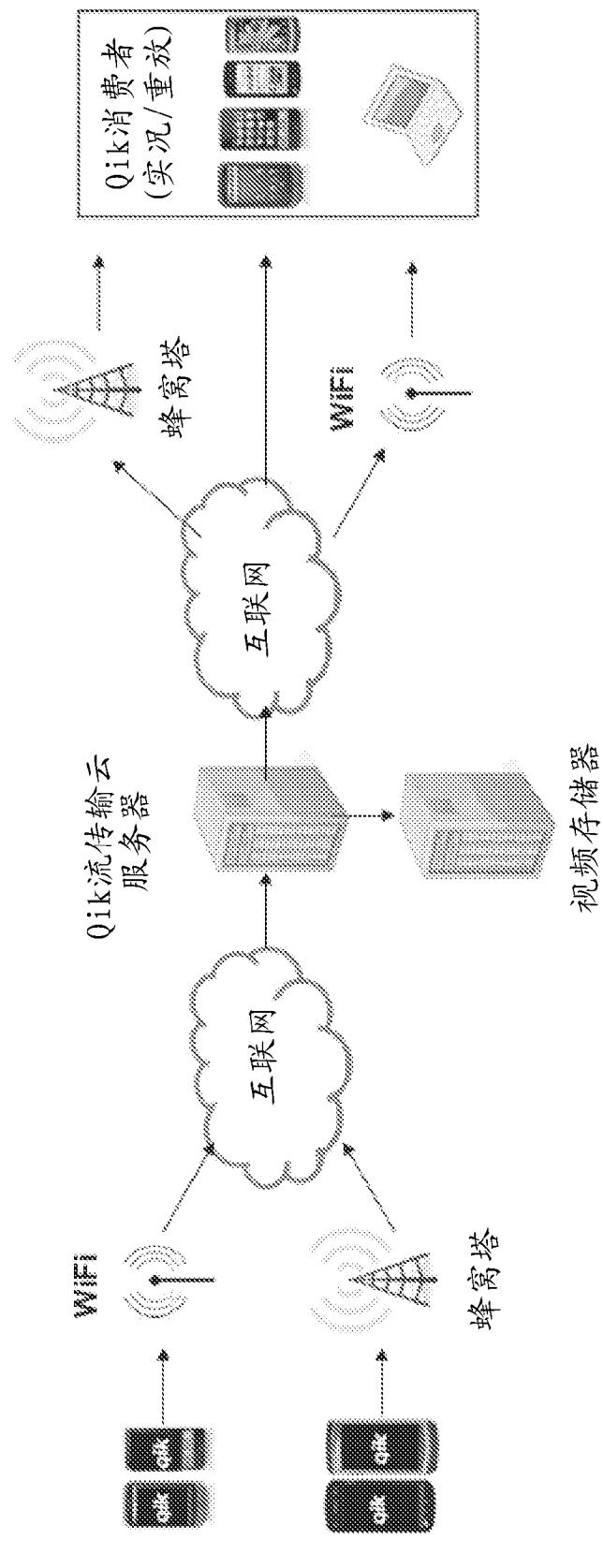


图 7

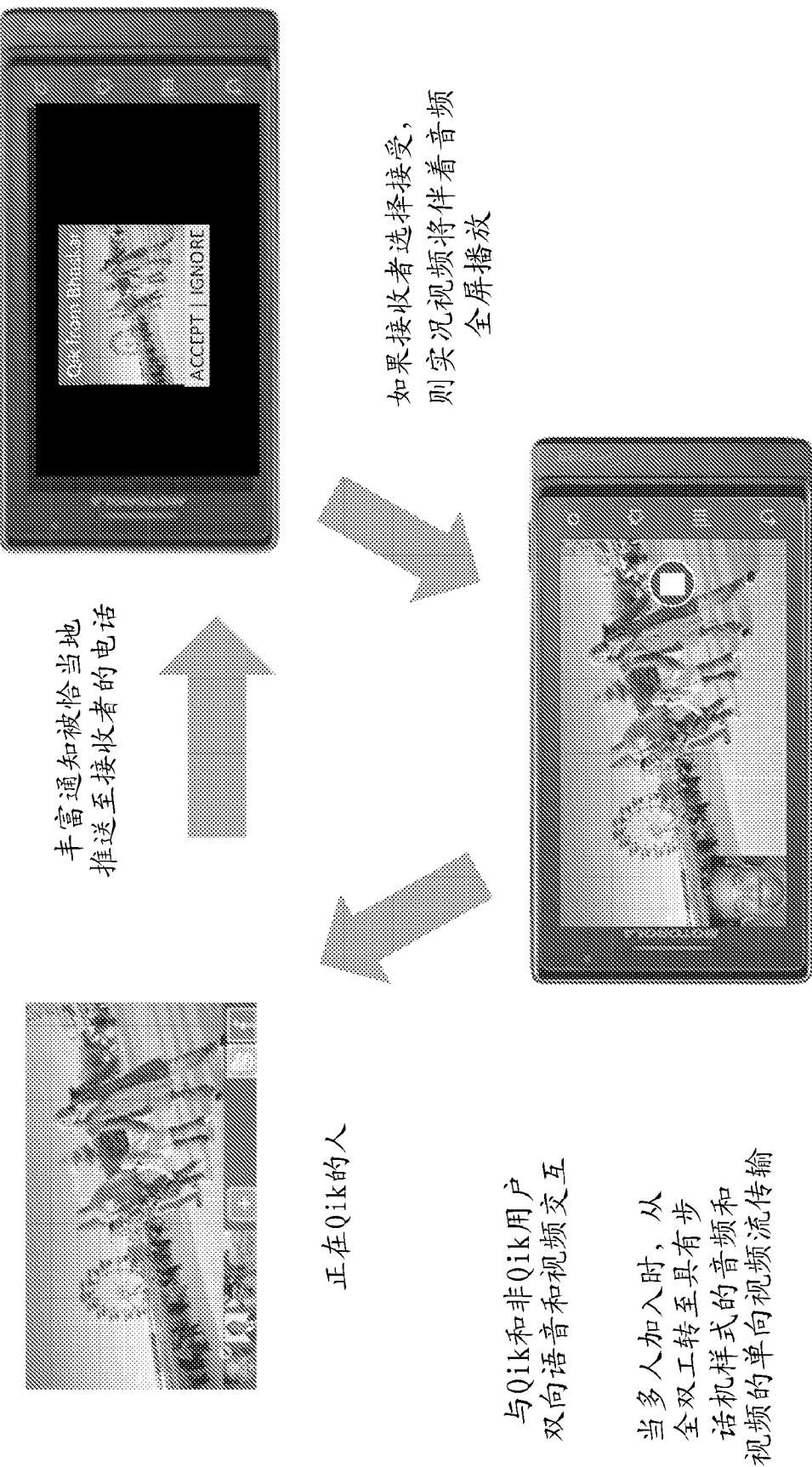


图 8

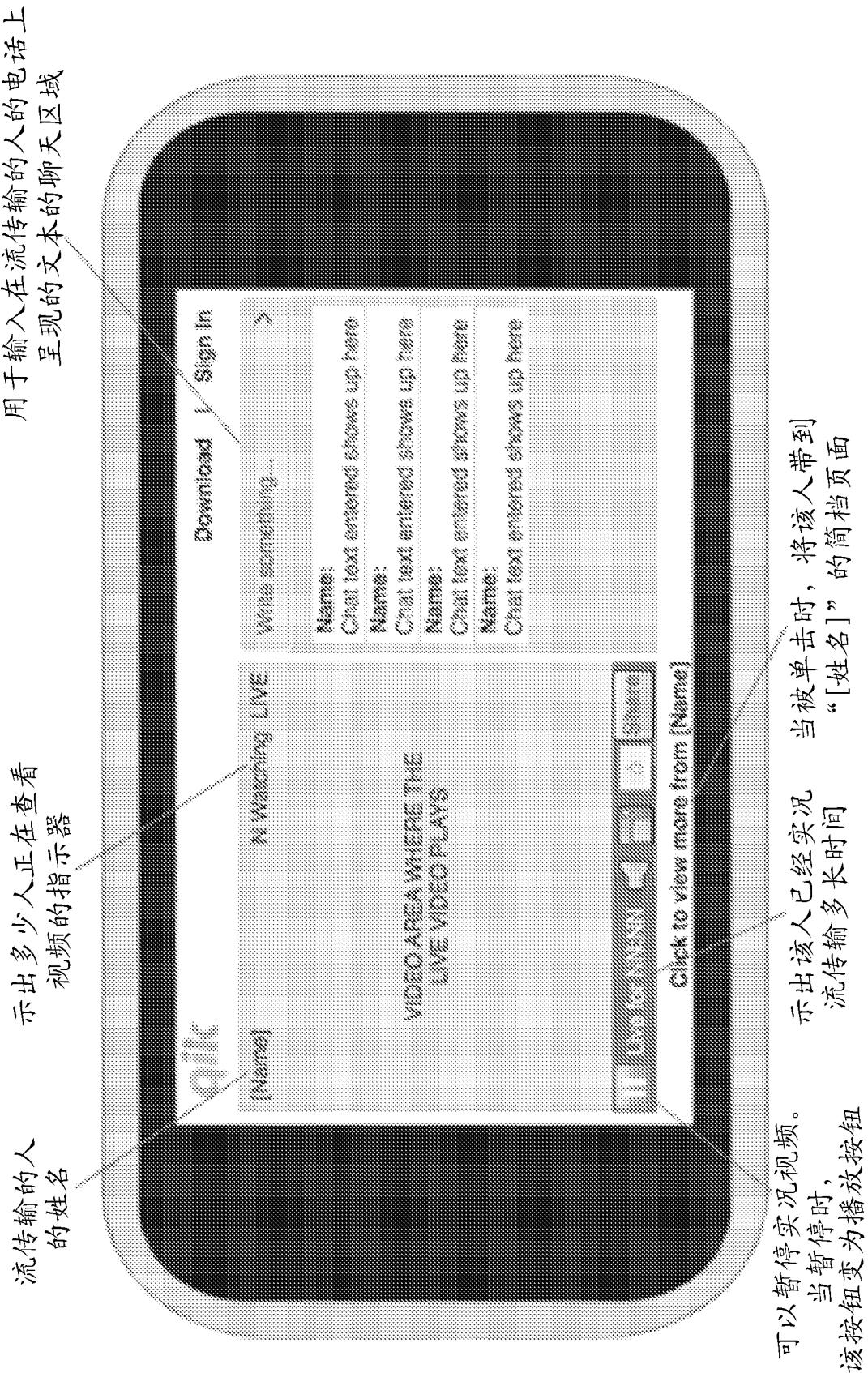


图 9