



(12)发明专利申请

(10)申请公布号 CN 106471802 A

(43)申请公布日 2017. 03. 01

(21)申请号 201580034991.8

(22)申请日 2015.06.26

(30)优先权数据

14/481,079 2014.09.09 US

(85)PCT国际申请进入国家阶段日

2016.12.27

(86)PCT国际申请的申请数据

PCT/US2015/038122 2015.06.26

(87)PCT国际申请的公布数据

W02016/039835 EN 2016.03.17

(71)申请人 谷歌公司

地址 美国加利福尼亚州

(72)发明人 布赖恩·大卫·马什

(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 梁晓广 周亚荣

(51)Int.Cl.

H04N 9/74(2006.01)

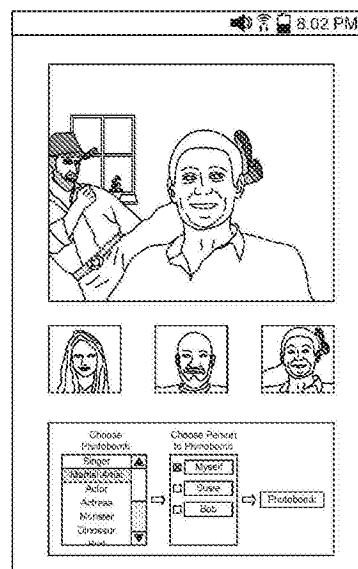
权利要求书2页 说明书10页 附图12页

(54)发明名称

视频会议中的实时视频变换

(57)摘要

公开了用于视频会议中的实时视频变换的系统和方法。方法包括由处理设备从视频会议的第一参与者接收修改视频流的请求。该方法进一步包括由所述处理设备识别视频流内的前景和背景。该方法进一步包括由所述处理设备生成修改的视频流,所述修改的视频流包括插入到所述背景中的视频或图像。



1. 一种方法,包括:

由处理设备从视频会议的第一参与者接收修改视频流的请求;

由所述处理设备识别所述视频流内的前景和背景;以及

由所述处理设备生成修改的视频流,所述修改的视频流包括插入到所述背景中的视频或图像,其中,所述视频或图像当被显示时出现在所述前景之后。

2. 如权利要求1所述的方法,进一步包括:

识别所述前景在视频的帧内的相对位置,其中,生成修改的视频流进一步包括基于识别的相对位置将视频或图像插入到所述背景中。

3. 如权利要求1所述的方法,其中,所述视频流由视频会议的第一参与者的第一客户端设备捕捉,并且其中所述处理设备是支持视频会议的服务器的一部分,所述方法进一步包括:

由所述处理设备在接收修改视频流的请求之前接收视频流。

4. 如权利要求3所述的方法,进一步包括:

由所述处理设备发送修改的视频流到视频会议的附加参与者的一个或多个客户端设备。

5. 如权利要求1所述的方法,进一步包括:

从视频会议的第一参与者接收对视频或图像的选择。

6. 如权利要求1所述的方法,其中,所述前景对应于第一参与者的面部。

7. 如权利要求1所述的方法,其中,所述视频或图像要在预定的持续时间出现在修改的视频流的前景之后。

8. 如权利要求1所述的方法,其中,所述视频或图像是名人的视频或图像、动物的视频或图像、或者卡通角色的视频或图像中的任何一种。

9. 一种系统,包括:

存储器;以及

处理设备,所述处理设备通信耦合到所述存储器,其中,所述处理设备用于:

从视频会议的第一参与者接收修改视频流的请求;

识别所述视频流内的前景和背景;以及

生成修改的视频流,所述修改的视频流包括插入到所述背景中的视频或图像,其中,所述视频或图像当被显示时出现在所述前景之后。

10. 权利要求9所述的系统,其中,所述处理设备进一步用于:

识别所述前景在视频的帧内的相对位置;以及

基于识别的相对位置将视频或图像插入到所述背景中。

11. 如权利要求9所述的系统,其中,所述视频流由视频会议的第一参与者的第一客户端设备捕捉,并且其中所述处理设备是支持视频会议的服务器的一部分,其中所述处理设备进一步用于:

在接收修改视频流的请求之前接收视频流。

12. 如权利要求11所述的系统,其中,所述处理设备进一步用于:

发送修改的视频流到视频会议的附加参与者的一个或多个客户端设备。

13. 如权利要求9所述的系统,其中,所述处理设备进一步用于:

从视频会议的第一参与者接收对视频或图像的选择。

14. 如权利要求9所述的系统,其中,所述前景对应于第一参与者的面部。

15. 如权利要求9所述的系统,其中,所述视频或图像要在预定的持续时间出现在修改的视频流的前景之后。

16. 如权利要求9所述的系统,其中,所述视频或图像是名人的视频或图像、动物的视频或图像、或者卡通角色的视频或图像中的任何一种。

17. 一种非瞬时计算机可读存储介质,具有存储在其上的指令,所述指令当被处理设备执行时,使得所述处理设备执行包括以下的操作:

从视频会议的第一参与者接收修改视频流的请求;

识别所述视频流内的前景和背景;以及

生成修改的视频流,所述修改的视频流包括插入到所述背景中的视频或图像,其中,所述视频或图像当被显示时出现在所述前景之后。

18. 如权利要求17所述的非瞬时计算机可读存储介质,其中,所述操作进一步包括:

识别所述前景在视频的帧内的相对位置,其中,生成修改的视频流进一步包括基于识别的相对位置将视频或图像插入到所述背景中。

19. 如权利要求17所述的非瞬时计算机可读存储介质,其中,所述视频流由视频会议的第一参与者的第一客户端设备捕捉,并且其中所述处理设备是支持视频会议的服务器的一部分,其中所述操作进一步包括:

由所述处理设备在接收修改视频流的请求之前接收视频流。

20. 如权利要求17所述的非瞬时计算机可读存储介质,其中,所述操作进一步包括:

从视频会议的第一参与者接收对视频或图像的选择。

视频会议中的实时视频变换

技术领域

[0001] 本公开涉及视频会议的领域,更具体地涉及视频会议中的实时视频变换。

背景技术

[0002] 互联网上的视频协作(例如,视频会议、远程医疗等等)正随着宽带接入和利用了宽带连接性的应用激增而变得日益普通。支持视频协作的应用可以是基于浏览器的,或者可以独立于浏览器而运行。

发明内容

[0003] 下面呈现本公开的各个方面的简化概述,以便提供对这些方面的基本理解。本概述不是本公开的扩展概览。既不意图标识本公开的关键或临界元素,也不描绘本公开的特定实现的任何范围或权利要求的任何范围。其唯一用途在于用简化形式呈现本公开的一些概念,作为稍后呈现的更详细描述의序幕。

[0004] 在本公开的一个方面中,一种方法包括由处理设备从视频会议的第一参与者接收修改视频流的请求。所述方法进一步包括由所述处理设备识别视频流内的前景和背景。所述方法进一步包括由所述处理设备生成修改的视频流,所述修改的视频流包括插入到所述背景中的视频或图像,其中所述视频或图像当被显示时出现在所述前景之后。

[0005] 公开了用于执行上述方法的操作的计算设备以及这里所述的各种实现。还公开了存储用于执行与上述方法相关联的操作和这里所述的各种实现的指令的计算机可读介质。

附图说明

[0006] 通过举例的方式而不是限制的方式来以附图的图片图示说明本公开,在附图中:

[0007] 图1图示说明根据本公开的实现的示例系统架构;

[0008] 图2是图示说明根据本公开的实现的会议平台和视频变换部件的框图;

[0009] 图3是图示说明根据本公开的实现的媒体协作服务/应用的示例图形用户界面窗口的图;

[0010] 图4是图示说明根据本公开的实现的用于变换视频会议中的视频流的方法的流程图;

[0011] 图5是图示说明根据本公开的实现的媒体协作界面的图;

[0012] 图6A是图示说明根据本公开的实现的视频流中识别的前景的图;

[0013] 图6B是图示说明根据本公开的实现的视频流中识别的背景的图;

[0014] 图7A是图示说明根据本公开的实现的应用于视频流的背景的变换的图;

[0015] 图7B是图示说明根据本公开的实现的组合视频流的前景与修改后的背景的变换的图;

[0016] 图8A是图示说明根据本公开的实现的在视频会议期间的视频流的变换的图;

[0017] 图8B是图示说明根据本公开的实现的在视频会议期间的视频流的变换的图;

- [0018] 图8C是图示说明根据本公开的实现的在视频会议期间的视频流的变换的图
- [0019] 图8D是图示说明根据本公开的实现的在视频会议期间的视频流的变换的图；
- [0020] 图9是图示说明根据本公开的实现的示例计算机系统的框图。

具体实施方式

[0021] 本公开的实现的各个方面涉及向视频流施加实时变换。公开的系统和方法可以应用于媒体协作(例如,音频和/或视频会议、音频和/或视频聊天、音频和/或视频会议室、音频和/或视频聊天室等等),其中,由一个或多个参与者/用户的设备所发送的内容/媒体流(例如现场音频和/或视频内容、录制的音频和/或视频内容等等)可以被组合到合成内容/媒体流中。现有视频会议技术使得参与者能够观看视频会议的其他参与者所提供的视频流,且这样的流通常布置为(例如在单个界面内)使得单个用户(例如主呈现者或当前讲话者)被分配有界面的相对较大或较显著的部分。

[0022] 这里在各种实现中描述的是允许媒体协作的一个或多个参与者请求将与参与者相关联的视频流进行实时变换的技术。例如,如这里所述,在视频会议的过程期间,接收将图像或视频插入到视频流(请求者的视频流和/或不同参与者的视频流)的背景中的请求。插入的图像或视频可以出现在背景中(例如在参与者之后)、在前景中(例如在参与者之前或之前)、或二者。修改的视频流可以被生成(例如,由内容处理服务器、会议服务器、和/或在参与视频会议的客户端设备本地),且被发送到视频会议的每个参与者,以便在视频会议期间产生审美上的吸引力和/或幽默效果。

[0023] 图1图示说明根据本公开的实现的示例系统架构100。系统架构100包括客户端设备110A-110Z、网络105、数据存储装置(data store)106、会议服务器120和内容处理服务器130。系统架构100的各种设备中的每个可以经由网络105互相连接。在一个实现中,网络105可以包括公共网络(例如互联网)、专用网络(例如局域网(LAN)或广域网(WAN))、有线网络(例如以太网)、无线网络(例如802.11网络或Wi-Fi网络)、蜂窝网络(例如长期演进(LTE)网络)、路由器、集线器、交换机、服务器计算机和/或上述的组合。

[0024] 在一个实现中,数据存储装置106可以是存储器(例如随机存取存储器)、缓存、驱动器(例如硬盘驱动器)、闪存驱动器、数据库系统或者能够存储数据的另一类型的部件或设备。数据存储装置106还可以包括多个存储部件(例如多个驱动器或多个数据库),所述多个存储部件还可以跨越多个计算设备(例如多个服务器计算机)。在一些实现中,数据存储装置106可以是基于云的。系统架构100的一个或多个设备可以利用数据存储装置106来存储公用和专用数据,且数据存储装置106可以被配置为向专用数据提供安全存储。数据存储装置106可以是客户端设备110A-110Z、会议服务器120和内容处理服务器130的一部分或分布在其任意之间。

[0025] 在一个实现中,客户端设备110A-110Z可以包括一个或多个计算设备,诸如个人电脑(PC)、笔记本、移动电话、智能电话、平板电脑、上网本计算机等等。客户端设备110A-110Z也可以被称为“用户设备”或“移动设备”。单个用户可以与一个或多个客户端设备(例如,客户端设备110A-110Z中的一个或多个)相关联(例如,拥有和/或使用)。客户端设备110A-110Z可以每个都被在不同地理位置上的不同用户所拥有和利用。参与视频协作(例如视频会议)的用户在此可被称为“会议参与者”。

[0026] 客户端设备110A-110Z每个可以分别实现用户界面111A-111Z。每个用户界面111A-111Z可以允许相应客户端设备110A-110Z的用户发送和接收信息到会议服务器120和内容处理服务器130中的一个或多个。例如,用户界面111A-111Z中的一个或多个可以是web浏览器界面,其可以接入、检索、呈现和/或导航由会议服务器120所提供的内容(例如web页面,诸如超文本标记语言(HTML)页面)。在一个实现中,用户界面111A-111Z中的一个或多个可以是独立应用(例如,移动app、会议软件等等),其允许相应客户端设备110A-110Z的用户发送和接收信息到会议服务器120和/或内容处理服务器130。在一个实现中,用户界面111A-111Z中的一个或多个可以是会议接口,其允许会议参与者从事视频会议、音频会议、聊天和/或即时消息传递中的一个或多个。示例用户界面(例如图形用户界面窗口)结合图3来更详细地描述。

[0027] 每个客户端设备110A-110Z进一步分别包括媒体观看器112A-112Z。在一个实现中,媒体观看器112A-112Z可以是应用,其允许用户观看内容,诸如图像、视频、web页面、文档等等。例如,媒体观看器112A可以是web浏览器,其可以接入、检索、呈现和/或导航由web服务器所服务的内容(例如web页面,诸如超文本标记语言(HTML)页面,数字媒体项等等)。媒体观看器112A可以渲染、显示和/或呈现内容给用户。媒体观看器112A还可以显示嵌入式媒体播放器,其嵌入在web页面中(例如,可以提供关于在线商户所销售的产品的信息的web页面)。在另一示例中,媒体观看器112A可以是独立应用(例如移动app),其允许用户观看数字媒体项(例如数字视频、数字图像、电子书等等)。在一个实现中,媒体观看器112A-112Z可以进一步允许视频从一个或多个源接收且在相应的用户界面111A-111Z内显示。例如,客户端设备110A可以从会议服务区120接收视频数据,其可对应于由一个或多个附加客户端设备所生成的视频流。媒体观看器112A可以基于接收到的视频数据生成用于显示(例如在用户界面111A内)的对应于每个视频流的视频。

[0028] 在一个实现中,会议服务器120可以是一个或多个计算设备(诸如机架式服务器、路由器计算机、服务器计算机、个人电脑、大型计算机、笔记本电脑、平板电脑、桌上电脑等等)、数据存储装置(例如硬盘、存储器、数据库)、网络、软件部件、硬件部件、或者可以适用于实现这里所述的各种特征的上述的组合。在一些实现中,会议服务器120可以启用媒体协作服务,诸如音频和/或视频会议(例如,在客户端设备110A-110Z的用户之间),使用例如流视频或者IP语音(VoIP)技术,且可以用于面向个人、娱乐、商业、教育或学术的交互。会议服务器120可以专用于提供视频会议服务,或者可以连同其他服务(包括例如新闻服务、社交网络服务和/或内容托管服务)一起提供视频会议服务。

[0029] 在一个实现中,会议服务器120包括会议平台200。如图2中所图示,会议平台200包括托管模块202、接口模块204以及内容档案206。在会议平台200中可以包括更多或更少的部件而不失一般性。例如,可以将两个或更多的模块合并到单个模块中,或者模块之一可以被分割成两个或更多模块。在一个实现中,一个或多个模块可以驻留在不同计算设备上(例如,不同服务器计算机、单个客户端设备、在多个客户端设备之间分布,等等)。

[0030] 在一个实现中,托管模块202可以由会议平台200用来初始化并支持媒体协作,诸如音频/视频会议、聊天室、视频聊天等等。例如,托管模块202可以接收来自用户的请求以创建媒体协作会话,可以在从用户接收到请求后允许用户加入预先存在的媒体协作会话,可以便于发送对用户的邀请,许可用户加入预先存在的媒体协作会话,等等。

[0031] 在一个实现中,接口模块204可以由会议平台200用于从各参与者的客户端设备110A-110Z中的一个或多个接收多个音频和/或视频流,并且生成合成流,合成流可以包括与每个接收到的流相关联的数据。合成流可以随后被提供至媒体协作会话(例如,视频会议)中各个参与者的一个或多个客户端设备110A-110Z。在一些实现中,合成流可以包括格式化数据,其可以由客户端设备110A-110Z用来提取单个视频流,并布置视频流用于由相应的用户界面111A-111Z来呈现。因此,视频会议的每个用户/参与者可以在单个界面/屏幕内同时观看内容流中的一些或全部,如图5和8中所图示。

[0032] 在一些实现中,会议平台200所托管的媒体协作中的一个或多个参与者可以提供/放出音频内容(例如,参与者的讲话语音、音乐等等),其在被客户端设备(例如客户端设备110A-110Z之一)接收到/感知到之后,可以被客户端设备发送或提供作为音频流,该音频流可以合并到媒体协作中。在一些实现中,这样的音频流可以独立于视频流而提供(例如,在媒体协作期间参与者提供仅音频输入),而在其他实现中,这样的音频流可以结合视频流而提供(例如,合并的流,合并参与者讲话的同步的视频和音频)。

[0033] 在一些实现中,内容档案206可以由会议平台200利用以存储媒体协作数据(例如,存储视频和/或音频流数据、聊天数据等等),这可以由一个或多个参与者经其相应的客户端设备稍后访问。在一些实现中,内容档案可以与数据存储装置106合并。

[0034] 在一个实现中,内容处理服务器130包括内容变换部件250。如图2中所图示,内容变换部件250包括前景识别模块252、目标识别模块254以及内容生成模块256。更多或更少的部件可以包括在内容变换部件250中而不失一般性。例如,两个或更多模块可以合并成单一模块,或者一个模块可以分解成两个或更多模块。在一个实现中,一个或多个模块可以驻留在不同计算设备上(例如,不同服务器计算机上、在单个客户端设备上、在多个客户端设备之间分布,等等)。

[0035] 在一个实现中,前景识别模块252和目标识别模块254可以由内容变换部件250用来识别视频流的一个或多个帧的相关部分。例如,前景识别模块252可以将视频流的一个或多个帧分成前景和背景,例如基于视频图像的逐帧对比,以确定对象运动(例如使用运动结构算法(structure-from-motion algorithm))。在一些实现中,前景识别模块252可以利用一个或多个数字信号处理(DSP)芯片来执行实时图像处理。目标识别模块254可以自动在前景或背景内识别感兴趣的目标/区域(例如,视频会议参与者的面部)。各种算法/技术可以被利用来识别图像的前景和感兴趣的区域,如本领域普通技术人员所认识的。

[0036] 在一个实现中,内容生成模块256可以被内容变换部件250用来使用由前景识别模块和/或目标识别模块254生成的数据而生成修改的视频流。例如,视频流和要插入到视频流中的图像或视频可以基于所识别的视频流的前景和背景而被解析为分别的图像,且被合并到一起而产生修改的视频流。在一些实现中,内容生成模块256可以执行附加变换(例如抗混叠以改善修改的视频流的质量)。在一些实现中,图像或视频可以对应于视频/图像项280A-280Z之一,其可以从数据存储装置106提取。每个视频/图像项280A-280Z可以分别包括视频/图像数据284A-284Z,以及分别包括相关联的元数据284A-284Z。

[0037] 一般地,在一个实现中被描述为由会议服务器120或内容处理服务器130中的一个执行的功能,在其他实现中也可以由客户端设备110A-110Z中的一个或多个来执行,如果适当的话。例如,客户端设备110A可以实现软件应用,其执行内容变换部件250的功能。此

外,归因于特定部件的功能可以由一起操作的不同的或多个部件来执行。在一些实现中,内容处理服务器130可以作为通过适当应用编程接口提供给其他系统或服务的服务来被访问。

[0038] 会议平台200和内容变换部件250被描述为分别由会议服务器120和内容处理服务器130来实现,但是也可以由客户端设备110A-110Z、会议服务器120或内容处理服务器130中的任何一个来实现。作为示例,客户端设备(例如客户端设备110A)可以被编程,以执行会议平台200和/或内容变换部件250的一些或全部功能。作为另一示例,会议平台200和内容变换部件250可以在会议服务器中被合并到一起。在会议平台200和/或内容变换部件250在客户端设备上实现的实施方式中,针对会议平台200和/或内容变换部件250所述的“接收”、“发送”、“检索”、“识别”、“确定”等等的任何功能被理解为由客户端设备内的子系统或子模块所执行的功能,而不是跨越网络(例如网络105),本领域普通技术人员将认识到这一点。

[0039] 图3是图示说明根据公开的实现的媒体协作服务/应用的示例图形用户界面(GUI)窗口300的图。在一个实现中,GUI窗口300包括主区域302、缩略图区域304A-304C、标题区域306、聊天区域308以及选项区域310。在一些实现中,少于全部的缩略图区域304A-304C可以呈现在GUI窗口300中,或者附加的缩略图区域可以被呈现。可以生成GUI窗口300来用于显示,作为与客户端设备的媒体观看器相组合的用户界面(例如客户端设备110A的媒体观看器112A和用户界面111A)。

[0040] 在一些实现中,每个区域可以包含、描绘或者呈现媒体协作中的特定参与者所提供的媒体内容(例如视频内容)。例如,主区域302可以包含由第一参与者发送的视频流(例如,视频会议中第一或主参与者的“房间”、第一参与者所共享的视频剪辑等等),而每个缩略图区域304A-304C可以包含一个或多个附加参与者所发送的视频流的小型化版本、与附加参与者相关联的静态图像(例如头像/avatar等等)、与附加参与者相关联的视频流的小型化版本、或上述的组合。应该注意的是,尽管GUI窗口300中的区域被描绘为矩形,但是一个或多个区域可以具有一些其他形状(例如圆形、梯形等等)。此外,GUI窗口300的形状、尺寸和/或布局可以取决于要呈现该GUI窗口300的设备。例如,如图3中所图示,GUI窗口300具有的格式适合于移动设备。

[0041] 在一个实现中,主区域302和缩略图区域304A-304C中的每个可以与视频协作中的一个或多个参与者(例如视频会议参与者)的相应客户端设备所生成的视频流相关联。此外,在特定实现中,特定区域可以改变(例如从缩略图区域到主区域)或者切换在不同区域中显示的视频流,这取决于与该区域相关联的参与者是否正在讲话(例如,使用接口模块204)。这样的改变可以自动进行(例如,不接收任何用户请求或检测任何用户交互)。在一些实现中,会议平台200可以接收视频流,且识别(例如使用接口模块204)哪个视频流对应于当前讲话者或一般注意焦点(例如,视频流对应于正托管媒体协作的参与者,讲话者的音量水平,特定参与者的托管选择,等等),在一些实现中,实现GUI窗口300的客户端设备可以接收合成内容流(例如,其包括由客户端设备110A-110Z中的一个或多个所生成的视频/音频流数据),其由会议平台200的接口模块204生成。客户端设备可以被配置为从合成内容流提取内容(例如一个或多个视频/音频流)。在一个实现中,合成内容流包括元数据,其规定对应于GUI窗口300的区域的几何属性,使得视频流可以从合成内容流中提取并映射到GUI窗

口300内的适当区域。

[0042] 在一个实现中,缩略图区域304A-304C之一可以变换或转换为主区域302(例如,响应于参与者选择/点击该缩略图区域或激活该缩略图区域),反之亦然。在一些实现中,当特定参与者在视频会议期间讲话时,参与者的视频流可以显示在主区域302内。例如,如果对应于第一参与者的视频流显示在主区域302中且第一参与者不在讲话,第二参与者的视频流可以从缩略图区域304A-304C之一“提升”到主区域302,而第一参与者的视频流被“降级”到缩略图区域304A-304C之一。

[0043] 在一个实现中,合成内容流可以包括实时聊天数据。例如,聊天数据可以在聊天区域308内显示。聊天区域308可以向用户/参与者提供机会来输入聊天消息和发送聊天消息(例如,至会议服务器,其将聊天消息路由到其他参与者的客户端设备)。

[0044] 在一个实现中,选项区域310可以提供可选选项给用户。例如,可选选项可以允许用户调整媒体协作的设定(例如显示器特征、音量等等),邀请附加用户来参与,应用变换到接收到的/生成的视频流等等。与内容变换相关的选项结合图4-8来更详细地描述。

[0045] 图4是图示说明根据本公开的实现用于对视频会议中的视频流执行用户触发的变换的方法400的流程图。方法400可以通过包括硬件(例如电路、专用逻辑、可编程逻辑、微代码等等)、软件(例如运行在处理设备上以执行硬件仿真的指令)或者其组合的处理逻辑来执行。在一个实现中,方法400可以由如结合图1和2所述的内容变换部件250执行。在描述方法400时,对图5-8进行参考以便说明各种实现。应注意:图5-8中提供的示例是用于说明性目的,而不应被认为限制实现。

[0046] 方法400的操作可以根据系统架构的各种实现(例如系统架构100)而针对媒体协作来执行。在一个实现中,内容流,诸如由客户端设备(例如客户端设备110A-110Z)所生成/捕捉的视频流,可以被发送到会议服务器(例如会议服务器120的会议平台200),然后发送到内容处理服务器(例如内容处理服务器130的内容变换部件250)。内容流中的一个或多个被内容处理服务器变换且被发送(例如,作为合成内容流)到客户端设备。在另一实现中,内容流被发送到内容处理服务器,且内容流中的一个或多个被变换、发送到会议服务器,然后被发送到客户端设备(例如作为合成内容流)。在另一实现中,内容流被第一客户端设备(例如客户端设备110A)生成/捕捉,在该客户端设备处被变换(例如通过在客户端设备110A上本地实现内容变换部件250),且被发送到会议服务器120,或者直接到参与媒体协作的一个或多个附加客户端设备。在另一实现中,对变换视频流的指示可以从第一客户端设备发送到第二客户端设备,其中视频流的变换由第二客户端设备执行(例如,第二客户端设备实现内容变换部件250)。

[0047] 参看图4,方法400开始于框410,此时从视频会议的第一参与者接收对视频或图像的选择。例如,第一参与者可以经由客户端设备的用户界面(例如客户端设备110A的用户界面111A)选择视频或图像。在一个实现中,可以通过客户端设备将选择发送到内容处理服务器(例如,内容处理服务器130)。

[0048] 在一个实现中,用户界面(或媒体协作界面)可以对应于GUI窗口500,如图5中所示。GUI窗口500包括主区域502,其可以显示第一参与者的视频流。缩略图504、506、508可以包括视频会议中的附加参与者(例如与缩略图504相关联的第二参与者,以及与缩略图506相关联的第三参与者)的视频流。在一个实现中,主区域502的视频流的缩略图版本被显示

作为缩略图(例如,作为缩略图508)。在一些实现中,GUI窗口500可以包括聊天窗口(例如,聊天区域308,如图3中所图示)。GUI窗口500包括选项区域510,其包括可选选项,用于进行内容变换。例如,内容变换可以包括在视频会议期间将视频或图像插入到一个或多个视频流中,以提供被称为“photobomb(照片炸弹)”的滑稽效果。选项区域510包括图像/视频列表512,第一参与者可以从中选择图像或视频。例如,图像/视频列表512可以包括名人的视频或图像、动物的视频或图像、卡通角色的视频或图像、或其他视频或图像中的一种或多种。如所图示说明的,第一参与者从列表中选择了“martial artist(武术家)”,其可能对应于知名武术家的视频剪辑。GUI窗口还可以包括参与者列表516,其列出视频会议中的每个参与者。在一个实现中,第一参与者可以选择一个或多个参与者(包括对应于“myself(自己)”的第一参与者)。如所图示的,第一参与者已经选择了复选框518,其表示他/她希望变换他/她的视频流以包括在图像/视频列表512中选择的“martial artist”。在一些实现中,少于全部的选项可以被呈现在选项区域510中,或者附加选项可以呈现在选项区域510中。

[0049] 回头来参看图4,在框420,从视频会议的第一参与者接收修改视频流的请求。在一个实现中,在客户端设备接收请求,并且将请求发送到内容处理服务器的内容变换部件(例如,内容变换部件250)。在一个实现中,请求可以由客户端设备接收,形式为由第一参与者选择可选选项(例如,选择照片炸弹按钮520)。在一些实现中,第一参与者可以从图像/视频列表512中选择图像或视频,并且将所选图像或视频“拖拽”到GUI窗口500中显示视频流之一的区域(其可以由内容变换部件250解释为修改视频流的请求)。例如,第一参与者可以拖拽“martial artist”的指示符到在主区域502中显示的他/她的视频流中。在其他实现中,不同方法可以被用来接收对视频或图像的选择以及修改一个或多个视频流的请求(例如,语音激活、从口头或文本对话中提取的话题,等等)。

[0050] 在框430,在视频流内识别前景和背景。在一些实现中,如果视频流是现场视频流,框430对于现场视频流的每帧都执行。在一些实现中,框430由前景识别模块252来执行。如图6A中所图示,视频流的前景602被识别,且可以由将视频的一个或多个帧进行分割的边界604来限定。如图6B中所图示,背景610可以对应于视频流的剩余部分,其在边界604之外。在一个实现中,目标检测算法(例如,面部识别算法)可以由目标识别模块254用来识别位于前景602中的目标区域606。例如,目标区域606可以对应于面部(例如,第一参与者的面部)。前景602的相对位置(例如,对应于目标区域606的目标)可以基于视频帧内的目标区域606的位置608A和608B来限定。在一些实现中,目标(例如面部)可以被跟踪,使得位置608A和608B可在视频流里逐帧地变化。

[0051] 回头参看图4,在框404,包括插入到背景中的视频或图像的修改的视频流被生成,使得该视频或图像当被显示时出现在前景之后。在一个实现中,内容生成模块256将变换应用于视频流以产生修改的视频流。如图7A中所图示,对应于所选“martial artist”的视频帧被覆盖/叠加到图6B的背景610上,以产生修改的背景702。如图7中所图示的,图6A的前景602随后覆盖到修改的背景702上,以产生修改的视频帧704。在一些实现中,视频流的每个视频帧被变换,产生修改的视频流。

[0052] 在一些实现中,变换可以持续应用一段时间(例如,5秒、10秒、30秒等等的预定持续时间)。在一些实现中,如果视频流要被修改以包括背景中的图像(例如,静态图像),则在持续时间内视频流的每个帧可以被修改以包括静态图像。在一些实现中,如果视频流被修

改以包括背景中的视频,则视频流的每个帧可以被修改以包括视频的适当帧(例如,视频流的帧N可以被修改以包括视频的帧M,视频流的帧N+1可以被修改以包括视频的帧M+1,等等)。图8A和8B分别表示在应用变换之后与第一参与者相关联的修改的视频流的帧,根据框450,其在修改的视频流已经被发送到客户端设备时可出现在附加参与者的一个或多个客户端设备的相应GUI窗口内。例如,图8A和8B对应于视频会议的第一参与者的已经变换的视频流,其变换为产生武术家进入第一参与者的房间且给予强力回旋踢到第一参与者脑后的假象。在应用变换时,内容变换部件250可以考虑视频流的目标区域(例如,目标区域606),以便将武术家的视频插入到视频流的适当位置(例如,通过平移武术家的视频帧),使得对应于武术家靴子的视频的预定区域直接位于第一参与者的脑后。

[0053] 在一个实现中,附加变换可以应用于视频流。例如,如图8C中所图示的,修改的背景802以类似于结合图7A所述的方式生成,而视频流的前景可以覆盖在修改的背景802上以产生第一修改的视频帧804,如结合图7B所述。此外,视频的第一修改的视频帧部分806(例如对应于手)也可以覆盖在第一修改的视频帧804上以产生第二修改的视频帧808,如图8D中所图示。因此,变换可以包括出现在视频流的前景之后的一部分图像或视频和出现在前景之前的一部分图像或视频。在一些实现中,视频流的目标区域可以在确定要插入图像或视频的位置时被考虑(例如,用来定位武术家的视频,以便他的手出现在第一参与者的肩膀上)。

[0054] 为了简化解释,本公开的方法的各种实现被描绘且描述为一系列动作。但是,根据本公开的动作可以以各种次序和/或同时出现,以及与此未呈现和描述的其他动作一起出现。此外,并非所有图示的动作可能都是实现根据所公开主题的方法所需。此外,本领域技术人员将理解并认识到,方法可以可替换地被表示为经由状态图或事件的一系列相关的状态。此外,应该认识到,本说明书中所公开的方法的实现能够存储在制造物品中,以便于转移和传送这样的方法到计算设备。如这里所使用的,术语“制造物品”意图包含可从任何计算机可读设备或存储介质访问的计算机程序。

[0055] 图9以示例形式的计算机系统900来图示说明机器的图形表示,在计算机系统900中可以执行指令集,使得机器执行任何一个或多个这里所讨论的方法。在可替换实现中,机器可以连接到(例如联网到)LAN、内部网、外部网或互联网中的其他机器。机器可以以客户端-服务器网络环境中的服务器或客户端机器的资格操作,或者在对等式(或分布式)网络环境中作为对等机器。机器可以是个人电脑(PC)、平板PC、机顶盒(STB)、个人数字助理(PDA)、蜂窝电话、web应用、服务器、网络路由器、交换机或桥接器、或者能够执行规定要由该机器采取的行动的指令集(顺序的或其他的)的任何机器。此外,尽管只有一个机器被图示,术语“机器”还应该被认为包括任何机器集合,其单个或联合执行指令集(或多个指令集),以执行任何一个或多个这里所讨论的方法。计算机系统900的一些或全部部件可以由客户端设备110A-110Z、数据存储装置106、会议服务器120或内容处理服务器130中的一个或多个使用或者图示。

[0056] 示例计算机系统900包括处理设备(处理器)902、主存储器904(例如只读存储器(ROM)、闪存存储器、诸如同步DRAM(SDRAM)或Rambus DRAM(RDRAM)的动态随机存取存储器(DRAM)等等)、静态存储器906(例如闪存存储器、静态随机存取存储器(SRAM)等等)、以及数据存储设备918,其经由总线908互相通信。

[0057] 处理器902表示一个或多个通用处理设备,诸如微处理器、中央处理单元等。更特别地,处理器902可以是复杂指令集计算(CISI)微处理器、精减指令集计算(RISC)微处理器、超长指令字(VLIW)微处理器、或者实现其他指令集的处理器、或实现指令集组合的处理器。处理器902还可以是一个或多个专用处理设备,诸如专用集成电路(ASIC)、现场可编程门阵列(FPGA)、DSP、网络处理器等。处理器902被配置为执行指令926,以执行这里所讨论的操作和步骤。

[0058] 计算机系统900可以进一步包括网络接口设备922。计算机系统900还可以包括视频显示单元910(例如液晶显示器(LCD)、阴极射线管(CRT)或触摸屏)、字母数字输入设备912(例如键盘)、光标控制设备914(例如鼠标)以及信号生成设备920(例如扬声器)。

[0059] 数据存储设备918可以包括计算机可读存储介质924,其上存储一个或多个指令集926(例如软件),体现任何一个或多个这里所述的方法或功能。在其由计算机系统900执行期间,指令926还可以完全或至少部分地驻留在主存储器904内和/或处理器902内,主存储器904和处理器902也构成计算机可读存储介质。指令926可以进一步在网络974(例如网络105)上经由网络接口设备922被发送或接收。

[0060] 在一个实现中,指令926包括用于一个或多个内容变换部件250的指令,其可以对应于关于图1和2所述的相同命名的对应物。尽管计算机可读存储介质924在示例实现中被示出为单个介质,但术语“计算机可读存储介质”或“机器可读存储介质”应该被认为包括一个介质或多个介质(例如,集中式或分布式数据库,和/或相关联的缓存和服务),其存储一个或多个指令集。术语“计算机可读存储介质”或“机器可读存储介质”还应被认为包括任何瞬时或非瞬时介质,其能够存储、编码或携带指令集以便由机器执行,且使得机器执行任何一个或多个本公开的方法。术语“计算机可读存储介质”应该因此被认为包括(但不限于)固态存储器、光介质、以及磁介质。

[0061] 在之前的描述中,阐述了许多细节。但是,对于从本公开中受益的本领域普通技术人员来说很显然的是,本公开可以不用这些具体细节而实践。在一些情形中,公知结构和设备以框图形式示出,而不是细节,以便避免对本公开的混淆。

[0062] 详细描述的一些部分已经根据在计算机存储器内针对数据比特的操作的算法和符号表示进行了呈现。这些算法描述和表示是数据处理领域的技术人员用来最有效地将其工作的实质传递给本领域其他技术人员所使用的手段。算法在此通常被构思为导致所需结果的步骤的自治序列。这些步骤是需要物理量的物理操作的步骤。通常,尽管并非必要,这些量采用了能够被存储、传输、合并、比较和其他操作的电或磁信号的形式。主要出于通常使用的原因,已证明有时将这些信号作为比特、值、元素、符号、字符、项、数字等是便利的。

[0063] 但是,应该记住的是,所有这些和类似术语要与适当物理量相关联,且仅仅是施加于这些量的便利标签。除非从前文讨论中显然看出的具体陈述,应该认识到,在整个描述中,利用了诸如“接收”、“发送”、“生成”、“添加”、“减少”、“插入”、“去除”、“分析”、“确定”、“启用”、“识别”、“修改”等术语的讨论,指的是计算机系统的动作和过程,或者类似的电子计算设备的动作和过程,其操作和变换被表示为计算机系统的寄存器和存储器内的物理(例如电子)量的数据成为其他类似地被表示为计算机系统存储器或寄存器内或其他这样的信息存储、传输或显示设备内的物理量的数据。

[0064] 本公开还涉及用于执行这里的操作的装置、设备或者系统。该装置、设备或系统可

以特别构建用于所需用途,或者其可以包括通用计算机,可选择地由存储在计算机中的计算机程序激活或重新配置。这样的计算机程序可以存储在计算机或机器可读存储介质中,诸如(但不限于)任何类型的盘,包括软盘、光盘、只读光盘存储器(CD-ROM)、以及磁光盘、只读存储器(ROM)、随机存取存储器(RAM)、EPROM、EEPROM、磁或光卡、或者任何类型的适合存储电子指令的媒体。

[0065] 单词“示例”或“示例性”在此用于表示用作示例、例子或图示。这里被描述为“示例”或“示例性”的任何方面或设计不必被解释为相比其他方面或设计是优选的或有利的。而是,单词“示例”或“示例性”的使用是希望呈现具体形式的概念。如本说明书中使用的,术语“或”意味着包含性的“或”而非排他性的“或”。也就是,除非另外指出,或者从上下文中明显看出,“X包括A或B”意味着任何自然的包含性排列。也就是,如果X包括A;X包括B;或者X包括A和B二者,则“X包括A或B”在任何前述情况下都是满足的。此外,冠词“一”和“一个”在本说明书和所附权利要求中的使用应该通常被解释为意味着“一个或多个”,除非另外指出或从上下文中明显看出是指单数形式。本说明书中通篇对“实现”或“一个实现”的引用,意味着结合该实现所描述的特定特征、结构或特性包括在至少一个实现中。因此,词组“实现”或“一个实现”在本说明书通篇各种地方的出现不必都指相同的实现。此外,应注意:在对附图的特定元素的引用中所使用的“A-Z”的注释并不希望被限制为特定数量的元素。因此,“A-Z”应被解释为具有呈现在特定实现中的一个或多个元素。

[0066] 应该理解,上面的描述意在是说明性的,而非限制性的。在阅读和理解上面的描述之后,许多其他实现对于本领域技术人员而言是显然的。因此,本公开的范围应该结合所附权利要求以及与这样的权利要求等价的全部范围而确定。

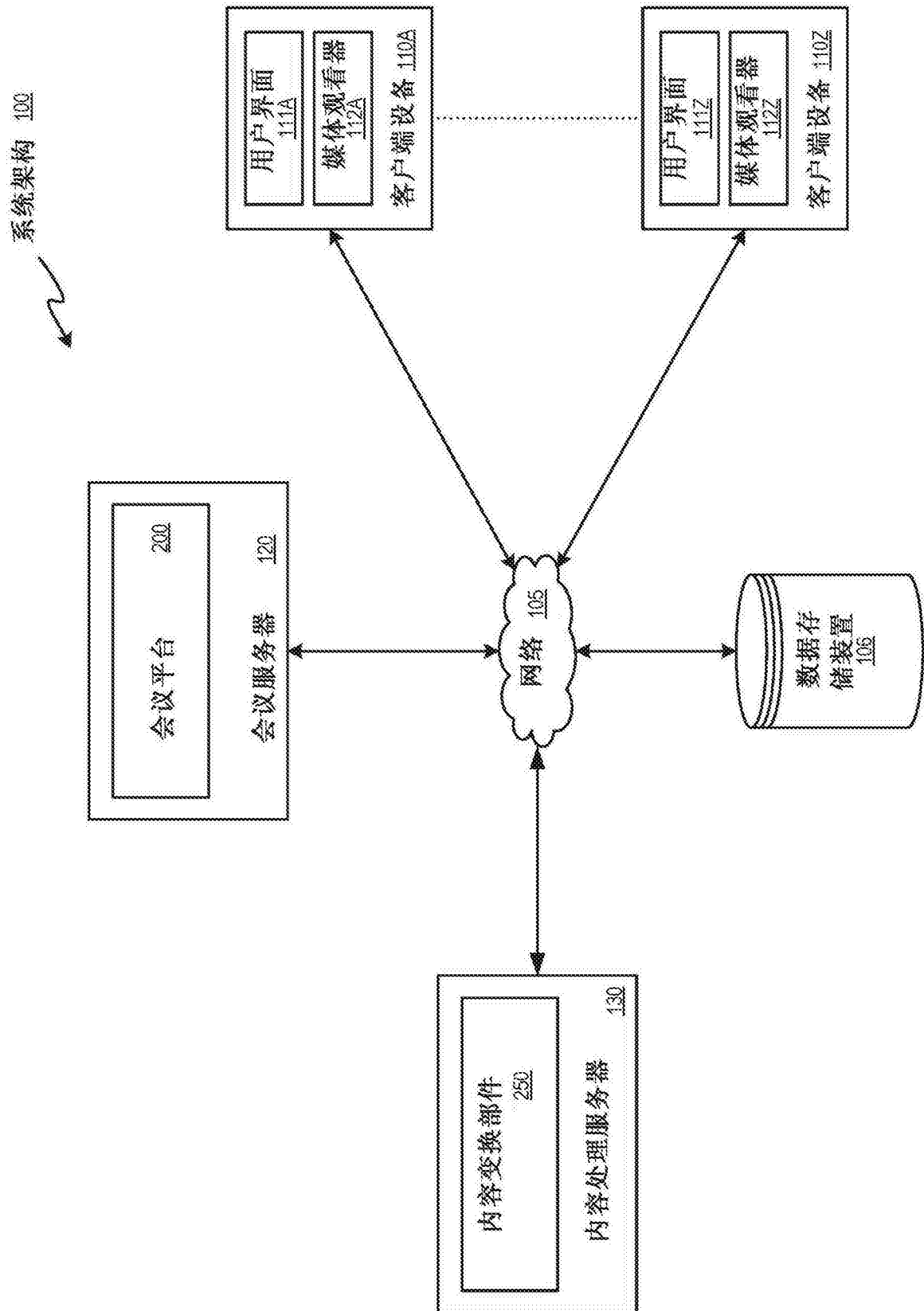


图1

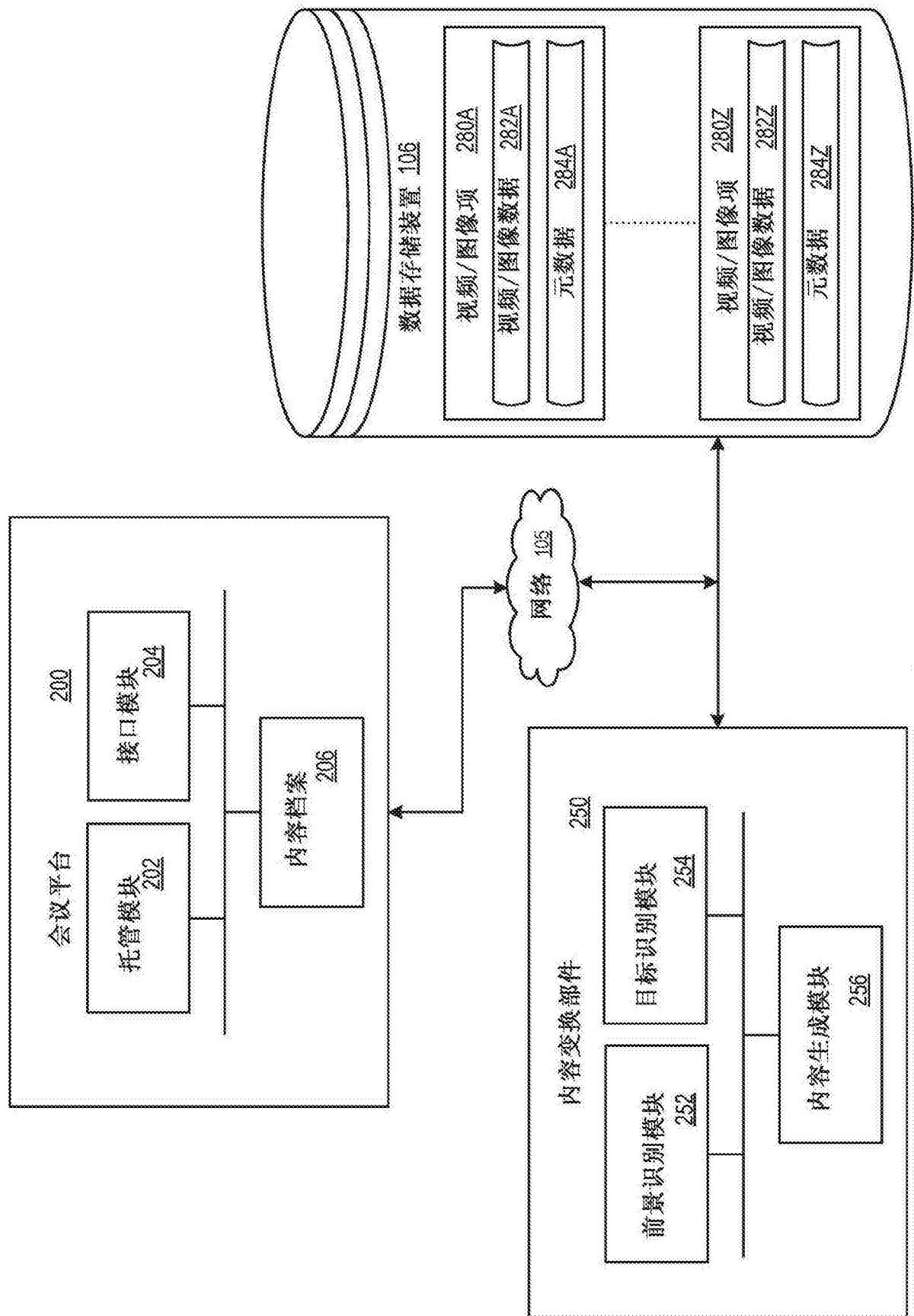


图2

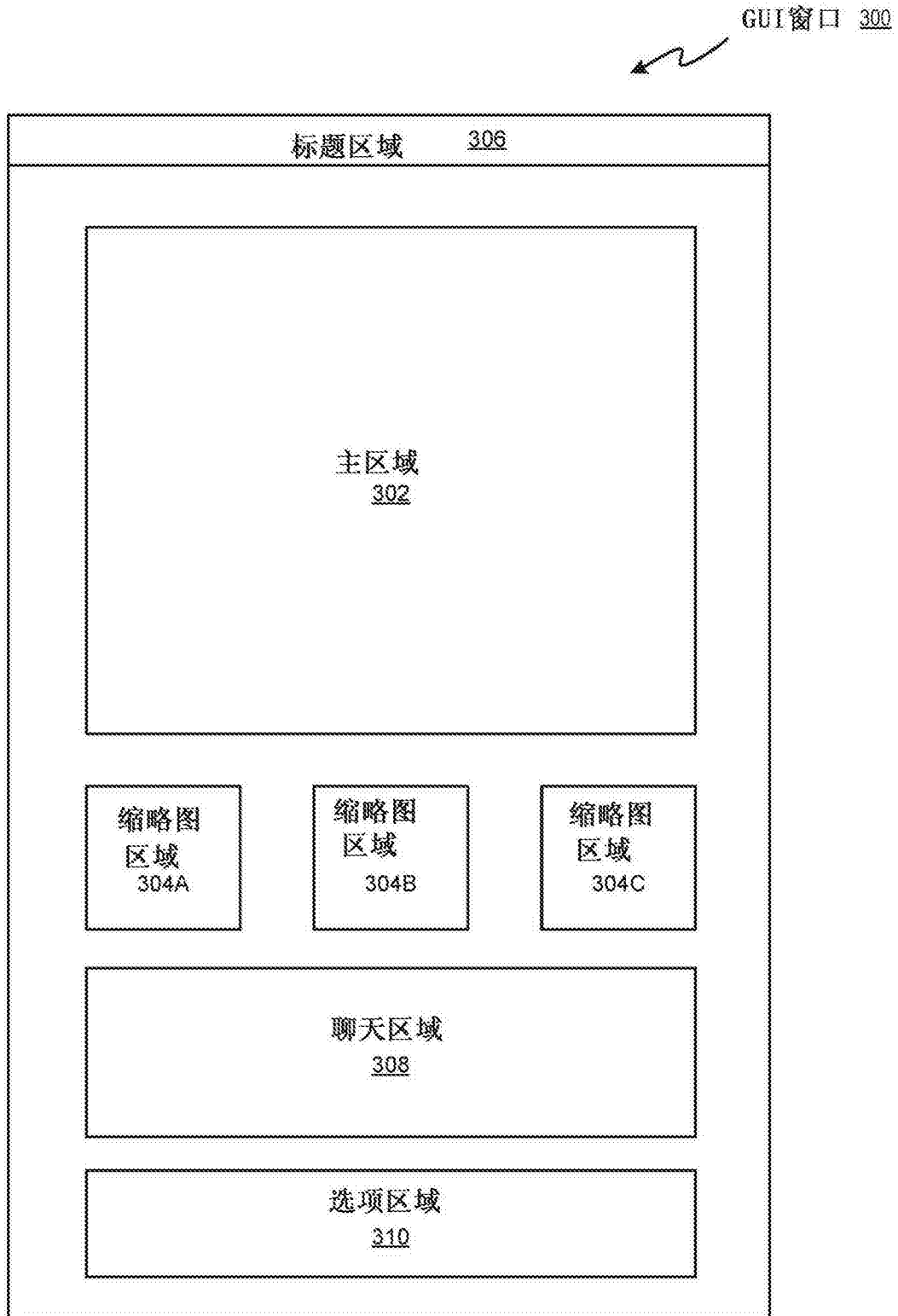


图3

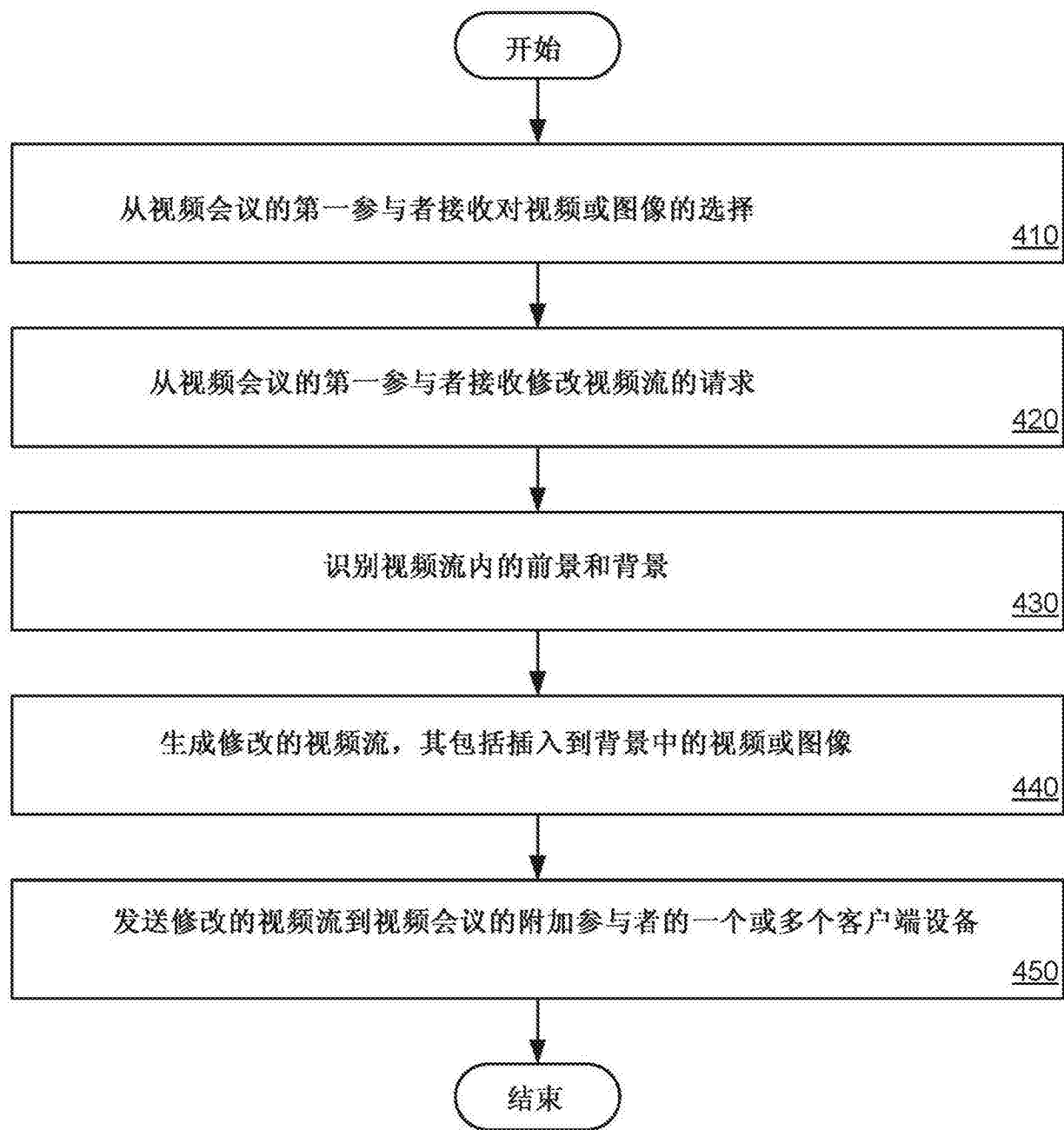
400

图4

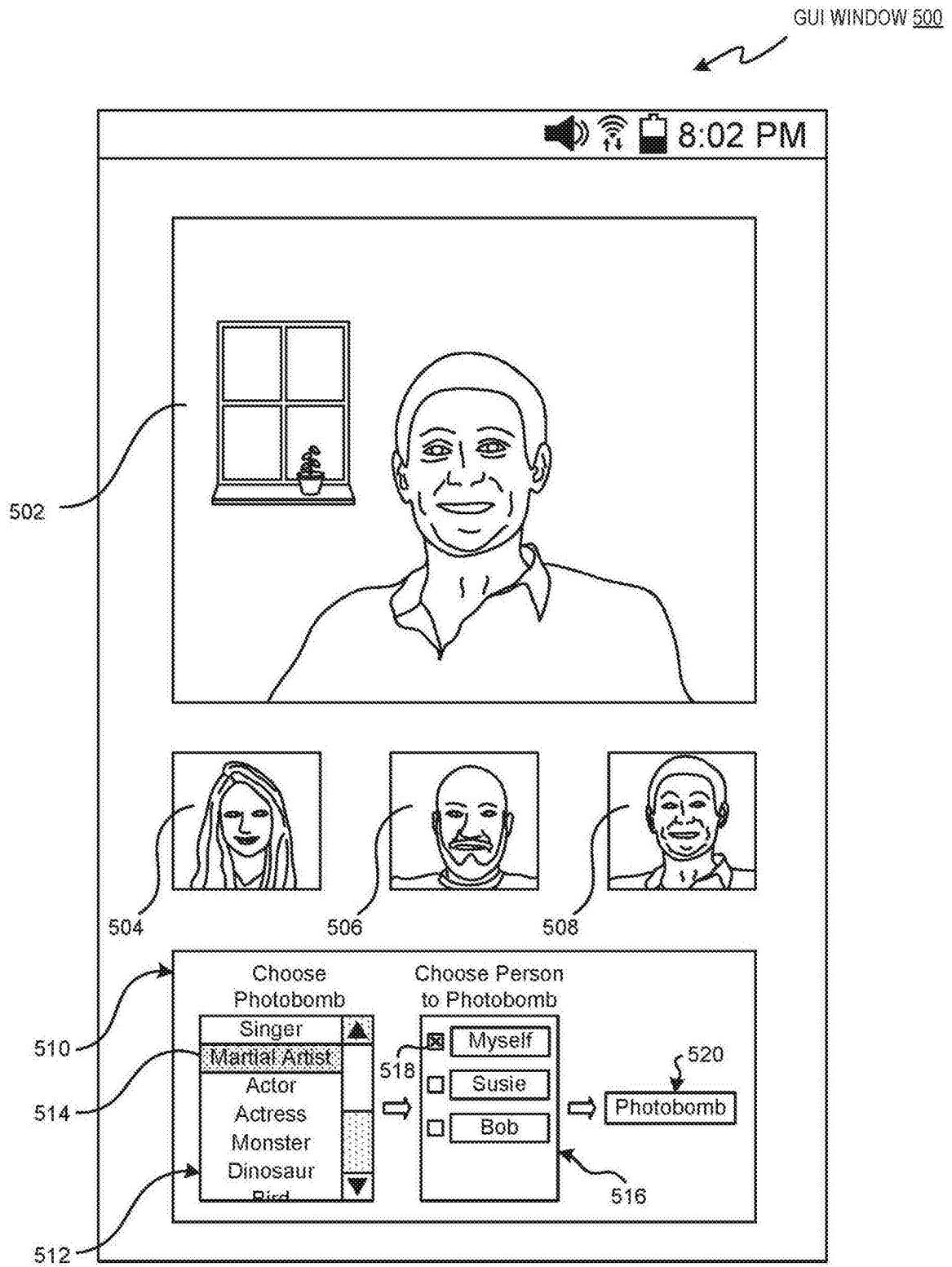


图5

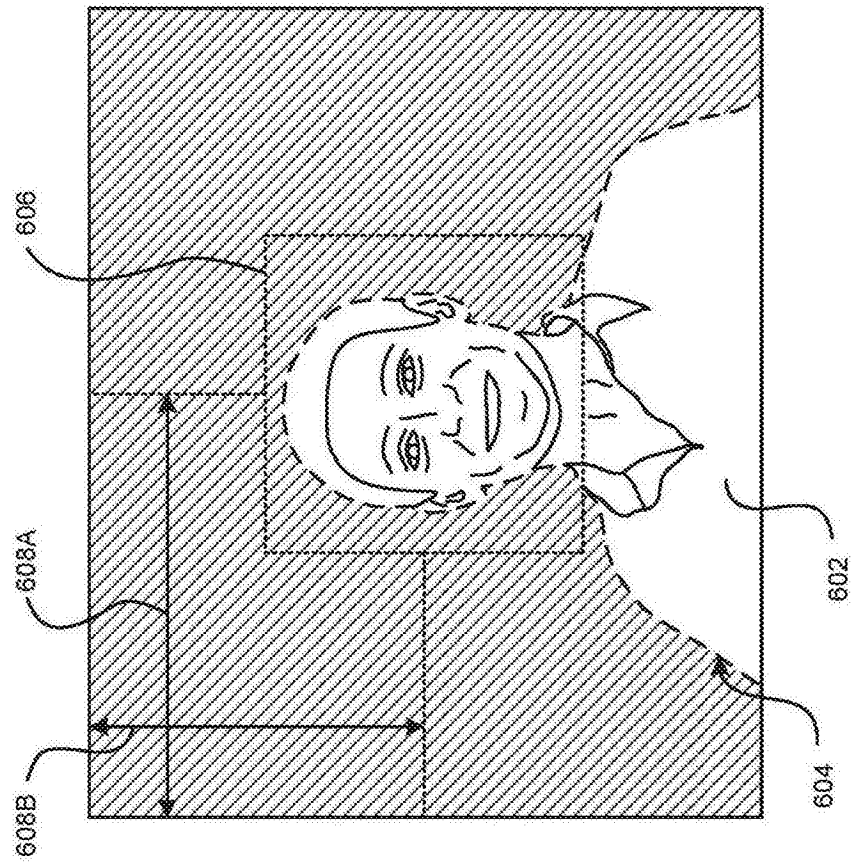


图6A

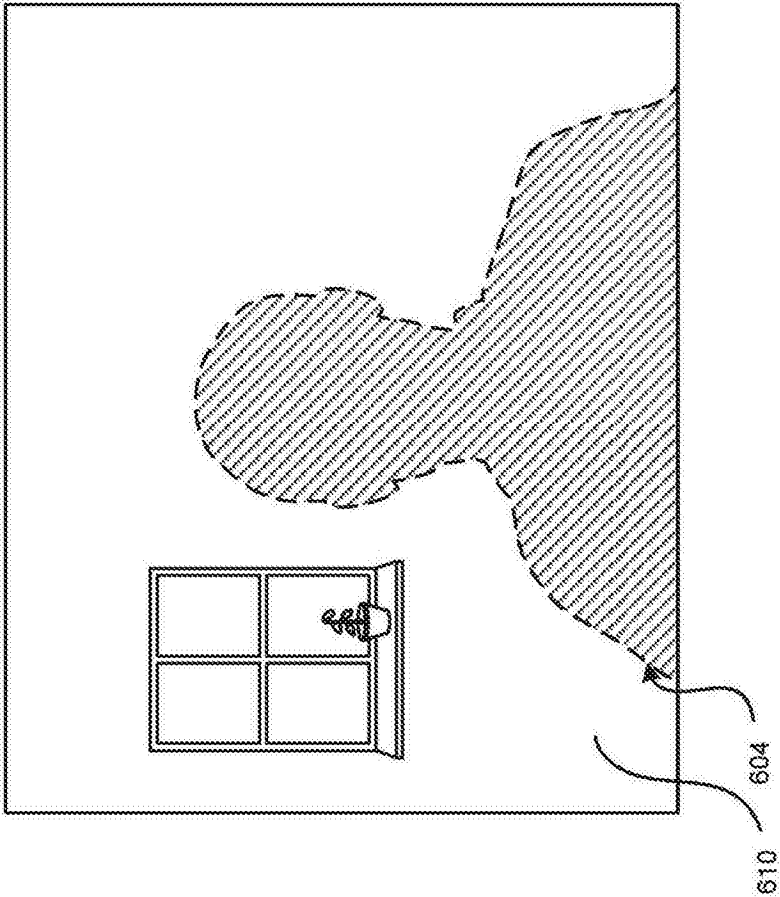


图6B

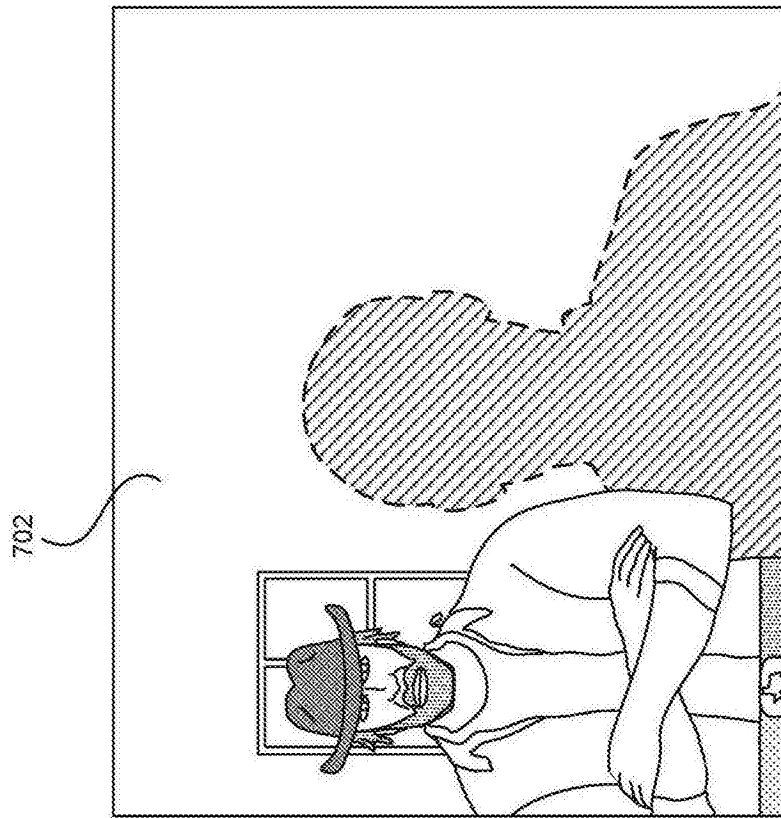


图7A

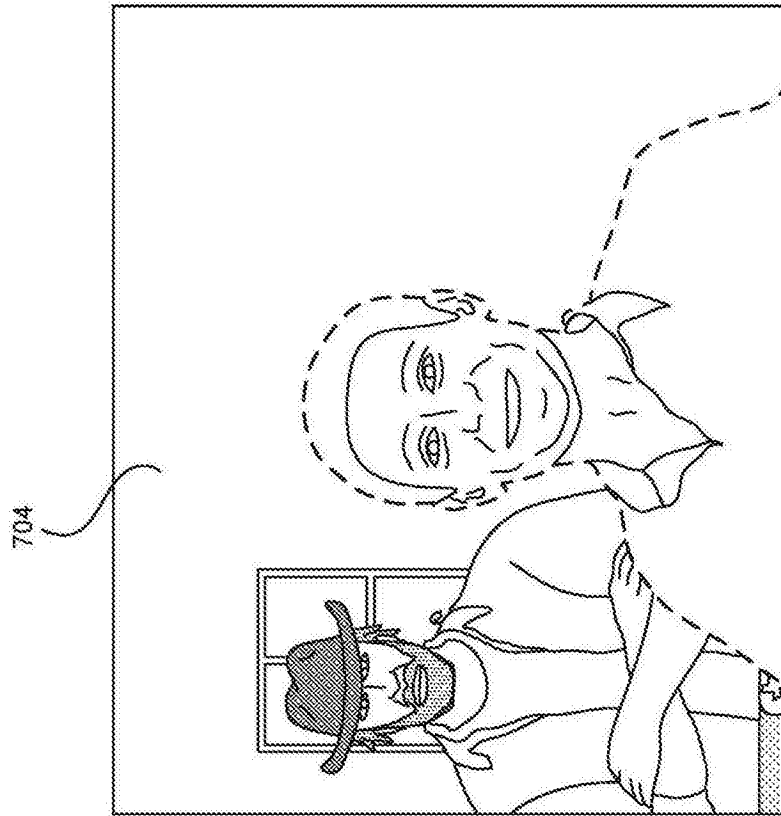


图7B

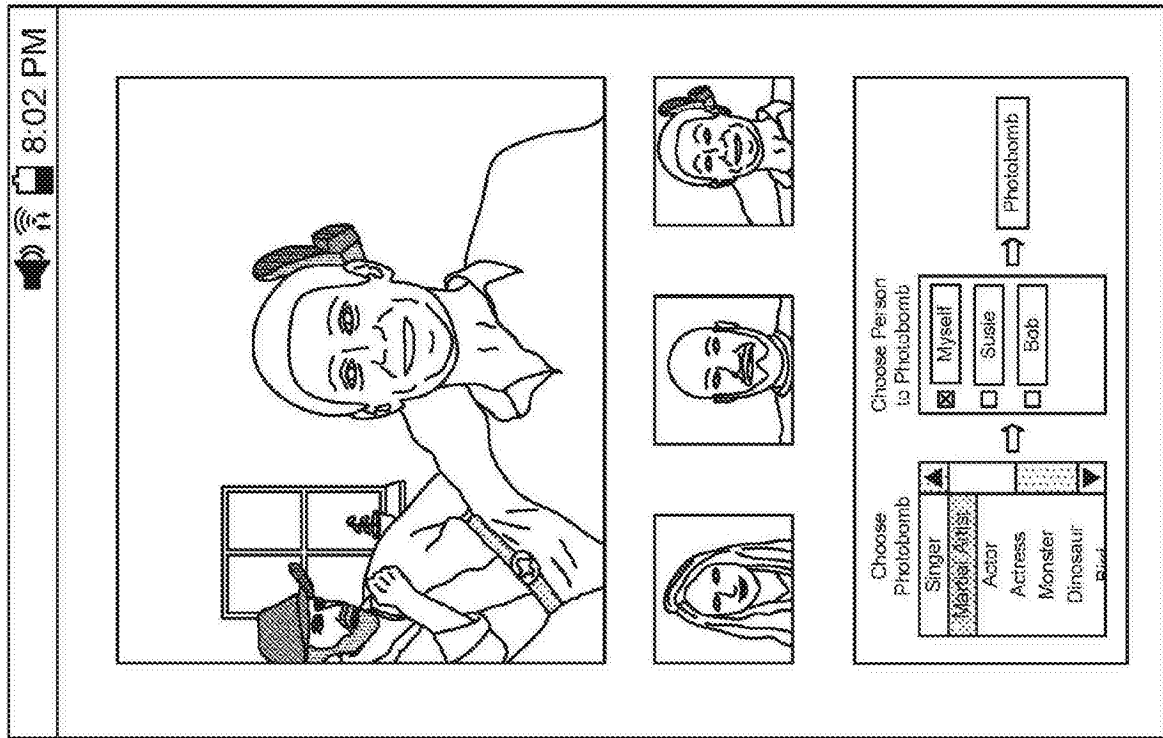


图 8B

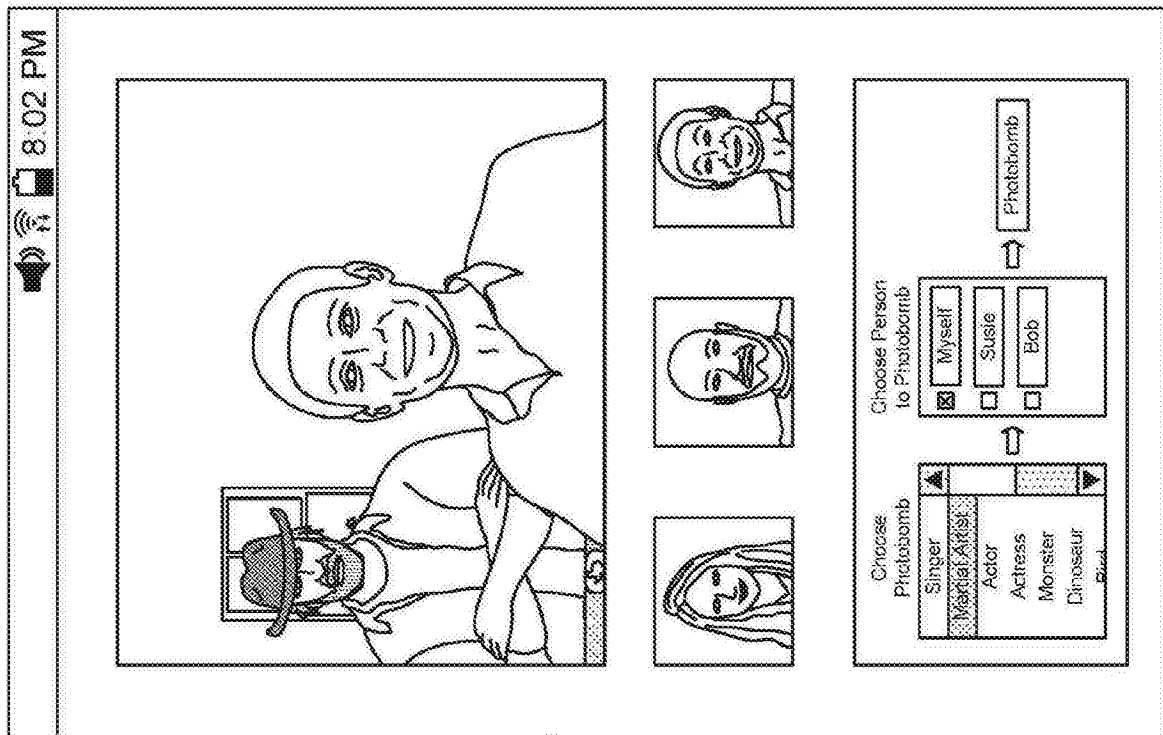


图 8A

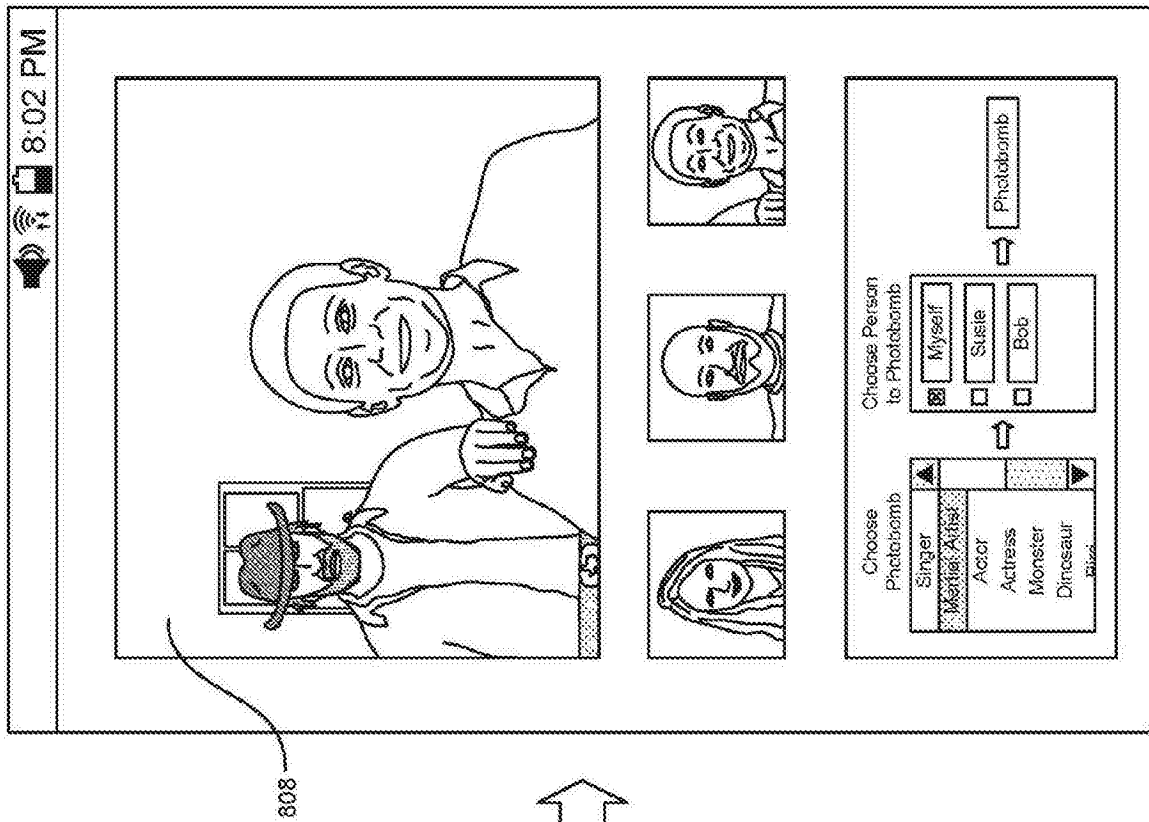


图8D

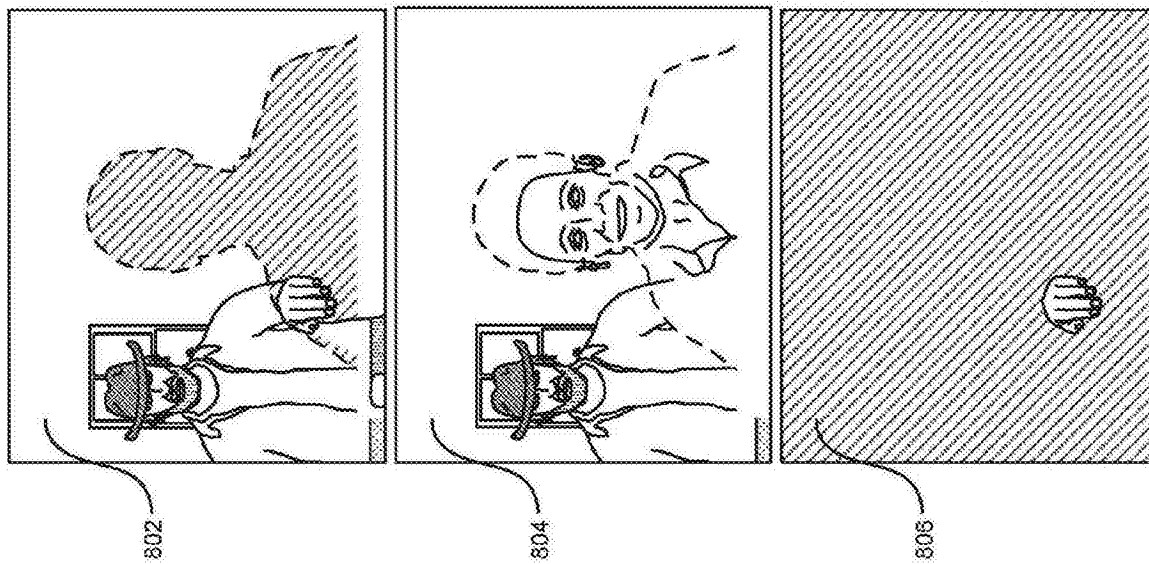


图8C

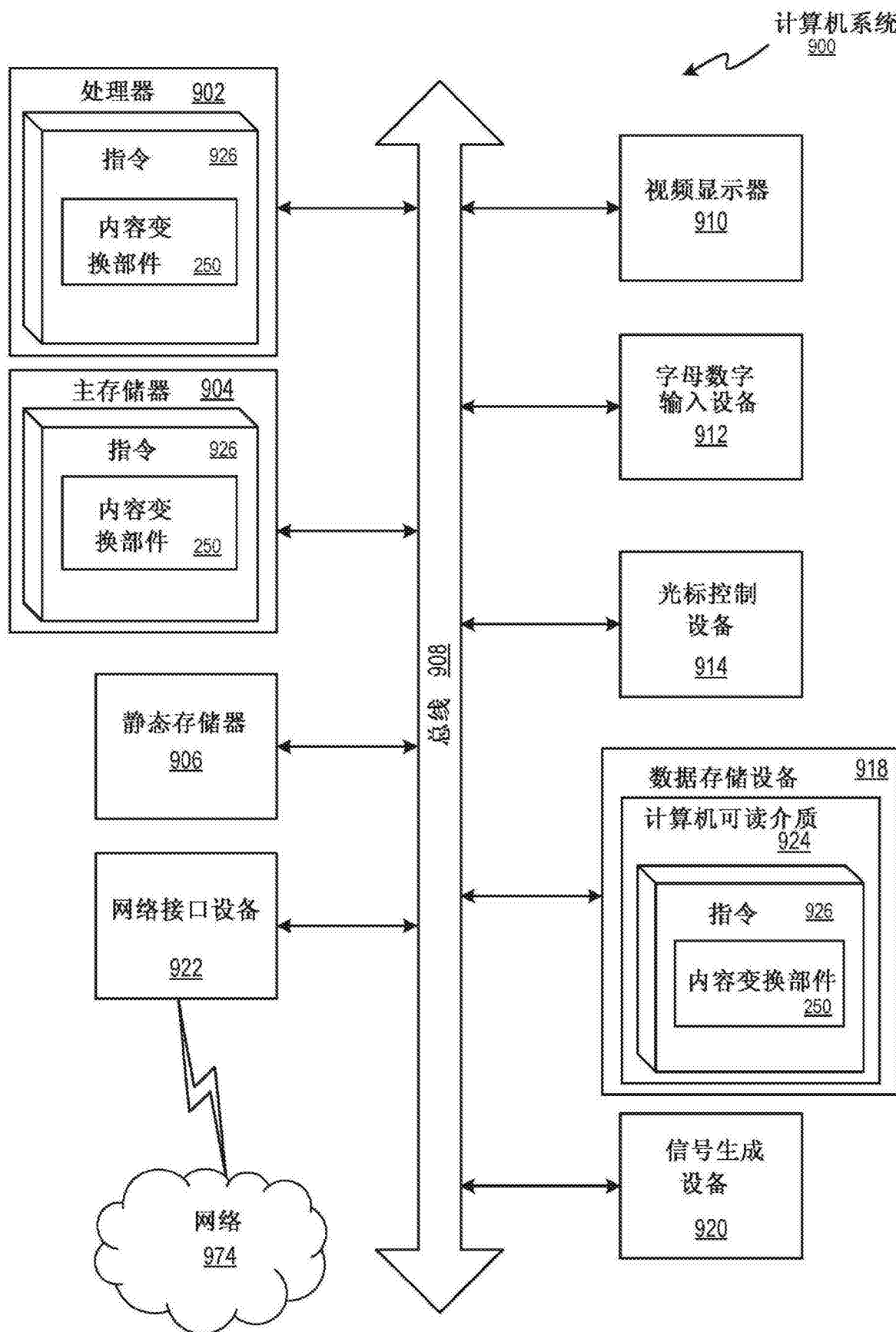


图9