

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022 年 3 月 31 日 (31.03.2022)



(10) 国际公布号
WO 2022/062896 A1

(51) 国际专利分类号:
H04N 21/2187 (2011.01)

(21) 国际申请号: PCT/CN2021/117040

(22) 国际申请日: 2021 年 9 月 7 日 (07.09.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011001739.6 2020 年 9 月 22 日 (22.09.2020) CN

(71) 申请人: 北京达佳互联信息技术有限公司(BEIJING
DAJIA INTERNET INFORMATION TECHNOLOGY
CO., LTD.) [CN/CN]; 中国北京市海淀区上地西路
6 号 1 幢 1 层 101D1-7, Beijing 100085 (CN)。

(72) 发明人: 张水发 (ZHANG, Shuifa); 中国北京
市海淀区上地西路 6 号 1 幢 1 层 101D1-7,
Beijing 100085 (CN)。

(74) 代理人: 北京律智知识产权代理有限公司 (BEIJING INTELLEGAL INTELLECTUAL

PROPERTY AGENT LTD.); 中国北京市朝阳区
慧忠路 5 号 B1605、B1606、B1607, Beijing
100101 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家
保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,
BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU,
CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT,
JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,
LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,

(54) Title: LIVENSTEAMING INTERACTION METHOD AND APPARATUS

(54) 发明名称: 直播互动方法及装置

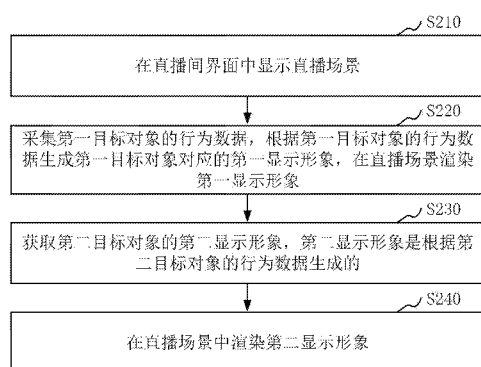


图2

S210 Display a live streaming scene in a live streaming room interface
S220 Collect behavior data of a first target object, generate, according to the behavior data of the first target object, a first display image corresponding to the first target object, and render the first display image in the live streaming scene
S230 Obtain a second display image of a second target object, the second display image being generated according to behavior data of the second target object
S240 Render the second display image in the live streaming scene

(57) Abstract: A live streaming interaction method and apparatus, an electronic device, and a storage medium. The method comprises: displaying a live streaming scene in a live streaming room interface; collecting behavior data of a first target object, generating, according to the behavior data of the first target object, a first display image corresponding to the first target object, and rendering the first display image in the live streaming scene; obtaining a second display image of a second target object, the second display image being generated according to behavior data of the second target object; and rendering the second display image in the live streaming scene.

(57) 摘要: 一种直播互动方法、装置、电子设备及存储介质。所述方法包括: 在直播间界面中显示直播场景; 采集第一目标对象的行为数据, 根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象, 在直播场景渲染第一显示形象; 获取第二目标对象的第二显示形象, 第二显示形象是根据第二目标对象的行为数据生成的; 在直播场景中渲染第二显示形象。

RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

直播互动方法及装置

相关申请的交叉引用

5 本公开基于申请日为2020年9月22日、申请号为202011001739.6号的中国专利申请，并要求该中国专利申请的优先权，在此全文引用上述中国专利申请公开的内容以作为本公开的一部分。

技术领域

10 本公开涉及互联网技术领域，尤其涉及一种直播互动方法、装置、电子设备及存储介质。

背景技术

互动直播是视频直播的增强应用，是在视频直播中增加互动功能。

相关技术中，互动直播中的互动功能包括在视频直播中增加语音、视频的互动。

15

发明内容

本公开提供一种直播互动方法、装置、电子设备及存储介质。本公开的技术方案如下：

根据本公开实施例的第一方面，提供一种直播互动方法，包括：

在直播间界面中显示直播场景；

20

采集第一目标对象的行为数据，根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象，在直播场景渲染第一显示形象；

获取第二目标对象的第二显示形象，第二显示形象是根据第二目标对象的行为数据生成的；

在直播场景中渲染第二显示形象。

25

根据本公开实施例的第二方面，提供一种直播互动装置，包括：

显示模块，被配置为在直播间界面中显示直播场景；

采集模块，被配置为采集第一目标对象的行为数据；

显示形象生成模块，被配置为根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象；

30

第一渲染模块，被配置为在直播场景渲染第一显示形象；

获取模块，被配置为获取第二目标对象的第二显示形象，第二显示形象是根据第二目标对象的行为数据生成的；

第二渲染模块，还被配置为在直播场景中渲染第二显示形象。

根据本公开实施例的第三方面，提供一种电子设备，包括：

35

处理器；用于存储所述处理器可执行指令的存储器；

其中，所述处理器被配置为执行所述指令，以实现第一方面的任一项实施例中所述的直播互动方法。

根据本公开实施例的第四方面，提供一种存储介质，当所述存储介质中的指令由电子设备的处理器执行时，使得所述电子设备能够执行第一方面的任一项实施例中所述的直播互动方法。

根据本公开实施例的第五方面，提供一种计算机程序产品，所述程序产品包括计算机程序，所述计算机程序存储在可读存储介质中，设备的至少一个处理器从所述可读存储介质读取并执行所述计算机程序，使得设备执行第一方面的任一项实施例中所述的直播互动方法。

根据本公开的方案，预先建立直播场景，在主播端和观众端显示同一个直播场景；主播端和观众端同时采集各自用户的行为数据生成显示形象，并同时双向传播显示形象，使主播端和观众端能够在同一虚拟场景中以真实世界行为进行直播互动，从而使直播互动方式更加全面。

附图说明

图 1 是根据一示例性实施例示出的一种直播互动方法的应用环境图。

图 2 是根据一示例性实施例示出的一种直播互动方法的流程图。

图 3 是根据一示例性实施例示出的一种采集行为数据步骤的流程图。

图 4 是根据一示例性实施例示出的一种直播互动方法的流程图。

图 5 是根据一示例性实施例示出的一种直播场景的示意图。

图 6 是根据另一示例性实施例示出的一种直播互动方法的流程图。

图 7 是根据一示例性实施例示出的一种直播互动装置的框图。

图 8 是根据一示例性实施例示出的一种电子设备的内部结构图。

具体实施方式

为了使本领域普通人员更好地理解本公开的技术方案，下面将结合附图，对本公开实施例中的技术方案进行清楚、完整地描述。

需要说明的是，本公开的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本公开的实施例能够以除了在这里图示或描述的那些以外的顺序实施。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本公开的一些方面相一致的装置和方法的例子。

需要说明的是，本公开实施例中所描述的获取用户信息以及用户账户的相关信息，包括社交关系身份信息之类的，均已获得用户许可，在取得用户充分许可授权的前提下，本

公开所涉及的方法，装置，设备，存储介质可以获取用户的相关信息。

相关技术中，在互动过程中通常只能对主播端影像进行处理，并由主播端向观众端进行单向展示主播端影像，存在互动方式单一的问题。

据此，本公开提供一种直播互动方法。

5 本公开所提供的直播互动方法，可以应用于如图 1 所示的应用环境中。其中，主播端 110 和服务器 120 通过网络进行通信，至少一个观众端 130 和服务器 120 通过网络进行通信。观众端 130 中至少包含参与直播互动的观众端（以下称为互动观众端）。主播端 110 中安装有能够用于进行直播的应用程序。观众端 130 中安装有能够用于观看直播的应用程序。主播端 110 中安装的用于进行直播的应用程序和观众端 130 中安装的用于观看直播的应用程序可以是相同的应用程序。主播端 110 创建直播间时，可以获取主播选择的直播场景素材，建立直播间。主播端 110 进行直播的过程中，主播端 110 采集主播的行为数据，根据主播的行为数据生成主播对应的主播显示形象，在直播场景渲染主播显示形象。在主播端 110 进行直播的过程中，观众端 130 进入该直播间，并在观众端 130 的屏幕上显示包含主播显示形象的直播场景。观众端 130 中的部分或者全部观众端（互动观众端）可以向
10 主播端 110 请求进行直播互动。互动观众端采集互动观众的行为数据，根据互动观众的行为数据生成互动观众对应的互动观众显示形象，在直播场景渲染互动观众显示形象。互动观众端将互动观众显示形象发送至服务器 120，以使服务器 120 将互动观众显示形象发送至主播端 110 以及未参与互动的其他观众端，使主播端 110 以及未参与互动的其他观众端在直播场景中渲染互动观众显示形象。其中，主播端 110 可以但不限于是各种个人计算机、
15 笔记本电脑、智能手机、平板电脑，服务器 120 可以用独立的服务器或者是多个服务器组成的服务器集群来实现，观众端 130 可以但不限于是各种个人计算机、笔记本电脑、智能手机、平板电脑。

图 2 是根据一示例性实施例示出的一种直播互动方法的流程图，如图 2 所示，直播互动方法用于图 1 中的主播端 110 或者观众端 130 中的互动观众端，包括以下步骤。

25 在步骤 S210 中，在直播间界面中显示直播场景。

其中，直播场景是指为直播间设置的虚拟场景。直播场景的素材可以预先配置，例如，可以是游戏场景、虚拟图像背景等，或者可以是用户在终端设备的相册中选择得到；或者通过图像采集设备实时拍摄图像得到，在此不做限定。在一些实施例中，主播可以通过主播端触发直播间的创建请求。主播端响应于直播间的创建请求，获取直播场景的素材；根据所获取的直播场景的素材创建直播场景。主播端显示已创建的直播场景。观众端可以通过搜索、热点推荐等方式进入该直播间，在观众端的屏幕上显示与主播端相同的直播场景。

在步骤 S220 中，采集第一目标对象的行为数据，根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象，在直播场景渲染第一显示形象。

其中，第一目标对象可以是主播或者参与直播互动的互动观众。互动观众可以是正在
35 观看直播的全部或者部分观众。在一些实施例中，对于第一目标对象对应的第一客户端，

实时通过图像采集设备采集第一目标对象的行为数据。对第一目标对象的行为数据进行相应的处理，生成第一目标对象对应的第一显示形象，并在第一客户端的直播场景渲染第一显示形象。第一目标对象的行为数据不限于是第一目标对象的视频数据、语音数据或者文字评论数据等。第一目标对象对应的第一显示形象可以基于深度学习理论得到。

5 示例性地，若第一目标对象的行为数据是对第一目标对象进行拍摄得到的行为图像，第一显示形象则可以是对行为图像进行语义分割处理得到的第一目标对象图像，也可以是通过第一目标对象的人体姿态估计结果驱动的三维模型；若第一目标对象的行为数据是对第一目标对象进行语音采集得到的语音数据，第一显示形象则可以是对语音数据进行语音识别得到的相关文字内容。

10 在步骤 S230 中，获取第二目标对象的第二显示形象，第二显示形象是根据第二目标对象的行为数据生成的。

在步骤 S240 中，在直播场景中渲染第二显示形象。

其中，第二目标对象可以是主播或者参与直播互动的互动观众。当第一目标对象是主播时，第二目标对象可以是互动观众；当第一目标对象是互动观众时，第二目标对象可以是主播和/或其他互动观众。在一些实施例中，对于第二目标对象对应的第二客户端，可以参照步骤 S220，根据第二目标对象的行为数据生成第二目标对象对应的第二显示形象，并在第二客户端显示的直播场景中渲染第二显示形象。第二客户端将已获取的第二显示形象发送至服务器，通过服务器将第二目标对象对应的第二显示形象发送第一客户端。第一客户端接收服务器发送的第二目标对象的第二显示形象，在显示的直播场景中渲染第二显示形象。

同理，对于第二目标对象对应的第二客户端，可以从服务器接收第一目标对象对应的第一显示形象，并在第二客户端显示的直播场景中渲染第一显示形象，从而使第二客户端与第一客户端呈现相同的直播场景。

25 进一步地，对于未参与直播互动的账户对应的观众端，可以从服务器获取第一目标对象对应的第一显示形象，以及第二目标对象对应的第二显示形象，并在观众端显示的直播场景中渲染第一显示形象和第二显示形象，从而使未参与直播互动的观众端、第一客户端和第二客户端呈现相同的直播场景。

上述直播互动方法中，预先建立直播场景，在主播端和观众端显示同一个直播场景；主播端和观众端同时采集各自用户的行为数据生成显示形象，并同时进行双向传播显示形象，使主播端和观众端能够在同一虚拟场景中以真实世界行为进行直播互动，从而可以使直播互动方式更加全面。

30 在一示例性实施例中，在步骤 S220 中，采集第一目标对象的行为数据，根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象，在直播场景渲染第一显示形象，包括：采集第一目标对象的多帧行为图像，对每帧行为图像进行语义分割处理，得到
35 每帧第一显示形象，在直播场景中渲染每帧第一显示形象。

在一些实施例中，第一目标对象的行为数据可以通过图像采集设备实时采集到的第一目标对象的连续多帧行为图像。第一客户端每获取一帧行为图像，调用预先配置的已训练的语义分割模型。通过已训练的语义分割模型对每帧行为图像进行语音分割处理，得到第一目标对象图像，将所得到的第一目标对象图像作为第一显示形象。第一客户端在直播场景中渲染所获取的每帧第一显示形象。其中，语义分割模型不限于采用 DeepLab（一种语义分割网络）、FCN（Fully Convolution Networks，全卷积网络）、SegNet（Sementic Segmentation，语义分割网络）、BiSeNet（Bilateral Segmentation Network for Real-time Semantic Segmentation，一种双通道实时语义分割模型）等。在一些实施例中，当第一目标对象是主播或者互动观众时，第一目标对象图像则可以是对应的真实主播人像或者真实互动观众人像。

进一步地，对于第一客户端所获取的第二目标对象的第二显示形象，同样可以上述方式对第二目标对象的每帧行为图像进行语义分割处理得到的第二显示形象。服务器将语义分割处理得到的第二显示形象发送至第一客户端，以使第一客户端能够在直播场景中渲染第二显示形象。

一些实施例中，通过采集参与直播互动的主播和/或互动观众的行为图像，对所得到的行为图像进行语义分割处理，得到真实人像，并在直播场景中渲染所得到的真实人像，使得虚拟直播场景更接近真实世界场景，从而可以提升直播互动的真实性，有助于提升用户在直播间的停留时间、提高直播应用的用户留存率。

在一示例性实施例中，对每帧行为图像进行语义分割处理，包括：将多帧行为图像发送至服务器；接收服务器发送的对每帧行为图像进行语音分割处理得到的每帧第一显示形象。

在一些实施例中，对第一客户端和/或第二客户端采集到的多帧行为图像进行语义分割处理，还可以通过服务器端执行。在第一客户端和/或第二客户端获取各自的图像采集设备采集到的多帧行为图像后，第一客户端和/或第二客户端实时将所获取的多帧行为图像发送至服务器。服务器调用预先部署的语义分割模型。通过语义分割模型对每帧行为图像进行语音分割处理，得到第一目标对象图像和第二目标对象图像，将所得到的第一目标对象图像作为第一显示形象，将所得到的第二目标对象图像作为第二显示形象。服务器可以将第一显示形象和第二显示形象发送至直播间的关联客户端（可以是指已进入直播间的账户对应的客户端），以使关联客户端能够在当前显示的直播场景中渲染第一显示形象和第二显示形象。

一些实施例中，通过在服务器中预先部署语义分割模型，用于对第一客户端和第二客户端获取的多帧行为图像进行语义分割处理，可以减轻终端设备的运行压力，提升终端设备的响应速度。

在一示例性实施例中，在直播场景渲染第一显示形象，包括：对多帧行为图像进行跟踪处理，得到第一目标对象的运动轨迹信息；根据第一目标对象的运动轨迹信息，在直播

场景中渲染第一显示形象的运动轨迹。

在一些实施例中，为了使直播场景中渲染的第一显示形象能够更加接近真实世界的人物行为，一些实施例中，预先在第一客户端部署已训练的目标跟踪算法。通过目标跟踪算法对第一客户端采集多帧行为图像进行跟踪处理，得到第一目标对象的运动轨迹信息。进而根据第一目标对象的运动轨迹信息，在直播场景中渲染第一显示形象的运动轨迹。其中，目标跟踪算法可以采用基于相关滤波器的跟踪算法，例如 KCF Tracker（Kernel Correlation Filter，核相关滤波跟踪算法）、MOSSE Tracker（Minimum Output Sum of Squared Error Tracker，误差最小平方和滤波器跟踪算法）等。

在一些实施例中，在直播场景中渲染第二显示形象，包括：获取第二目标对象的运动轨迹信息，根据第二目标对象的运动轨迹信息，在直播场景中渲染每帧第二显示形象的运动轨迹。

在一些实施例中，第一客户端还可以接收服务器发送的第二目标对象的运动轨迹信息。根据第二目标对象的运动轨迹信息，在当前显示的直播场景中渲染第二显示形象的运动轨迹。第二目标对象的运动轨迹信息可以通过预先配置在第二客户端的目标跟踪算法，对第二目标对象的多帧行为图像进行跟踪处理得到。

进一步地，第一客户端和第二客户端还可以通过服务器将第一显示形象、第一目标对象的运动轨迹信息，以及第二显示形象、第二目标对象的运动轨迹信息发送至直播间的其他关联客户端，以使其他关联客户端在当前显示的直播场景中渲染第一显示形象、第一显示形象的运动轨迹，以及渲染第二显示形象、第二显示形象的运动轨迹。

一些实施例中，通过预先部署目标跟踪算法，通过目标跟踪算法得到真实世界的目标对象的运动轨迹信息，并根据真实世界的目标对象的运动轨迹信息在直播场景中渲染显示形象的运动轨迹，使得直播场景中显示形象可按照真实世界人物的行为进行互动，可以使直播互动方式更加全面，且可以提高直播互动的真实性，有助于提高用户的停留时间。

在一示例性实施例中，对多帧行为图像进行跟踪处理，得到第一目标对象的运动轨迹信息，包括：将多帧行为图像发送至服务器；接收服务器发送的对多帧行为图像进行跟踪处理得到的第一目标对象的运动轨迹信息。

在一些实施例中，对第一客户端和/或第二客户端采集到的多帧行为图像进行跟踪处理，还可以通过服务器端执行。在第一客户端和/或第二客户端各自获取图像采集设备采集到的多帧行为图像后，第一客户端和/或第二客户端实时将所获取的多帧行为图像发送至服务器。服务器调用预先部署的目标跟踪算法。通过目标跟踪算法对多帧行为图像进行跟踪处理，得到第一目标对象和第二目标对象各自对应的运动轨迹信息。服务器可以将第一目标对象和第二目标对象各自对应的运动轨迹信息发送至直播间的关联客户端，以使关联客户端能够在当前显示的直播场景中渲染第一显示形象和第二显示形象各自对应的运动轨迹。

一些实施例中，通过在服务器中预先部署已训练的目标跟踪算法，用于对第一客户端

和第二客户端获取的多帧行为图像进行跟踪处理，可以减轻终端设备的运行压力，提升终端设备的响应速度。

在一示例性实施例中，在对每帧行为图像进行语义分割处理之前，还包括：获取直播场景的场景显示参数以及图像采集设备的设备参数；根据场景显示参数和设备参数，对每

其中，直播场景的场景显示参数不限于包括直播场景的亮度、对比度等信息。直播场景的场景显示参数可以是主播在创建直播间时手动配置的，或者采用预先配置的默认参数。设备参数是指用于采集行为图像的图像采集设备的参数。设备参数不限于包括光照、对比度、摄像头分辨率、镜头畸变等系数。第一客户端和第二客户端各自对应的图像采集设备的设备参数可能不同。

在一些实施例中，第一客户端获取直播场景的场景显示参数。在采集第一目标对象的行为图像时，第一客户端获取图像采集设备的设备参数。第一客户端根据直播场景的场景显示参数，对所获取的每帧行为图像进行调整。示例性地，若已获取的直播场景的场景显示参数和图像采集设备的设备参数都包含亮度，且场景显示参数的亮度小于设备参数中的亮度，则可以根据场景显示参数的亮度，减小第一目标对象的行为图像的亮度。

同样地，对于第二客户端，在采集第二目标对象的行为图像时，第二客户端获取直播场景的场景显示参数以及图像采集设备的设备参数。第二客户端根据直播场景的场景显示参数，对所获取的每帧行为图像进行调整。

在一些实施例中，对每帧行为图像进行语义分割处理，具体包括：对调整后的每帧行为图像进行语义分割处理。在一些实施例中，在对第一目标对象的每帧行为图像进行调整后，第一客户端调用预先部署的语义分割模型对第一目标对象的每帧行为图像进行语义分割处理，得到第一目标对象图像，并将所得到的第一目标对象图像作为第一显示形象。

一些实施例中，通过根据直播场景的场景显示参数以及图像采集设备的设备参数对所获取的行为图像进行调整，可以使不同的客户端采集到的行为图像在直播场景中呈现的效果更加一致。

在一示例性实施例中，在直播场景渲染第一显示形象，包括，对第一显示形象进行行为分析，得到第一显示形象的行为类别，按照与行为类别对应的渲染方式在直播场景中渲染第一显示形象。

其中，行为类别不限于跳舞、对唱、跳跃、击掌、激励等。与行为类别对应的渲染方式可以是指与行为类别对应的相关特效渲染方式，例如，与行为类别为跳舞对应的渲染方式可以是灯光特效，与行为类别为击掌对应的渲染方式可以是将同为击掌的至少一个显示形象靠近，并在击掌部位增加相应的特效。

在一些实施例中，对第一显示形象进行行为分析可以基于深度学习理论执行。示例性地，若第一显示形象是对行为图像进行语义分割处理得到的第一目标对象图像，则可以采用深度学习模型对第一显示形象进行动作识别，得到第一显示形象的行为类别；若第一显

示形象是对语音数据进行语音识别得到的相关文字内容,则可以对相关文字内容进行关键字识别,得到第一显示形象的行为类别。行为类别与渲染方式的对应关系可以预先配置在第一客户端。在第一客户端获取第一显示形象的行为类别后,可以从行为类别与渲染方式的对应关系中查找与行为类别对应的渲染方式,并按照与行为类别对应的渲染方式在直播场景中渲染第一显示形象。

在一些实施例中,在直播场景中渲染第二显示形象,包括:获取第二显示形象的行为类别,按照与第二显示形象的行为类别对应的渲染方式,在直播场景中渲染第二显示形象。

同样地,对于第二客户端,在获取第二显示形象后,第二客户端可以基于深度学习理论对第二显示形象进行行为分析,得到第二显示形象的行为类别。第二客户端可以将第二显示形象的行为类别发送至服务器。通过服务器将第二显示形象的行为类别发送至第一客户端,以使第一客户端能够在直播场景中按照与第二显示形象的行为类别对应的渲染方式,在直播场景中渲染第二显示形象。

一些实施例中,通过对直播场景中的显示形象进行行为分析,得到显示形象的行为类别,并按照与行为类别对应的渲染方式在直播场景中渲染显示形象,进一步丰富了直播互动方式,且可以使直播场景中的显示形象在视觉效果上更加生动形象,有助于增加直播间的观众数量,提高直播间观众的停留时长。

在一示例性实施例中,第一目标对象为主播,第二目标对象为观众;获取第二目标对象的第二显示形象,包括:响应于第二目标对象的互动请求,根据互动请求获取第二目标对象的第二显示形象。

其中,若第一目标对象为主播,则第二目标对象为参与直播互动的互动观众。在一些实施例中,第二目标对象可以通过第二客户端触发互动请求。第二客户端响应于互动请求,采集第二目标对象的行为数据,并根据第二目标对象的行为数据生成第二目标对象对应的第二显示形象。第二客户端可以通过服务器将第二显示形象发送至主播对应的第一客户端,以使第一客户端获取第二显示形象,并在当前显示的直播场景中渲染所获取的第二显示形象。

一些实施例中,通过使观看直播间的观众能够参与直播互动;主播端和观众端同时采集各自用户的行为数据生成显示形象,并同时双向传播显示形象,使主播端和观众端能够在同一虚拟场景中以真实世界行为进行直播互动,从而可以使直播互动方式更加全面。

在一示例性实施例中,第一目标对象为观众,第二目标对象为主播或观众;采集第一目标对象的行为数据,包括:响应于第一目标对象的互动请求,接收互动请求的确认消息,根据互动请求的确认消息采集第一目标对象的行为数据。

其中,若第一目标对象为互动观众,则第二目标对象可以为参与直播互动的其他互动观众或者主播。在一些实施例中,第一目标对象可以通过第一客户端触发互动请求。第一客户端可以通过服务器将互动请求发送至主播对应的第二客户端。主播可以通过第二客户

端触发许可指令。服务器响应于该许可指令，向第一客户端发送互动请求的确认消息，以使第一客户端能够根据互动请求的确认消息开始采集第一目标对象的行为数据。一些实施例中，通过使观众端在接收到主播端的确认消息后，才能够采集观众的行为数据，便于主播对互动观众进行统一管理。

5 在一示例性实施例中，如图3所示，响应于第一目标对象的互动请求，接收互动请求的确认消息，根据互动请求的确认消息采集第一目标对象的行为数据，包括：

在步骤S310中，响应于第一目标对象的互动请求，获取直播场景中的显示形象数量。

在步骤S320中，当显示形象数量未达到数量阈值时，上传互动请求；

10 在步骤S330中，接收互动请求的确认消息，根据确认消息采集第一目标对象的行为数据。

其中，直播场景中的显示形象数量可以是指当前直播场景中互动观众对应的显示形象数量。数量阈值是指允许参与直播互动的最大互动观众数量。数量阈值可以是主播在创建直播间时手动配置的，也可以是预先配置的默认阈值。在一些实施例中，若第一目标对象为观众，则第二目标对象可以其他互动观众或者主播。第一目标对象可以通过第一客户端触发互动请求。第一客户端响应于互动请求，获取当前直播场景中的显示形象数量。将当前直播场景中的显示形象数量与预先获取的数量阈值进行比较。若显示形象数量未达到数量阈值，则通过服务器将第一客户端的互动请求发送至主播的第二客户端。主播可以通过第二客户端触发许可指令。服务器响应于该许可指令，向第一客户端发送互动请求的确认消息，以使第一客户端能够根据互动请求的确认消息采集第一目标对象的行为数据。一些
15 实施例中，通过为直播场景配置相应的数量阈值，控制参与直播互动的观众人数，可以改善直播场景中显示形象的展示效果。

在一示例性实施例中，在步骤S220中，采集第一目标对象的行为数据，根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象，包括：采集第一目标对象的行为数据；当根据第一目标对象的行为数据识别出第一目标对象的全身形象时，根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象。
25

在一些实施例中，若第一目标对象为观众，则第二目标对象可以为其他互动观众或者主播。第一目标对象可以通过第一客户端触发互动请求。第一客户端可以通过服务器将第一客户端的互动请求发送至主播的第二客户端。主播可以通过第二客户端触发许可指令。服务器响应于该许可指令，向第一客户端发送互动请求的确认消息，以使第一客户端能够根据互动请求的确认消息采集第一目标对象的行为数据。第一目标对象的行为数据中包含第一目标对象的行为图像。第一客户端可以对第一目标对象的行为图像进行识别，判断行为图像中是否包含第一目标对象的全身形象。若包含第一目标对象的全身形象，则获取第一显示形象，并将第一显示形象渲染至直播场景中。
30

一些实施例中，通过在判断互动观众的客户端能够采集互动观众的全身形象后，再允许互动客户端继续采集互动观众的行为数据，可以提升直播互动的合规性。
35

图 4 是根据一示例性实施例示出的一种直播互动方法的流程图,如图 4 所示,直播互动方法用于主播端中,包括以下步骤。

在步骤 S401 中,主播端创建直播间,配置直播间的直播场景以及直播场景中的显示形象的数量阈值。

5 在步骤 S402 中,主播端在直播间界面中显示直播场景。

在步骤 S403 中,采集主播的行为数据,主播的行为数据可以通过摄像头采集的连续多帧主播行为图像。

在步骤 S404 中,对每帧主播行为图像进行语义分割处理、跟踪处理以及行为分析,得到主播显示形象、主播的运动轨迹信息以及主播显示形象的行为类别。

10 在一些实施例中,通过语义分割模型对每帧主播行为图像进行语义分割处理,得到每帧主播人像的分割结果,作为每帧主播显示形象。通过动作识别模型对主播显示形象进行识别,得到主播显示形象的行为类别。通过目标跟踪算法对多帧主播行为图像进行跟踪处理,得到的运动轨迹信息。进一步地,主播显示形象的行为类别也可以通过目标跟踪算法对多帧主播行为图像进行行为检测得到,在此不做具体限定。

15 在步骤 S405 中,在主播端的直播场景中渲染主播显示形象以及主播显示形象的运动轨迹,并根据与主播显示形象的行为类别对应的渲染方式,对直播场景中的主播显示形象进行渲染。

在步骤 S406 中,将主播显示形象、主播的运动轨迹信息以及主播显示形象的行为类别发送至服务器,以使服务器将主播显示形象、主播的运动轨迹信息以及主播显示形象的行为类别发送至所有观众端。通过观众端在直播场景中渲染主播显示形象以及主播显示形象的运动轨迹,并根据与主播显示形象的行为类别对应的渲染方式,对直播场景中的主播显示形象进行渲染。

20 在步骤 S407 中,响应于互动观众端的互动请求,获取许可指令以及为互动观众端的观众显示形象分配的初始位置信息。

25 在步骤 S408 中,向互动观众端发送互动请求的确认消息。

在步骤 S409 中,获取互动观众的观众显示形象,根据初始位置信息将观众显示形象渲染至对应的初始位置。其中,互动观众的观众显示形象是互动观众端或者主播端在检测直播场景中的显示形象数量未超过数量阈值、且互动观众端确定摄像头可以采集到观众的全身图像时,根据采集的观众行为图像得到的。

30 在步骤 S410 中,继续获取观众显示形象、互动观众的运动轨迹信息以及观众显示形象的行为类别。其中,观众显示形象、互动观众的运动轨迹信息以及观众显示形象的行为类别可以参照步骤 S404 得到,在此不做具体阐述。

在步骤 S411 中,主播端在直播场景中渲染观众显示形象、观众显示形象的运动轨迹,并按照与观众显示形象的行为类别对应的渲染方式,对直播场景中的观众显示形象进行渲染。图 5 示例性示出了一个实施例中主播端显示的直播场景。其中,直播场景为预先选择

35

的虚拟场景，主播显示形象和观众显示形象为通过语音分割模型得到的真实主播人像和真实观众人像。

图 6 是根据一示例性实施例示出的一种直播互动方法的流程图，如图 5 所示，直播互动方法用于互动观众端中，包括以下步骤。

5 在步骤 S601 中，互动观众端在直播间界面中显示直播场景。

在步骤 S602 中，获取主播显示形象、主播的运动轨迹信息以及主播显示形象的行为类别。

10 在步骤 S603 中，在直播场景中渲染主播显示形象以及主播显示形象的运动轨迹，并根据与主播显示形象的行为类别对应的渲染方式，对直播场景中的主播显示形象进行渲染。

在步骤 S604 中，响应于互动观众触发的互动请求，获取直播场景中的显示形象数量，并在显示形象数量未达到数量阈值时，向主播端发送互动请求。

15 在步骤 S605 中，接收主播端发送的互动请求的确认消息，确认消息携带初始位置信息，根据确认消息采集互动观众的行为数据。互动观众的行为数据可以通过摄像头采集的互动观众的观众行为图像。

在步骤 S606 中，当根据观众行为图像能够识别出互动观众的全身形象时，对互动观众的行为图像进行语义分割处理，得到观众显示形象，并根据初始位置信息将观众显示形象渲染至对应的初始位置。

20 在步骤 S607 中，将观众显示形象发送至服务器，以使服务器将该观众显示形象发送至主播端和其他所有观众端。

在步骤 S608 中，继续采集观众的连续多帧观众行为图像。

在步骤 S609 中，对每帧观众行为图像进行语义分割处理、跟踪处理以及行为分析，得到观众显示形象、互动观众的运动轨迹信息以及观众显示形象的行为类别。

25 在一些实施例中，通过语义分割模型对每帧观众行为图像进行语义分割处理，得到每帧观众人像分割结果，作为每帧观众显示形象。通过动作识别模型对观众显示形象进行识别，得到观众显示形象的行为类别。通过目标跟踪算法对多帧观众行为图像进行跟踪处理，得到观众的运动轨迹信息。进一步地，观众形象的行为类别也可以通过目标跟踪算法对多帧观众行为图像进行行为检测得到，在此不做具体限定。

30 在步骤 S610 中，在直播场景中渲染观众显示形象以及观众显示形象的运动轨迹，并根据与观众显示形象的行为类别对应的渲染方式，对直播场景中的观众显示形象进行渲染。其中，互动观众端与主播端呈现相同的直播场景，具体可以参照图 5 的直播场景示意图。

35 在步骤 S611 中，将观众显示形象、互动观众的运动轨迹信息以及观众显示形象的行为类别发送至服务器，以使服务器将观众显示形象、观众的运动轨迹信息以及观众显示形象的行为类别发送至主播端和其他所有观众端。通过主播端和其他所有观众端在直播场景

中渲染观众显示形象以及观众显示形象的运动轨迹，并根据与观众显示形象的行为类别对应的特效渲染方式，对直播场景中的观众显示形象进行渲染。

应该理解的是，虽然上述流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，上述流程图的至少一部分步骤可以包括多个步骤或者多个阶段，这些步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤中的步骤或者阶段的至少一部分轮流或者交替地执行。

图 7 是根据一示例性实施例示出的一种直播互动装置 700 框图。参照图 7，该装置 700 包括显示模块 701、采集模块 702、显示形象生成模块 703、第一渲染模块 704、获取模块 705 和第二渲染模块 706。

显示模块 701，被配置为在直播间界面中显示直播场景；采集模块 702，被配置为采集第一目标对象的行为数据；显示形象生成模块 703，被配置为根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象；第一渲染模块 704，被配置为在直播场景渲染第一显示形象；获取模块 705，被配置为获取第二目标对象的第二显示形象，第二显示形象是根据第二目标对象的行为数据生成的；第二渲染模块 706，还被配置为在直播场景中渲染第二显示形象。

在一示例性实施例中，采集模块 702，被配置为采集第一目标对象的多帧行为图像；所述装置 700 还包括：图像分割模块，被配置为对每帧行为图像进行语义分割处理，得到每帧第一显示形象；第一渲染模块 704，还被配置为在直播场景中渲染每帧第一显示形象。

在一示例性实施例中，图像分割模块，包括：发送单元，被配置为将多帧行为图像发送至服务器；接收单元，被配置为接收服务器发送的对每帧行为图像进行语音分割处理得到的每帧第一显示形象。

在一示例性实施例中，第一渲染模块 704，包括：跟踪单元，被配置为对多帧行为图像进行跟踪处理，得到第一目标对象的运动轨迹信息；第一渲染单元，被配置为根据第一目标对象的运动轨迹信息，在直播场景中渲染第一显示形象的运动轨迹；第二渲染模块 706，包括：轨迹信息获取单元，被配置为获取第二目标对象的运动轨迹信息；第二渲染单元，被配置为根据第二目标对象的运动轨迹信息，在直播场景中渲染每帧第二显示形象的运动轨迹。

在一示例性实施例中，跟踪单元，被配置为将多帧行为图像发送至服务器；接收服务器发送的对多帧行为图像进行跟踪处理得到的第一目标对象的运动轨迹信息。

在一示例性实施例中，获取模块 705，还被配置为获取直播场景的场景显示参数以及图像采集设备的设备参数；所述装置 700 还包括：图像调整模块，被配置为根据场景显示参数和设备参数，对每帧行为图像进行调整；图像分割模块，被配置为对调整后的每帧行为图像进行语义分割处理。

在一示例性实施例中，第一渲染模块 704，包括：行为分析单元，被配置为对第一显示形象进行行为分析，得到第一显示形象的行为类别；第三渲染单元，被配置为按照与行为类别对应的渲染方式在直播场景中渲染第一显示形象；第二渲染模块 706，包括：行为类别获取单元，被配置为获取第二显示形象的行为类别；第四渲染单元，被配置为按照与第二显示形象的行为类别对应的渲染方式，在直播场景中渲染第二显示形象。

在一示例性实施例中，第一目标对象为主播，第二目标对象为观众；获取模块 705，被配置为响应于第二目标对象的互动请求，根据互动请求获取第二目标对象的第二显示形象。

在一示例性实施例中，第一目标对象为观众，第二目标对象为主播或观众；采集模块 702，被配置为响应于第一目标对象的互动请求，接收互动请求的确认消息，根据互动请求的确认消息采集第一目标对象的行为数据。

在一示例性实施例中，采集模块 702，包括：数量获取单元，被配置为响应于第一目标对象的互动请求，获取直播场景中的显示形象数量；上传单元，被配置为当显示形象数量未达到数量阈值时，上传互动请求；采集单元，被配置为接收互动请求的确认消息，根据确认消息采集第一目标对象的行为数据。

在一示例性实施例中，采集模块 702，被配置为采集第一目标对象的行为数据；当根据第一目标对象的行为数据识别出第一目标对象的全身形象时，根据第一目标对象的行为数据生成第一目标对象对应的第一显示形象。

关于上述实施例中的装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

图 8 是根据一示例性实施例示出的一种用于直播互动的设备 800 的框图。例如，设备 800 可以是移动电话、计算机、数字广播终端、消息收发设备、游戏控制台、平板设备、医疗设备、健身设备、个人数字助理等。

参照图 8，设备 800 可以包括以下一个或多个组件：处理组件 802、存储器 804、电源组件 806、多媒体组件 808、音频组件 810、输入/输出（I/O）的接口 812、传感器组件 814 以及通信组件 816。

处理组件 802 通常控制设备 800 的整体操作，诸如与显示、电话呼叫、数据通信、相机操作和记录操作相关联的操作。处理组件 802 可以包括一个或多个处理器 820 来执行指令，以完成上述的方法的全部或部分步骤。此外，处理组件 802 可以包括一个或多个模块，便于处理组件 802 和其他组件之间的交互。例如，处理组件 802 可以包括多媒体模块，以方便多媒体组件 808 和处理组件 802 之间的交互。

存储器 804 被配置为存储各种类型的数据以支持在设备 800 的操作。这些数据的示例包括用于在设备 800 上操作的任何应用程序或方法的指令、联系人数据、电话簿数据、消息、图片、视频等。存储器 804 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM）、电可擦除可编程只读存储器（EEPROM）、

可擦除可编程只读存储器 (EPROM)、可编程只读存储器 (PROM)、只读存储器 (ROM)、磁存储器、快闪存储器、磁盘或光盘。

电源组件 806 为设备 800 的各种组件提供电力。电源组件 806 可以包括电源管理系统，一个或多个电源，及其他与为设备 800 生成、管理和分配电力相关联的组件。

5 多媒体组件 808 包括在所述设备 800 和用户之间的提供一个输出接口的屏幕。在一些实施例中，屏幕可以包括液晶显示器 (LCD) 和触摸面板 (TP)。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中，
10 多媒体组件 808 包括一个前置摄像头和/或后置摄像头。当设备 800 处于操作模式，如拍摄模式或视频模式时，前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

音频组件 810 被配置为输出和/或输入音频信号。例如，音频组件 810 包括一个麦克风 (MIC)，当设备 800 处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风
15 被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 804 或经由通信组件 816 发送。在一些实施例中，音频组件 810 还包括一个扬声器，用于输出音频信号。

I/O 接口 812 为处理组件 802 和外围接口模块之间提供接口，上述外围接口模块可以是键盘，点击轮，按钮等。这些按钮可包括但不限于：主页按钮、音量按钮、启动按钮和
20 锁定按钮。

传感器组件 814 包括一个或多个传感器，用于为设备 800 提供各个方面的状态评估。例如，传感器组件 814 可以检测到设备 800 的打开/关闭状态，组件的相对定位，例如所述组件为设备 800 的显示器和小键盘，传感器组件 814 还可以检测设备 800 或设备 800 一个组件的位置改变，用户与设备 800 接触的存在或不存在，设备 800 方位或加速/减速和
25 设备 800 的温度变化。传感器组件 814 可以包括接近传感器，被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件 814 还可以包括光传感器，如 CMOS 或 CCD 图像传感器，用于在成像应用中使用。在一些实施例中，该传感器组件 814 还可以包括加速度传感器、陀螺仪传感器、磁传感器、压力传感器或温度传感器。

通信组件 816 被配置为便于设备 800 和其他设备之间有线或无线方式的通信。设备 800
30 可以接入基于通信标准的无线网络，如 WiFi，运营商网络（如 2G、3G、4G 或 5G），或它们的组合。在一个示例性实施例中，通信组件 816 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，所述通信组件 816 还包括近场通信 (NFC) 模块，以促进短程通信。例如，在 NFC 模块可基于射频识别 (RFID) 技术，红外数据协会 (IrDA) 技术，超宽带 (UWB) 技术，蓝牙 (BT) 技术和其他技术来实现。

35 在示例性实施例中，设备 800 可以被一个或多个应用专用集成电路 (ASIC)、数字

信号处理器（DSP）、数字信号处理设备（DSPD）、可编程逻辑器件（PLD）、现场可编程门阵列（FPGA）、控制器、微控制器、微处理器或其他电子元件实现，用于执行上述方法。

5 在示例性实施例中，还提供了一种包括指令的非临时性计算机可读存储介质，例如包括指令的存储器 804，上述指令可由设备 800 的处理器 820 执行以完成上述方法。例如，所述非临时性计算机可读存储介质可以是 ROM、随机存取存储器（RAM）、CD-ROM、磁带、软盘和光数据存储设备等。

本公开所有实施例均可以单独被执行，也可以与其他实施例相结合被执行，均视为本公开要求的保护范围。

权利要求

1.一种直播互动方法，其中，所述方法应用于直播端或者观众端，包括：

在直播间界面中显示直播场景；

采集第一目标对象的行为数据，根据所述第一目标对象的行为数据生成所述第一目标

5 对象对应的第一显示形象，在所述直播场景渲染所述第一显示形象；

获取第二目标对象的第二显示形象，所述第二显示形象是根据所述第二目标对象的行为数据生成的；

在所述直播场景中渲染所述第二显示形象。

2.根据权利要求1所述的直播互动方法，其中，所述采集第一目标对象的行为数据，
10 根据所述第一目标对象的行为数据生成所述第一目标对象对应的第一显示形象，在所述直播场景渲染所述第一显示形象，包括：

采集所述第一目标对象的多帧行为图像，对每帧行为图像进行语义分割处理，得到所述每帧第一显示形象，在所述直播场景中渲染所述每帧第一显示形象。

3.根据权利要求2所述的直播互动方法，其中，所述对每帧行为图像进行语义分割处
15 理，包括：

将所述多帧行为图像发送至服务器；

接收所述服务器发送的对所述每帧行为图像进行语音分割处理得到的所述每帧第一显示形象。

4.根据权利要求2所述的直播互动方法，其中，所述在所述直播场景渲染所述第一显示
20 形象，包括：对所述多帧行为图像进行跟踪处理，得到所述第一目标对象的运动轨迹信息；根据所述第一目标对象的运动轨迹信息，在所述直播场景中渲染所述第一显示形象的运动轨迹；

所述在所述直播场景中渲染所述第二显示形象，包括：获取所述第二目标对象的运动
25 轨迹信息，根据所述第二目标对象的运动轨迹信息，在所述直播场景中渲染每帧第二显示形象的运动轨迹。

5.根据权利要求4所述的直播互动方法，其中，所述对所述多帧行为图像进行跟踪处理，得到所述第一目标对象的运动轨迹信息，包括：

将所述多帧行为图像发送至服务器；

30 接收所述服务器发送的对所述多帧行为图像进行跟踪处理得到的所述第一目标对象的运动轨迹信息。

6.根据权利要求2所述的直播互动方法，其中，所述方法还包括：

获取所述直播场景的场景显示参数以及所述直播端或观众端的图像采集设备的设备
参数；

根据所述场景显示参数和所述设备参数，对所述每帧行为图像进行调整；

35 所述对每帧行为图像进行语义分割处理，包括：对调整后的所述每帧行为图像进行语

义分割处理。

7.根据权利要求1所述的直播互动方法,其中,所述在所述直播场景渲染所述第一显示形象,包括,对所述第一显示形象进行行为分析,得到所述第一显示形象的行为类别,按照与所述行为类别对应的渲染方式在所述直播场景中渲染所述第一显示形象;

5 所述在所述直播场景中渲染所述第二显示形象,包括:

获取所述第二显示形象的行为类别,按照与所述第二显示形象的行为类别对应的渲染方式,在所述直播场景中渲染所述第二显示形象。

8.根据权利要求1所述的直播互动方法,其中,所述第一目标对象为主播,所述第二目标对象为观众;所述获取第二目标对象的第二显示形象,包括:

10 响应于第二目标对象的互动请求,根据所述互动请求获取第二目标对象的第二显示形象。

9.根据权利要求1所述的直播互动方法,其中,所述第一目标对象为观众,所述第二目标对象为主播或观众;所述采集第一目标对象的行为数据,包括:

15 响应于第一目标对象的互动请求,接收所述互动请求的确认消息,根据所述互动请求的确认消息采集所述第一目标对象的行为数据。

10.根据权利要求9所述的直播互动方法,其中,所述响应于第一目标对象的互动请求,接收所述互动请求的确认消息,根据所述互动请求的确认消息采集所述第一目标对象的行为数据,包括:

响应于所述第一目标对象的互动请求,获取所述直播场景中的显示形象数量;

20 响应于所述显示形象数量未达到数量阈值,上传所述互动请求;

接收所述互动请求的确认消息,根据所述确认消息采集所述第一目标对象的行为数据。

11.根据权利要求9所述的直播互动方法,其中,所述采集第一目标对象的行为数据,根据所述第一目标对象的行为数据生成所述第一目标对象对应的第一显示形象,包括:

25 采集所述第一目标对象的行为数据;

响应于根据所述第一目标对象的行为数据识别出所述第一目标对象的全身形象,根据所述第一目标对象的行为数据生成所述第一目标对象对应的第一显示形象。

12.一种直播互动装置,其中,所述装置包括:

显示模块,被配置为在直播间界面中显示直播场景;

30 采集模块,被配置为采集第一目标对象的行为数据;

显示形象生成模块,被配置为根据所述第一目标对象的行为数据生成所述第一目标对象对应的第一显示形象;

第一渲染模块,被配置为在所述直播场景渲染所述第一显示形象;

35 获取模块,被配置为获取第二目标对象的第二显示形象,所述第二显示形象是根据所述第二目标对象的行为数据生成的;

第二渲染模块，还被配置为在所述直播场景中渲染所述第二显示形象。

13.根据权利要求 12 所述的直播互动装置，其中，所述采集模块，被配置为采集所述第一目标对象的多帧行为图像；

所述装置还包括：图像分割模块，被配置为对每帧行为图像进行语义分割处理，得到

5 所述每帧第一显示形象；

所述第一渲染模块，还被配置为在所述直播场景中渲染所述每帧第一显示形象。

14.根据权利要求 13 所述的直播互动装置，其中，所述图像分割模块，包括：

发送单元，被配置为将所述多帧行为图像发送至服务器；

10 接收单元，被配置为接收所述服务器发送的对所述每帧行为图像进行语音分割处理得到的所述每帧第一显示形象。

15.根据权利要求 13 所述的直播互动装置，其中，所述第一渲染模块，包括：

跟踪单元，被配置为对所述多帧行为图像进行跟踪处理，得到所述第一目标对象的运动轨迹信息；

15 第一渲染单元，被配置为根据所述第一目标对象的运动轨迹信息，在所述直播场景中渲染所述第一显示形象的运动轨迹；

所述第二渲染模块，包括：

轨迹信息获取单元，被配置为获取所述第二目标对象的运动轨迹信息；

第二渲染单元，被配置为根据所述第二目标对象的运动轨迹信息，在所述直播场景中渲染每帧第二显示形象的运动轨迹。

20 16.根据权利要求 15 所述的直播互动装置，其中，所述跟踪单元，被配置为将所述多帧行为图像发送至服务器；接收所述服务器发送的对所述多帧行为图像进行跟踪处理得到的所述第一目标对象的运动轨迹信息。

17.根据权利要求 13 所述的直播互动装置，其中，所述获取模块，还被配置为获取所述直播场景的场景显示参数以及图像采集设备的设备参数；

25 所述装置还包括：图像调整模块，被配置为根据所述场景显示参数和所述设备参数，对所述每帧行为图像进行调整；

所述图像分割模块，被配置为对调整后的所述每帧行为图像进行语义分割处理。

18.根据权利要求 12 所述的直播互动装置，其中，所述第一渲染模块，包括：行为分析单元，被配置为对所述第一显示形象进行行为分析，得到所述第一显示形象的行为类别；

30 第三渲染单元，被配置为按照与所述行为类别对应的渲染方式在所述直播场景中渲染所述第一显示形象；

所述第二渲染模块，包括：

行为类别获取单元，被配置为获取所述第二显示形象的行为类别；

35 第四渲染单元，被配置为按照与所述第二显示形象的行为类别对应的渲染方式，在所述直播场景中渲染所述第二显示形象。

19.根据权利要求 12 所述的直播互动装置,其中,所述第一目标对象为主播,所述第二目标对象为观众;所述获取模块,被配置为响应于第二目标对象的互动请求,根据所述互动请求获取第二目标对象的第二显示形象。

20.根据权利要求 12 所述的直播互动装置,其中,所述第一目标对象为观众,所述第二目标对象为主播或观众;所述采集模块,被配置为响应于第一目标对象的互动请求,接收所述互动请求的确认消息,根据所述互动请求的确认消息采集所述第一目标对象的行为数据。

21.根据权利要求 209 所述的直播互动装置,其中,所述采集模块,包括:

数量获取单元,被配置为响应于所述第一目标对象的互动请求,获取所述直播场景中的显示形象数量;

上传单元,被配置为响应于所述显示形象数量未达到数量阈值,上传所述互动请求;

采集单元,被配置为接收所述互动请求的确认消息,根据所述确认消息采集所述第一目标对象的行为数据。

22.根据权利要求 20 所述的直播互动装置,其中,所述采集模块,被配置为采集所述第一目标对象的行为数据;响应于根据所述第一目标对象的行为数据识别出所述第一目标对象的全身形象,根据所述第一目标对象的行为数据生成所述第一目标对象对应的第一显示形象。

23.一种电子设备,其中,所述电子设备包括:

处理器;

用于存储所述处理器可执行指令的存储器;

其中,所述处理器被配置为执行所述指令,以实现一种直播互动方法,所述方法包括:在直播间界面中显示直播场景;

采集第一目标对象的行为数据,根据所述第一目标对象的行为数据生成所述第一目标对象对应的第一显示形象,在所述直播场景渲染所述第一显示形象;

获取第二目标对象的第二显示形象,所述第二显示形象是根据所述第二目标对象的行为数据生成的;

在所述直播场景中渲染所述第二显示形象。

24.一种非易失性存储介质,当所述存储介质中的指令由电子设备的处理器执行时,使得所述电子设备能够执行一种直播互动方法,所述方法包括:

在直播间界面中显示直播场景;

采集第一目标对象的行为数据,根据所述第一目标对象的行为数据生成所述第一目标对象对应的第一显示形象,在所述直播场景渲染所述第一显示形象;

获取第二目标对象的第二显示形象,所述第二显示形象是根据所述第二目标对象的行为数据生成的;

在所述直播场景中渲染所述第二显示形象。

25.一种计算机程序产品，包括计算机程序，其中，所述计算机程序被处理器执行时实现一种直播互动方法，所述方法包括：

在直播间界面中显示直播场景；

5 采集第一目标对象的行为数据，根据所述第一目标对象的行为数据生成所述第一目标对象对应的第一显示形象，在所述直播场景渲染所述第一显示形象；

获取第二目标对象的第二显示形象，所述第二显示形象是根据所述第二目标对象的行为数据生成的；

在所述直播场景中渲染所述第二显示形象。

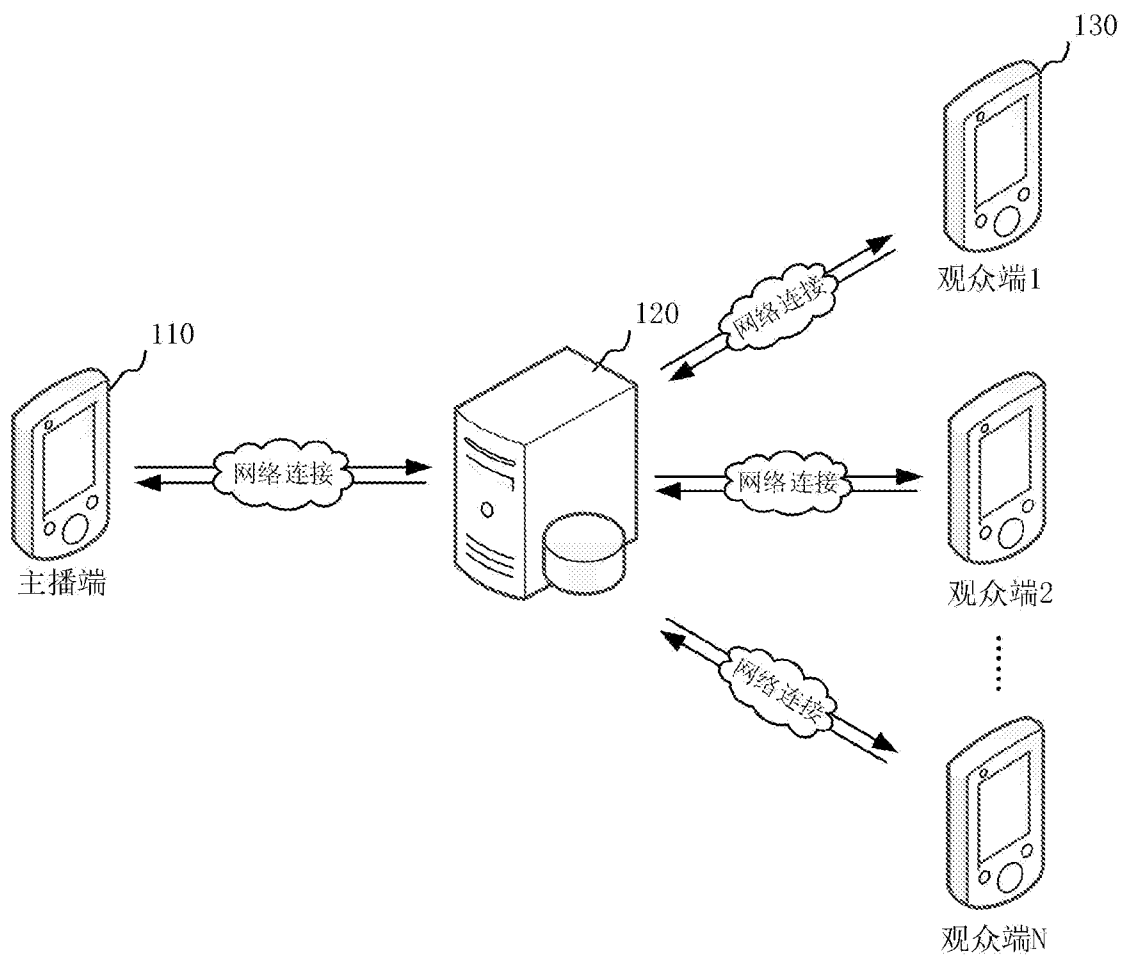


图1

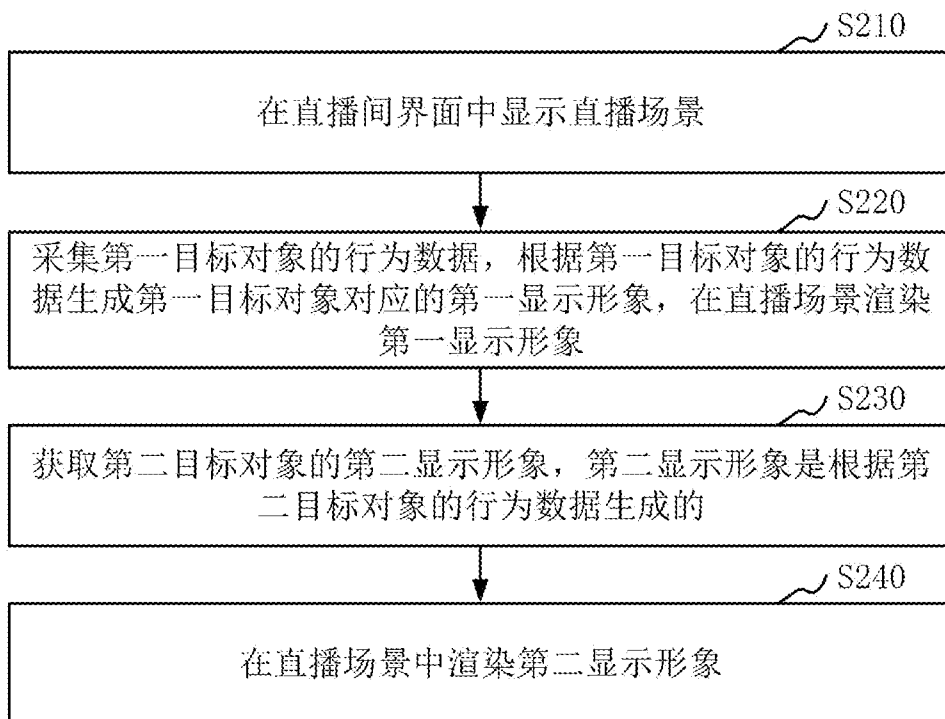


图2

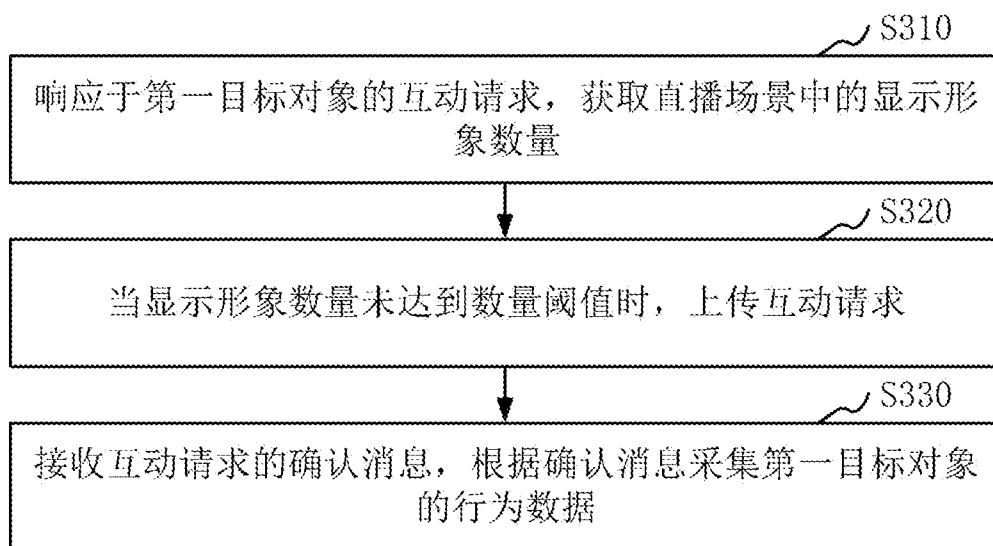


图3

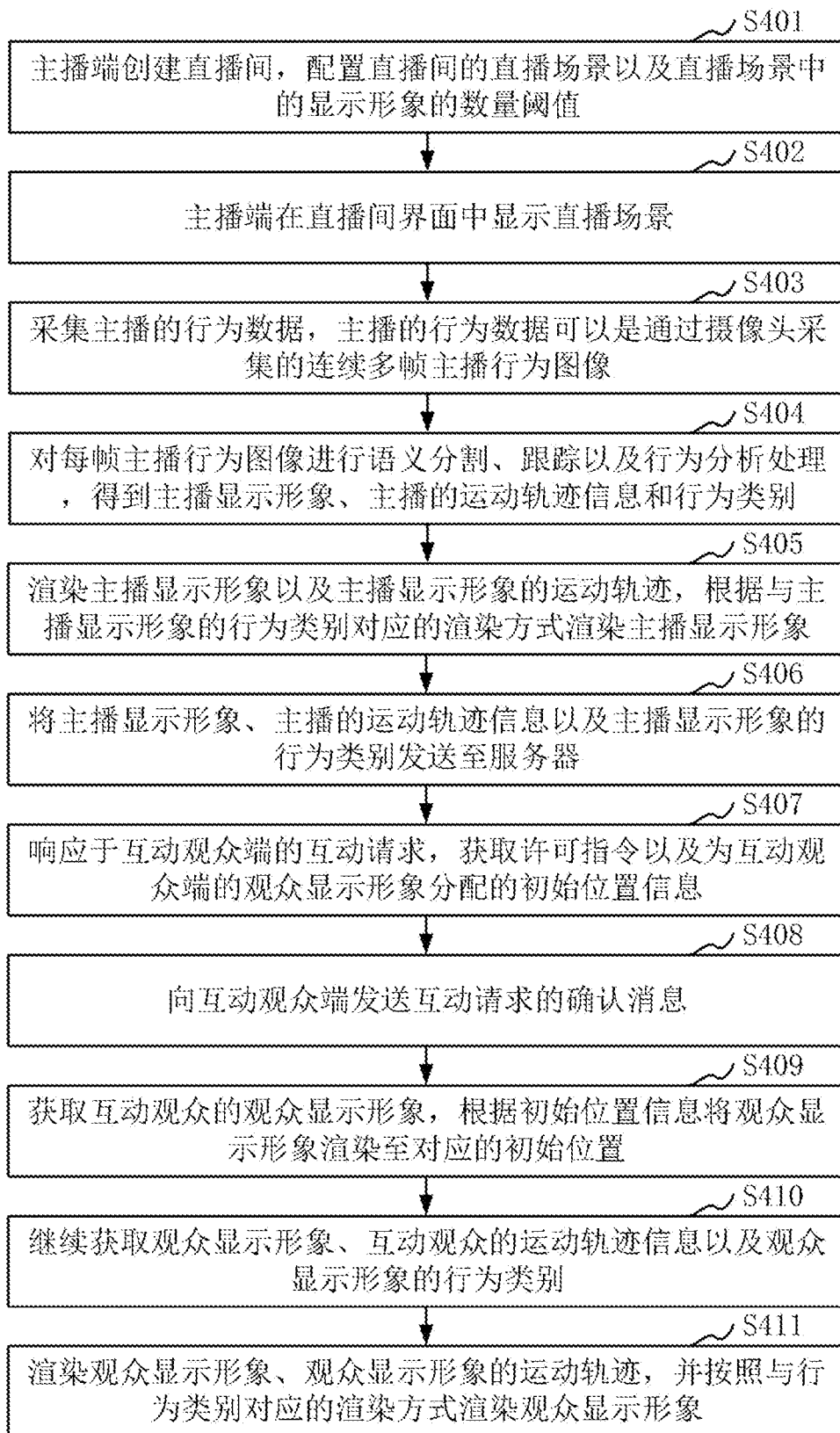


图4



图5

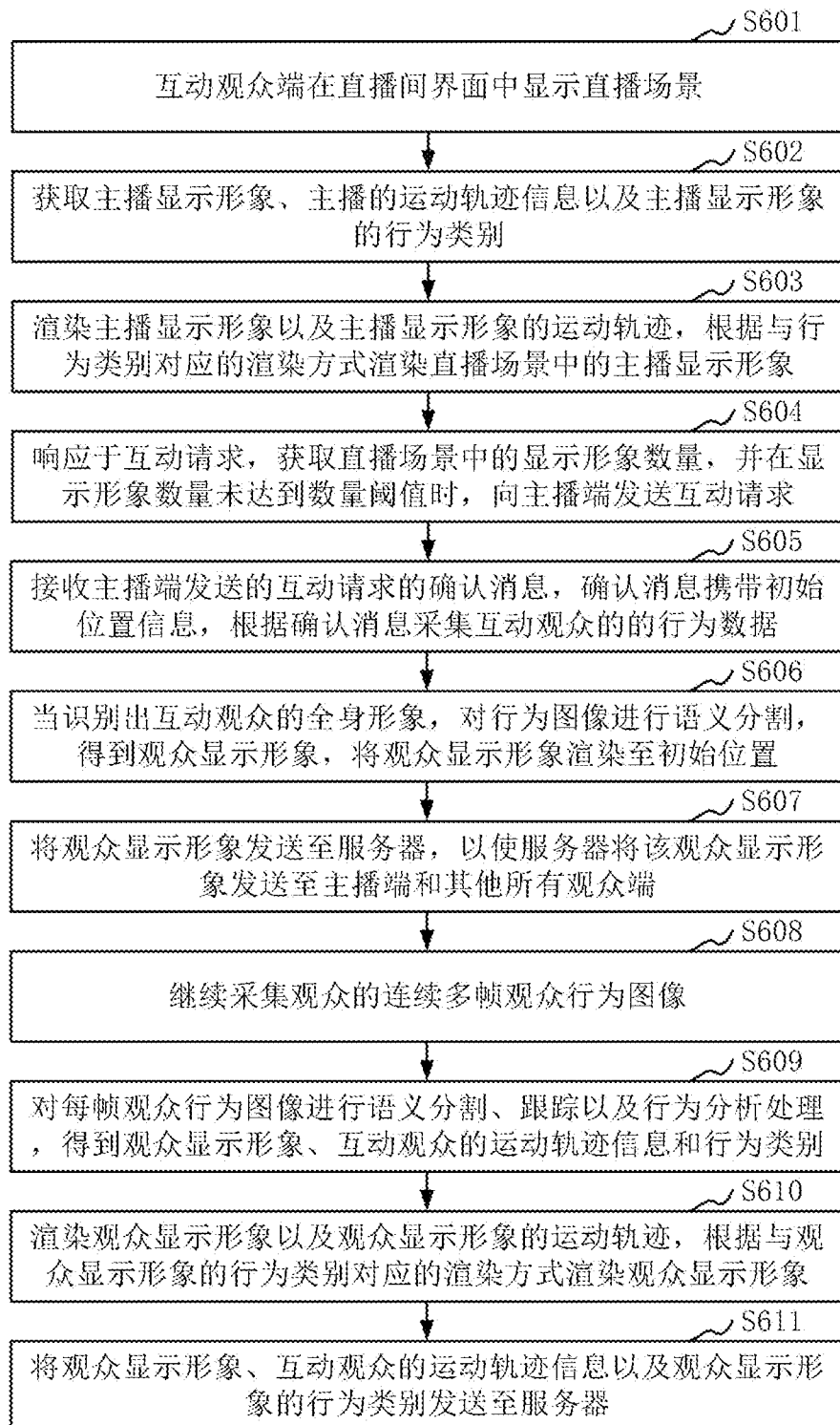


图6

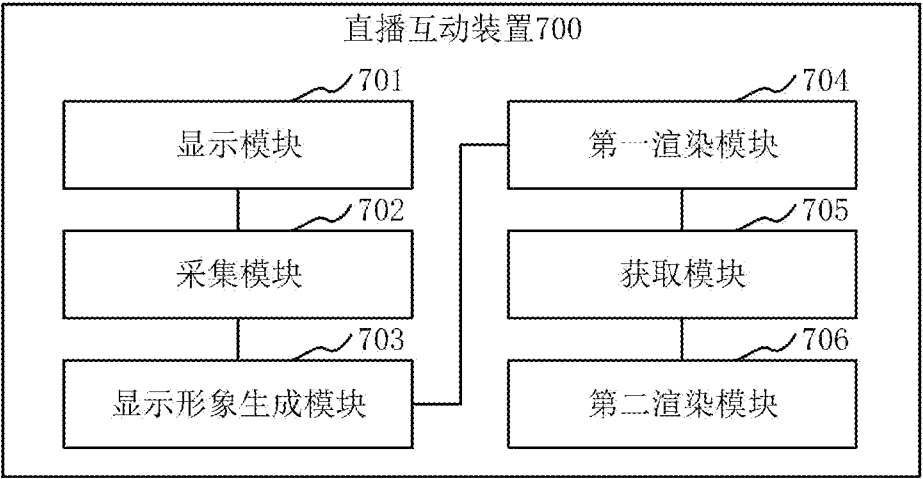


图7

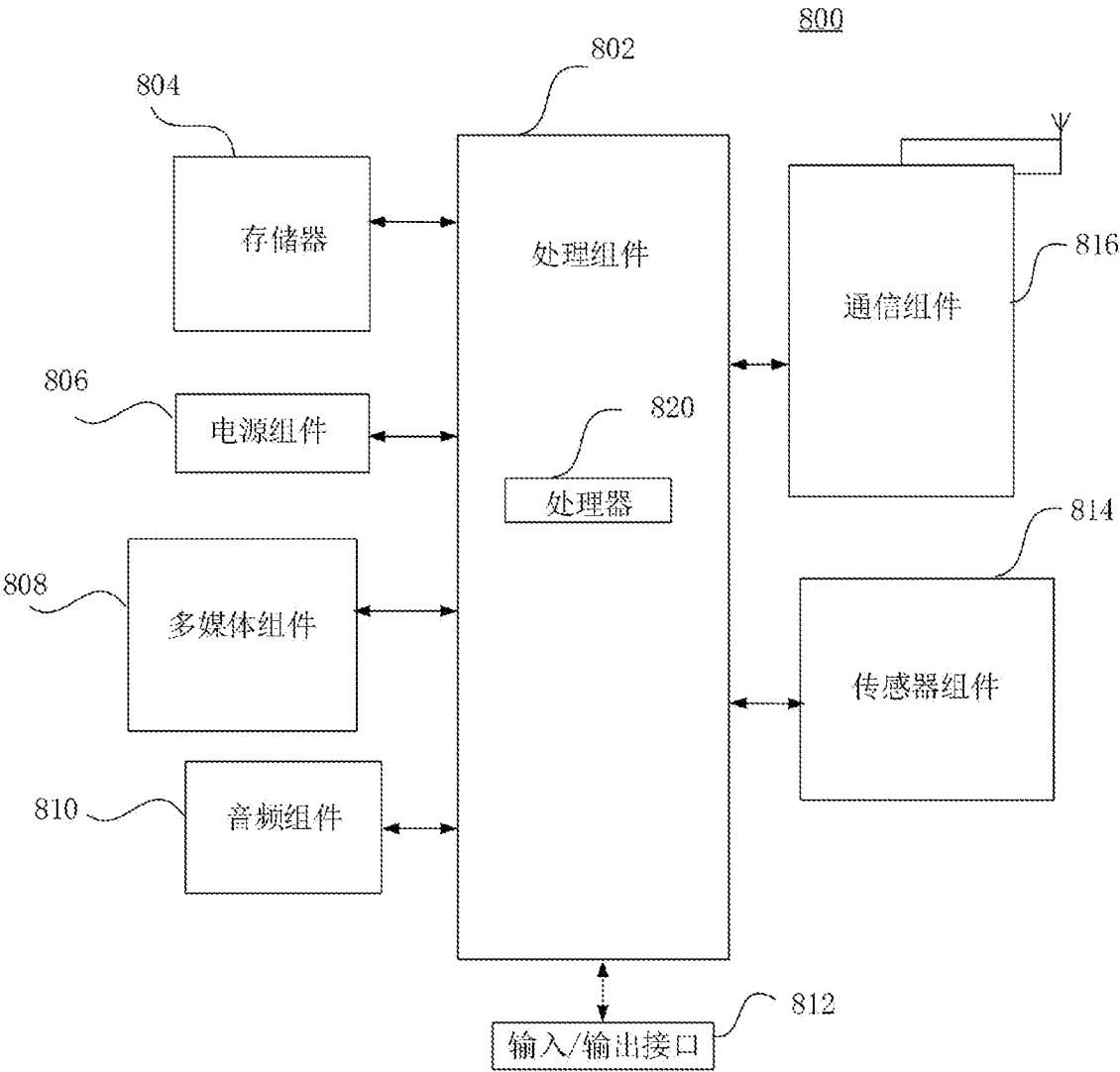


图8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/117040

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/2187(2011.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L; H04N; G06T; G06K; A63F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNKI, CNPAT: live, interact+, scene, behavior+, target, object, render+, semantic, segmentat+, track+, motion, trajectory, 直播, 互动, 行为, 运动, 轨迹, 语义, 分割, 渲染, 形象, 对象, 目标, 跟踪, 追踪

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 106789991 A (FUJIAN STAR-NET E-VIDEO INFORMATION SYSTEM CO., LTD. et al.) 31 May 2017 (2017-05-31) description paragraphs 0002-0004, 0050-0106, figures 1-4	1-25
Y	CN 107613310 A (GUANGZHOU HUADUO NETWORK TECHNOLOGY CO., LTD.) 19 January 2018 (2018-01-19) description, paragraphs 0073-0129, and figures 1-10	1-25
PX	CN 112153400 A (BEIJING DAJIA INTERCONNECTION INFORMATION TECHNOLOGY CO., LTD.) 29 December 2020 (2020-12-29) claims 1-10, description paragraphs 0084-0190	1-25
A	CN 110519611 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 29 November 2019 (2019-11-29) entire document	1-25
A	US 2020099960 A1 (GUANGZHOU HUYA INFORMATION TECHNOLOGY CO., LTD.) 26 March 2020 (2020-03-26) entire document	1-25



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

17 November 2021

Date of mailing of the international search report

01 December 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/117040

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106789991	A	31 May 2017	None			
CN	107613310	A	19 January 2018	None			
CN	112153400	A	29 December 2020	None			
CN	110519611	A	29 November 2019	None			
US	2020099960	A1	26 March 2020	CN	111405299	A	10 July 2020
				WO	2018113405	A1	28 June 2018

国际检索报告

国际申请号

PCT/CN2021/117040

A. 主题的分类

H04N 21/2187(2011.01) i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

H04L; H04N; G06T; G06K; A63F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

WPI, EPDOC, CNKI, CNPAT:live, interact+, scene, behavior+, target, object, render+, semantic, segmentat+, track+, motion, trajectory, 直播, 互动, 行为, 运动, 轨迹, 语义, 分割, 渲染, 形象, 对象, 目标, 跟踪, 追踪

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 106789991 A (福建星网视易信息系统有限公司 等) 2017年 5月 31日 (2017 - 05 - 31) 说明书第0002-0004、0050-0106段, 附图1-4	1-25
Y	CN 107613310 A (广州华多网络科技有限公司) 2018年 1月 19日 (2018 - 01 - 19) 说明书第0073-0129段, 附图1-10	1-25
PX	CN 112153400 A (北京达佳互联信息技术有限公司) 2020年 12月 29日 (2020 - 12 - 29) 权利要求1-10, 说明书第0084-0190段	1-25
A	CN 110519611 A (腾讯科技深圳有限公司) 2019年 11月 29日 (2019 - 11 - 29) 全文	1-25
A	US 2020099960 A1 (GUANGZHOU HUYA INFORMATION TECHNOLOGY CO., LTD.) 2020年 3月 26日 (2020 - 03 - 26) 全文	1-25

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2021年 11月 17日

国际检索报告邮寄日期

2021年 12月 1日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国 北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

郎亦虹

电话号码 53961711

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/117040

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	106789991	A	2017年 5月 31日	无	
CN	107613310	A	2018年 1月 19日	无	
CN	112153400	A	2020年 12月 29日	无	
CN	110519611	A	2019年 11月 29日	无	
US	2020099960	A1	2020年 3月 26日	CN 111405299 A	2020年 7月 10日
				WO 2018113405 A1	2018年 6月 28日