

(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



(10) 国际公布号

WO 2021/121374 A1

(43) 国际公布日
2021年6月24日 (24.06.2021)

(51) 国际专利分类号:

H04N 5/232 (2006.01)

(21) 国际申请号: PCT/CN2020/137550

(22) 国际申请日: 2020年12月18日 (18.12.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:

201911315344.0 2019年12月19日 (19.12.2019) CN
202010753515.4 2020年7月30日 (30.07.2020) CN

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 陈艳花 (CHEN, Yanhua); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 刘宏马 (LIU, Hongma); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 张超 (ZHANG, Chao);

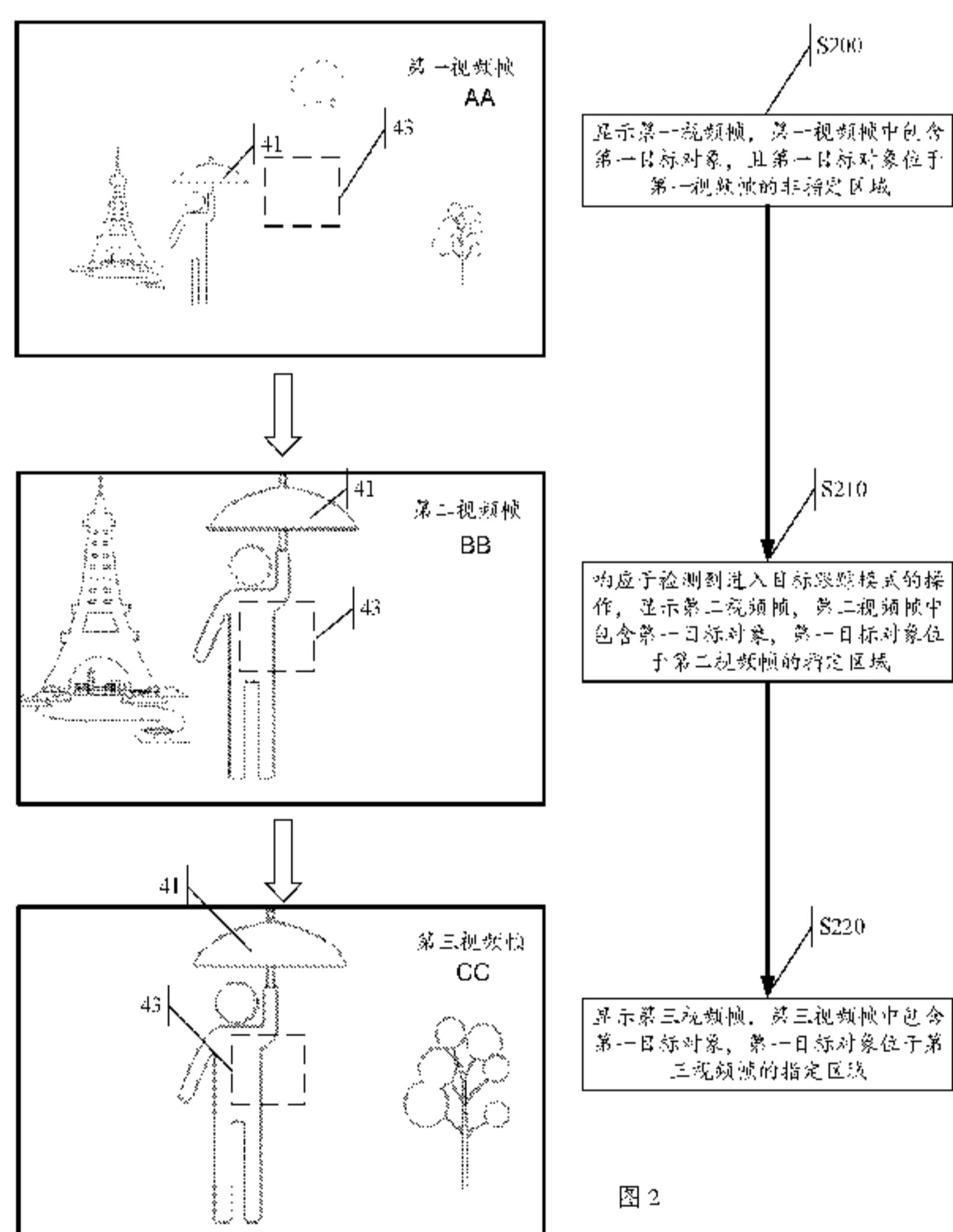
中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 张雅琪 (ZHANG, Yaqi); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 李宏禹 (LI, Hongyu); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(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,

(54) Title: VIDEO PROCESSING METHOD AND ELECTRONIC DEVICE

(54) 发明名称: 一种视频处理方法及电子设备



S200 Display a first video frame, the first video frame containing a first target object, and the first video frame being located in a non-designated region of the first video frame
S210 Display a second video frame in response to detection of an operation to enter a target tracking mode, the second video frame containing the first target object, the first target object being located in a designated region of the second video frame
S220 Display a third video frame, the third video frame containing the first target object, and the first target object being located in a designated region of the third video frame
AA First video frame
BB Second video frame
CC Third video frame

(57) Abstract: The present invention relates to the field of image processing, and disclosed are a video processing method, an electronic device, a computer readable storage medium, a computer program product, and a chip, so as to solve the technical problem in the prior art that a rotating camera needs to be provided for user tracking in a video call. The method comprises: obtaining a first video frame, and determining that the first video frame contains at least one target object; determining a cropping box with the at least one target object as a center; and obtaining content in the cropping box and using the content as a second video for display. The method can be used in an artificial intelligence device, and the method is related to deep learning and other technologies.

(57) 摘要: 本发明涉及图像处理领域, 公开了一种视频处理方法、电子设备、计算机可读存储介质、计算机程序产品及芯片, 以解决现有技术中视频通话对用户跟踪需要设置转动的摄像头的技术问题。该方法包括: 获取第一视频帧, 确定所述第一视频帧中包含有至少一个目标对象; 以所述至少一个目标对象为中心, 确定裁剪框; 获取所述裁剪框中的内容, 作为第二视频帧进行显示。该方法可用于人工智能设备, 该方法和深度学习等技术相关。

WO 2021/121374 A1

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, 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))。

一种视频处理方法及电子设备

本申请要求在 2019 年 12 月 19 日提交中国国家知识产权局、申请号为 201911315344.0 的中国专利申请的优先权，发明名称为“一种视频处理方法及电子设备”的中国专利申请的优先权，在 2020 年 7 月 30 日提交中国国家知识产权局、申请号为 202010753515.4 的中国专利申请的优先权，发明名称为“一种视频处理方法及电子设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及计算机视觉领域，尤其涉及一种电子设备的视频处理方法及电子设备。

背景技术

随着电子设备的普及，尤其是智能移动电子设备如手机的普及，手机上集成了越来越多的应用，涉及到人们日常生活中的方方面面。通常电子设备都包含摄像头，能够通过该摄像头进行图像采集、视频拍摄。

现有技术中，在进行视频采集时，如果需要对目标物体进行跟拍，则需要摄像头具备云台，通过云台旋转摄像头的拍摄角度，从而实现跟拍，该方案存在着成本较高的技术问题。

发明内容

本申请提供一种视频采集方法及电子设备，以解决现有技术中通过电子设备实现跟拍时成本较高的技术问题。

本申请实施例提供一种视频处理方法，其应用于电子设备，包括：获取第一视频帧，确定所述第一视频帧中包含有至少一个目标对象；以该确定的目标对象为中心，确定裁剪框；将该裁剪框中的内容作为第二视频帧。该电子设备可以显示该第二视频帧而不显示第一视频帧，或者显示第一视频帧后，响应于用户对预设功能的启动，显示第二视频帧。这样无需硬件改进，通过采用本视频处理方法，即可实现对目标对象的聚焦，在目标对象移动的场景下，能呈现自动跟踪目标对象，在视频帧中该目标对象始终在视觉中心位置，因此能低成本的实现自动追踪功能，降低对硬件的要求，因此降低功能的实现难度。

本申请实施例还提供一种电子设备，该电子设备执行本发明各方法实施例中的方法。具体的，该电子设备包括一个或多个处理器；一个或多个存储器；多个应用程序；以及一个或多个计算机程序，其中该一个或多个计算机程序被存储在该一个或多个存储器中，计算机程序包括指令，当指令被一个或多个处理器执行时，使得电子设备执行方法实施例中的方法，例如：获取第一视频帧，确定该第一视频帧中包含有至少一个目标对象；以该至少一个目标对象为中心，确定裁剪框；获取该裁剪框中的内容，作为第二视频帧指示进行显示。本发明提供的电子设备能提供视频的自动目标跟踪功能，在视频通话，视频拍摄，自播等场景下，

达到景随人动，移步换景的效果；其实现既不需要配置云台之类的硬件设施，也不需要摄影师手动操作，也无需安装专门图像处理应用，即能实现图像的自动处理，在不影响视频流畅度的前提下，实现目标对象的智能识别和图像的自动处理，提高了视频通信的画面质量，增加了人机交互的智能性。

附图说明

- 图 1A 为本发明实施例的电子设备的结构图；
图 1B 为本发明实施例的软件框架图；
图 2 为本发明实施例第一方面的视频采集方法的流程图；
图 3 为本发明实施例的系统通话应用的界面示意图；
图 4 为本发明实施例的视频控制方法的流程图；
图 5A 为本发明实施例的第一视频帧的示意图；
图 5B 为本发明实施例的第二视频帧的示意图；
图 5C 为本发明实施例的单人的坐标框的示意图；
图 5D 为本发明实施例的双人的坐标框的示意图；
图 6 为本发明实施例中确定裁剪框的第一种方式的流程图；
图 7 为本发明实施例中裁剪宽度与 $\delta W/Width$ 的对应关系；
图 8A-8C 为本发明实施例中确定裁剪框的第一种方式的示意图；
图 9 为本发明实施例中确定裁剪框的第二种方式的示意图；
图 10 为本发明实施例中检测出的人体关节节点的示意图；
图 11 为本发明实施例中关节节点与裁剪位置的对应关系的示意图；
图 12 为本发明实施例中调整目标对象之后所输出的第三视频帧的示意图；
图 13A 为本发明实施例中包含两个人物的原始视频帧的示意图；
图 13B 为本发明实施例中图 13A 中的两个人有一个不在原始视频帧的示意图；
图 13C 为本发明实施例中图 13B 中离开的目标对象又回到视频帧的示意图；
图 13D 为本发明实施例中基于图 13C 所示的原始视频帧所输出的视频帧的示意图；
图 13E 为本发明实施例中基于图 13B 所示的原始视频帧在预设时间段所输出的视频帧的示意图；
图 13F 为本发明实施例中基于图 13B 所示的原始视频帧在预设时间段之后所输出的视频帧的示意图；
图 14 为本发明实施例第二方面所介绍的视频采集方法的流程图；
图 15A 为本发明实施例所介绍的又一原始视频帧的示意图；
图 15B 为本发明实施例中基于图 15A 的原始视频帧输出的视频帧的示意图；
图 15C 为本发明实施例中图 15B 所示的画面向左切换之后的示意图；
图 15D 为本发明实施例中图 15B 所示的画面向左切换之后的示意图；
图 16 为本发明实施例中对视频帧进行放大操作的示意图；

图 17 为本发明实施例第三方面提供的视频采集方法的流程图；

图 18A 为本发明实施例中采集的视频帧中包含多个人的示意图；

图 18B 为本发明实施例中对 18A 中的人物添加聚光灯效果的示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行描述。其中，在本申请实施例的描述中，除非另有说明，“/”表示或的意思，例如，A/B 可以表示 A 或 B；本文中的“和/或”仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。

以下，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者更多个该特征。在本申请实施例的描述中，除非另有说明，“多个”的含义是两个或两个以上。

首先，介绍本申请以下实施例中提供的示例性的电子设备 100。

图 1A 示出了电子设备 100 的结构示意图。

下面以电子设备 100 为例对实施例进行具体说明。应该理解的是，图 1A 所示电子设备 100 仅是一个范例，并且电子设备 100 可以具有比图 1A 中所示的更多的或者更少的部件，可以组合两个或多个的部件，或者可以具有不同的部件配置。图中所示出的各种部件可以在包括一个或多个信号处理和/或专用集成电路在内的硬件、软件、或硬件和软件的组合中实现。

电子设备 100 可以包括：处理器 110，外部存储器接口 120，内部存储器 121，通用串行总线(universal serial bus, USB)接口 130，充电管理模块 140，电源管理模块 141，电池 142，天线 1，天线 2，移动通信模块 150，无线通信模块 160，音频模块 170，扬声器 170A，受话器 170B，麦克风 170C，耳机接口 170D，传感器模块 180，按键 190，马达 191，指示器 192，摄像头 193，显示屏 194，以及用户标识模块(subscriber identification module, SIM)卡接口 195 等。其中传感器模块 180 可以包括压力传感器 180A，陀螺仪传感器 180B，气压传感器 180C，磁传感器 180D，加速度传感器 180E，距离传感器 180F，接近光传感器 180G，指纹传感器 180H，温度传感器 180J，触摸传感器 180K，环境光传感器 180L，骨传导传感器 180M 等。

可以理解的是，本发明实施例示意的结构并不构成对电子设备 100 的具体限定。在本申请另一些实施例中，电子设备 100 可以包括比图示更多或更少的部件，或者组合某些部件，或者拆分某些部件，或者不同的部件布置。图示的部件可以以硬件，软件或软件和硬件的组合实现。电子设备 100 的详细结构介绍，请参考在先专利申请：CN201910430270.9。

图 1B 是本申请实施例的电子设备 100 的软件结构框图。分层架构将软件分成若干个层，每一层都有清晰的角色和分工。层与层之间通过软件接口通信。在

一些实施例中，将 Android 系统分为四层，从上至下分别为应用程序层，应用程序框架层，安卓运行时(Android runtime)和系统库，以及内核层。应用程序层可以包括一系列应用程序包。软件功能的详细介绍，请参考在先专利申请：CN201910430270.9。

第一方面，本发明实施例提供一种视频采集控制方法，应用于电子设备，该电子设备为不包含云台的电子设备，从而该电子设备的摄像头无法旋转，请参考图 3，该方法包括以下步骤：

S200：显示第一视频帧，第一视频帧中包含第一目标对象，且第一目标对象位于第一视频帧的非指定区域，在采集第一视频帧时，所述第一目标对象位于第一地理位置。

S210：响应于检测到进入目标跟踪模式的操作，显示第二视频帧，第二视频帧中包含第一目标对象，第一目标对象位于第二视频帧的指定区域，在采集第二视频帧时，第一目标对象位于第一地理位置；

S220：显示第三视频帧，第三视频帧中包含第一目标对象，第一目标对象位于第三视频帧的指定区域，在采集第三视频帧时，第一目标对象位于第二地理位置，第二地理位置与第一地理位置之间的距离大于预设距离。

在具体实施过程中，在基于步骤 S200 显示第一视频帧之前，需要先采集获得原始视频帧，基于该采集获得的原始视频帧获得输出的视频帧，例如：直接输出该原始视频帧、对该原始视频帧进行各种处理之后获得输出的视频帧等等。

在具体实施过程中，用户可以通过针对相机 APP (Application: 应用程序) 的视频采集操作，从而采集获得原始视频帧，例如：用户首先打开相机 APP，然后点击视频采集按钮（该点击视频采集按钮的操作即为视频采集操作），电子设备检测到该视频采集按钮的操作之后，控制摄像头进行视频采集获得原始视频帧。

电子设备也可以在用户视频聊天过程中，采集获得原始视频帧。例如：用户希望通过即时通信 APP 与对端用户进行视频通信，可以开启即时通讯软件（例如：微信、QQ 等等），然后进入联系人的聊天界面，然后点击“视频通信”按钮，电子设备检测到点击“视频通信”按钮的操作之后，开启与该联系人的视频通信功能，然后开启摄像头，采集获得原始视频帧。

又例如，用户还可以通过系统默认的视频聊天功能与对端用户进行视频聊天，例如，请参考图 3，电子设备的通信功能中包含有畅连通话功能（该畅连通话功能指的是视频通信功能），电子设备的用户打开系统通话应用之后，展示系统通话应用的主界面，该系统通话应用的主界面包含电话控件 31、联系人控件 32、收藏控件 33 和畅连通话控件 30，电话控件 31 用于触发显示电子设备显示最近的通话记录，包括所有通话和未接来电；联系人控件 32 用于触发电子设备显示通话应用的所有联系人信息；个人收藏控件 33 用于触发电子设备将一些联系人添加为个人收藏，通过将联系人添加为个人收藏，可以针对该联系人进行快捷通信，比如：快捷发送短信、快捷拨打电话等等；畅连通话控件 30 则用于触发电子设备开启视频通信功能。

在一种可选的实施方式中，该原始视频帧可以为通过电子设备的前置摄像头、后置摄像头采集的视频帧；在另一种可选的实施方式中，该原始视频帧也可以为与该电子设备存在通信的其他视频采集设备采集的视频帧，例如：该电子设备与家里的安防摄像头存在数据传输通道，则可以获得安防摄像头采集的视频帧；该电子设备与无人机存在数据传输通道，则可以获得无人机采集的视频帧。其中，可以将其他视频采集设备采集的视频流注册为系统的虚拟相机，在视频通话过程中，即时通信软件调用该虚拟相机采集的视频帧作为视频通信过程。

步骤 S200 中，在采集获得原始视频帧之后，可以直接输出该原始视频帧，则该原始视频帧即为第一视频帧，也可以先对该原始视频帧进行图像处理（例如：美化、裁剪等等），然后输出处理后的视频帧（处理后的视频帧即为第一视频帧等等）。

步骤 S200 中，第一物理位置可以为任意位置，其可以为经度、维度、海拔高度等参数表征。第一目标对象可以包含一个或至少两个目标对象。非指定区域指的是第一视频帧中除指定区域 43 之外的其他区域，指定区域 43 例如为：中心区域、黄金分割点区域、电子设备的用户指定其他任意区域。如图 2 所示，图 2 中的指定区域 43 指的是中心区域，其中，目标对象位于指定区域 43，例如为：目标对象的中心点位于指定区域 43，目标对象位于非指定区域，例如：目标对象的中心点位于非指定区域。

步骤 S210 中，在目标跟踪模式下可以通过对采集的原始视频帧进行裁剪的方式获得输出的视频帧，例如：通过对原始视频帧进行裁剪获得第二视频帧、第三视频帧等等。

如图 2 所示可知，第二视频帧中第一目标对象的地理位置并未发生变化，依然是第一地理位置，但是第一目标对象位于第二视频帧的中心区域（也即：指定区域 43）。其中，可以以第一目标对象为参考对原始图像帧进行裁剪，从而获得第二视频帧。具体如何以第一目标对象为参考对原始图像帧进行裁剪，将在后续介绍。

在具体实施过程中，所述第二视频帧中所述第一目标对象的显示尺寸大于所述第一视频帧中所述第一目标对象的显示尺寸。例如：第一视频帧中第一目标对象的显示尺寸是第二视频帧中第一目标对象的显示尺寸的 0.5 倍、0.6 倍等等。其中，基于第一目标对象在第一视频帧中的宽度不同，第一视频帧中第一目标对象的显示尺寸与第二视频帧中第一目标对象的显示尺寸的比例也不同。

步骤 S210 中，在进入目标跟踪模式之后，需要先确定出第一目标对象，以在后续视频帧中实现对第一目标对象的跟踪，第一目标对象可以采用多种方式确定，下面列举其中的两种进行介绍，当然，在具体实施过程中，不限于以下两种情况。

在具体实施过程中，视频帧中的目标对象可以由电子设备自动确定的目标对象，也可以由基于用户操作而确定的目标对象，下面针对这两种情况分别予以介绍，当然，在具体实施过程中，不限于以下两种情况。

第一种，该目标对象由电子设备自动确定。

示例来说，在进入主体跟踪模式后，电子设备基于目标对象的预设条件自动确定出视频帧中的目标对象，该预设条件例如为：视频帧中所有人、视频帧中的动物、视频帧中的其他活动物体等等。以该预设条件为视频帧中的人为例，可以基于人体检测技术识别出视频帧中包含的所有人，然后将所有人确定为目标对象。

又或者，该预设条件为：视频帧中满足跟踪条件的人（或动物、其他活动物体等），满足跟踪条件为距离原始视频帧（即通过摄像头采集到、尚未进行尺寸处理的视频帧）的左边缘大于第一预设距离、且距离原始视频帧的右边缘大于第二预设距离的人，第一预设距离与第二预设距离可以相同或不同，其例如为：150像素、200像素等等，或者第一预设距离、第二预设距离例如为图像宽度的0.1倍、0.2倍等等。基于该方案，能够保证被跟踪的目标对象为摄像头的视野中较为居中的人。满足跟踪条件还例如为：在原始视频帧中的面积大于预设面积，预设面积例如为：1万像素、2万像素等等。在这种情况下，对于原始视频帧中未采集全的人（动物、或其他活动物体）、或者位于原始视频帧的边缘不太显眼的人（或动物、或其他活动物体），则不属于被跟踪的目标对象，以使视频跟踪更具有针对性。

第二种，该目标对象基于用户操作确定。

(1) 基于电子设备的用户的第一预设操作确定目标对象。例如：在电子设备的显示单元上显示视频帧（可以为原始视频帧、也可以为经处理过的视频帧），用户通过手指点击视频帧中的某个人，电子设备检测到该点击操作之后，则该人为目标对象；又例如：电子设备的用户产生如下语音指令“跟随画面中穿黄衣服的人”，则电子设备识别出视频帧中穿黄衣服的人，并将该人确定为目标对象等等。

(2) 基于原始视频帧中的人物的第二预设操作确定出目标对象，该第二预设操作例如为：跟踪手势、语音指令等等。例如：电子设备采集获得原始视频帧之后，通过人体检测技术识别出原始视频帧中的每个人，然后通过关键点识别技术识别出每个人的关节点（例如：头、颈、肩、手掌、手腕、肘关节等等），基于各个关节点的位置关系判断是否存在跟踪手势，该跟踪手势例如为：举手、比心等等，其中，可以通过手掌、手腕、肘关节是否大致位于一条直线、以及手掌的纵坐标是否高于手腕的纵坐标、手腕的纵坐标是否高于肘关节的纵坐标来确定是否存在举手手势。又例如，电子设备检测到视频帧中的某用户产生语音指令“跟我”，则将该用户确定为目标对象等等。

其中，在目标对象为基于用户操作确定的情况下，该用户操作可以同时为触发进入主体跟踪模式的操作，例如：在电子设备进入视频采集模式之后，并未默认进入主体跟踪模式，也未检测到进入主体跟踪模式的其他操作，则电子设备在检测到用于确认目标对象的用户操作之后，通过响应该用户操作既进入主体跟踪模式、又确定目标对象。

又或者，电子设备在检测到确定目标对象的用户操作时，先判断视频采集过程是否已进入主体跟踪模式，如果已进入主体跟踪模式的话，则基于该用户操作确定出目标对象，然后以该目标对象为中心进行主体跟踪；如果视频采集过程没有进入主体跟踪模式，则不响应该该用户操作。

本发明实施例所确定出的目标对象，可以为一个或至少两个。

步骤 S220 中，由于电子设备并未接收到退出目标跟踪模式的操作，因此电子设备依然处理目标跟踪模式，在目标跟踪模式下，电子设备显示的视频帧随着第一目标对象的移动而移动，保持第一目标对象处于视频帧的指定位置 43。请继续参考图 2，从图 2 可以看出在第三视频帧中，第一目标对象已经远离塔，而朝向树走去，因此所输出的第三视频帧与第二视频帧的画面存在较大差异，但是第一目标对象依然处于视频帧的指定区域 43。

第二方面，请参考图 3，本发明实施例提供一种视频采集方法，包括以下步骤：

S300：采集获得视频帧；

在具体实施过程中，用户可以通过针对相机 APP（Application：应用程序）的视频采集操作，从而采集获得视频帧，例如：用户首先打开相机 APP，然后点击视频采集按钮（该点击视频采集按钮的操作即为视频采集操作），电子设备检测到该视频采集按钮的操作之后，控制摄像头进行视频采集获得视频帧。

电子设备也可以在用户视频聊天过程中，采集获得视频帧。例如：用户希望通过即时通信 APP 与对端用户进行视频通信，可以开启即时通讯软件（例如：微信、QQ 等等），然后进入联系人的聊天界面，然后点击“视频通信”按钮，电子设备检测到点击“视频通信”按钮的操作之后，开启与该联系人的视频通信功能，然后开启摄像头，采集获得视频帧。

又例如，用户还可以通过系统默认的视频聊天功能与对端用户进行视频聊天，例如，请参考图 2，电子设备的通信功能中包含有畅连通话功能（该畅连通话功能指的是视频通信功能），电子设备的用户打开系统通话应用之后，展示系统通话应用的主界面，该系统通话应用的主界面包含电话控件 31、联系人控件 32、收藏控件 33 和畅连通话控件 30，电话控件 31 用于触发显示电子设备显示最近的通话记录，包括所有通话和未接来电；联系人控件 32 用于触发电子设备显示通话应用的所有联系人信息；个人收藏控件 33 用于触发电子设备将一些联系人添加为个人收藏，通过将联系人添加为个人收藏，可以针对该联系人进行快捷通信，比如：快捷发送短信、快捷拨打电话等等；畅连通话控件 30 则用于触发电子设备开启视频通信功能。

在本发明实施例中，该视频帧可以为通过电子设备的前置摄像头、后置摄像头采集的视频帧；也可以为与该电子设备存在通信的其他视频采集设备采集的视频帧，例如：该电子设备与家里的安防摄像头存在数据传输通道，则可以获得安防摄像头采集的视频帧；该电子设备与无人机存在数据传输通道，则可以获得无

人机采集的视频帧。其中，可以将其他视频采集设备采集的视频流注册为系统的虚拟相机，在视频通话过程中，即时通信软件调用该虚拟相机采集的视频帧作为视频通信过程。

S310：在电子设备的显示单元上，输出第一视频帧；

请参考图 5A，为电子设备输出的第一视频帧的示意图，其中 40 为输出的视频帧的外边框。如果该方案用于视频采集过程中，输出第一视频帧例如为将输出在显示单元的视频预览界面上给用户预览；如果该方案应用于视频通信过程，输出第一视频帧例如为将第一视频帧传输至对端电子设备、将第一视频帧显示于视频通信界面等等。

在一种可选的实施例中，在视频采集过程中可以设置聚光灯模式，在聚光灯模式下可以针对特定对象设置聚光灯效果，也即对特定对象进行突出显示，例如：对特定对象设置聚光灯效果（设置高光）、控制特定对象彩色显示、将特定对象之外的其他内容黑白显示、将特定对象之外的其他内容虚化显示，为特定对象增加特效等等。

在具体实施过程中，该特定对象可以采用多种方式确定，下面列举其中的两种进行介绍，当然，在具体实施过程中，不限于以下两种情况。

第一种，通过针对视频帧中特定对象的选择操作，确定出该特定对象。示例来说，该选择操作例如为：点击操作、滑动操作等等。用户可以选择一个特定对象，也可以选择多个特定对象，例如：用户可以通过多次选择操作选择出多个特定对象，用户也可以通过一个操作选择多个特定对象，例如：用户通过双指选择，每个手指对应一个目标，从而同时选择两个特定对象等等。

第二种，通过电子设备的麦克风对声源进行定位，确定出声源所在区域的人物为特定对象。示例来说，该方案可以应用于多人讨论的场景、多人演唱的场景。以视频帧中包含人物甲 40、人物乙 41、人物丙 42、人物丁 43 为例，这五人在讨论问题，在第一时刻，由人物乙 41 发言，则确定人物乙 41 为特定对象，为人物乙 41 添加聚光灯效果；在第二时刻由人物丁 43 发言，则确定人物丁 43 为特定对象，为人物丁 43 添加聚光灯效果，且取消人物乙的聚光灯效果，依次类推。通过定位出特定对象，能够确定出当前发言的人。

第三种，通过人体识别技术确定出视频帧中包含的所有人物，将位于中间位置的人物确定为该特定对象。以视频帧中包含人物甲、人物乙、人物丙、人物丁、人物戊为例，则电子设备识别出这五个人物所在位置之后，确定出人物丙位于中间位置，故而将人物丙确定为特定对象。

第四种，接收用户的语音指令，通过语音指令确定出特定对象。例如：电子设备的用户说：“为中间的人设置聚光灯效果”，则确定特定对象为中间的人（例如：人物丙），又例如，电子设备的用户说“为最高的人设置聚光灯效果”，则确定特定对象为视频帧中最高的人等等。

第五种，对视频帧中采集的人物进行手势识别；将使用预设手势的人确定为特定对象，该预设手势例如为：举手、摆手等等。

在具体实施过程中，可以先检测到进入聚光灯模式的操作，再确定出特定对象，也可以先确定出特定对象，则检测到进入聚光灯模式的操作，从而对特定对象产生聚光灯效果，对此本发明实施例不再详细列举，并且不做限制。

S320：在视频采集模式下，进入主体跟踪模式。

在具体实施过程中，在该主体跟踪模式下输出的视频帧的画面随着目标对象的移动而移动，从而使目标对象位于视频帧的画面的中心位置、或者目标对象位于视频帧的黄金分割点位置、或者用户指定的位置等等。该目标对象可以为、动物、其他移动物体（例如：风筝、汽车、扫地机器人等等）等等。

电子设备在进入视频采集状态之后，可以响应预设操作进入主体跟踪模式，该预设操作例如为：点击显示单元显示的预设按钮的操作、选择视频帧中的特定人的操作、视频帧中的人产生预设手势的操作等等。

在具体实施过程中，电子设备在进入视频采集状态之后，也可以默认进入主体跟踪模式。

S330：在主体跟踪模式下，输出第二视频帧，第一视频帧和第二视频帧中都包含目标对象，第二视频帧中目标对象的显示比例与第一视频帧中目标对象的显示比例不同，且目标对象在第二视频帧中的相对位置与目标对象在第一视频帧中的相对位置不同。

在一种可选的实施例中，第二视频帧中目标对象的显示比例大于第一视频帧中目标对象的显示比例，例如：第二视频帧中目标对象的宽度占画面总宽度的50%以上、第一视频帧中目标对象的宽度占画面总宽度的20%；第二视频帧中目标对象的高度占画面总高度的50%以上、第一视频帧中目标对象的高度占画面总高度30%等等，当然，以上宽度、高度的比例仅仅作为举例，并不作为限制。

目标对象在第二视频帧中的相对位置与目标对象在第一视频帧中的相对位置不同，例如为：第二视频帧所对应的第二比值与第一视频帧所对应的第一比值不同，第二比值为：第二视频帧中目标对象的左边框50a与第二视频帧的左边框之间的距离比上第二视频帧的宽度，第一比值为：第一视频帧中目标对象的左边框50a与第一视频帧的左边框50a之间的距离比上第一视频帧的宽度。又例如，第二比值为：第二视频帧中目标对象的右边框50b与第二视频帧的右边框之间的距离比上第二视频帧的宽度，第一比值为：第一视频帧中目标对象的右边框50b与第一视频帧的右边框之间的距离比上第一视频帧的宽度等等。作为一种可选的实施方式，如果步骤S300中进入了聚光灯模式，则在进入S330中可以保持聚光灯模式，依然对S300中确定出的特定对象产生聚光灯效果；作为另一种可选的实施例，在进入主体跟踪模式之后，还可以保持聚光灯模式，但是调整具备聚光灯效果的特定对象，例如：将特定对象调整为主体跟踪模式下的目标对象；作为另一种可选的实施例，聚光灯模式与主体跟踪模式为多个并列的模式，在检测到进入主体跟踪模式之后，则退出聚光灯模式。

作为一种可选的实施例，如果在S330之前并未进入聚光灯模式，则在S330之后依然可以进入聚光灯模式而不会退出主体跟踪模式，对此，本发明实施例不

再详细列举，并且不做限制。

在一种可选的实施例中，为了保证由第一视频帧与第二视频帧之间切换时，不会出现画面突变，在第一视频帧与第二视频帧之间可以进行平滑处理，例如：在第一视频帧与第二视频帧之间还存在多帧过渡的视频帧，该过渡的视频帧例如为 10 帧、20 帧等等。在具体实施过程中，该目标对象可以为视频帧中的人、动物、其他移动物体（例如：无人机、玩具汽车、气球等等）等等。以图 5A 所示为例，第一视频帧中包含的目标对象例如为：人物 41、人物 42，所输出的第二视频帧例如如图 5B 所示。从图 5A 和图 5B 可知，在进入主体跟踪模式之后，放大了目标对象（人物 41、人物 42）在视频帧中的显示区域，且目标对象在视频帧中的相对位置发生变化，在第一视频帧中，目标对象位于画面的左边部分，而在第二视频帧中，目标对象位于画面的中间部分。

在具体实施过程中，视频帧中的目标对象可以由电子设备自动确定的目标对象，也可以由基于用户操作而确定的目标对象，由于前面已做介绍，故而在不再赘述。

在具体实施过程中，可以通过对原始视频帧进行裁剪，获得输出的第二视频帧，以目标对象为人为例，可以先通过人体检测模型确定出人体的坐标框 50，然后通过该坐标框 50 确定出对视频帧进行裁剪的裁剪框 81，该坐标框 50 可以由坐标框 50 中每个点的坐标表示、也可以由左上角的坐标加右下角的坐标表示、也可以由左下角加右上角的坐标表示等等。如图 5C 所示，为目标对象为一个时，所确定出的坐标框 50 的示意图；如图 5D 所示，为目标对象为两个人（多个人类似）时，所确定出的坐标框 50 的示意图，可以先基于人体检测技术确定出原始视频帧中的每个人的坐标框 50，然后对每个人的坐标框 50 进行合并，确定出目标对象的坐标框 50，图 5D 中，通过左上角的坐标（Xmin, Ymin）与右下角的坐标（Xmax, Ymax）表征该坐标框 50。Xmin 表示 X 轴的最小值，Ymin 表示 Y 轴的最小值，Xmax 表示 X 轴的最大值，Ymax 表示 Y 轴的最大值，视频帧的左上角为原点。

在确定出目标对象的坐标框 50 时，可以仅仅考虑目标对象本身，而不考虑其附着的其他东西，例如：撑的雨伞、戴的帽子、骑的自行车等等；而为了考虑画面的完整性，则在确定目标对象时，还可以考虑目标对象附着的其他东西。

在具体实施过程中，可以通过多种方式确定出裁剪框 81，下面列举其中的几种进行介绍，当然，在具体实施过程中，不限于以下几种情况。

第一种，请参考图 6，可以通过以下方式确定出对原始视频帧进行裁剪的裁剪框 81：

S600：通过坐标框 50 确定出目标对象的宽度 δW ；

示例来说，请参考图 5C 和图 5D，可以通过 Xmax 减去 Xmin 来获得目标对象的宽度 δW 。

S610：获得原始图像帧的宽度 Width；

S620：基于目标对象的宽度 δW 和原始图像帧的宽度 Width 确定出对原始

图像帧进行裁剪的裁剪宽度；

在具体实施过程中，可以通过目标对象的宽度 δW 和图像的宽度 $Width$ 的比值来确定出裁剪宽度，例如：请参考图 7，在 $\delta W/Width$ 小于等于第一预设比例时，裁剪宽度为第一预设比例乘以原始图像帧的宽度，在 $\delta W/Width$ 大于第一预设比例且小于等于第二预设比例时，裁剪宽度为目标对象的宽度 δW ；在 $\delta W/Width$ 大于第二预设比例时时，裁剪宽度为 $Width$ 。第一预设比例例如为 0.3、0.5 等等、第二预设比例例如为 0.6、0.8 等等。当然，第一预设比例还可以为其他值，本发明实施例不再详细列举，并且不做限制，

S630：通过裁剪宽度确定出左裁剪边 81a 和右裁剪边 81b。

示例来说，假设第一预设比例为 0.5，第二预设比例为 0.8，则当 δW 小于等于 0.5 时，裁剪宽度为原始图像帧的 0.5 倍，在 δW 大于 0.5 且小于等于 0.8 时，裁剪宽度为 δW ，在 δW 大于 0.8 时，裁剪宽度为原始图像帧的宽度。

示例来说，假设原始图像帧如图 8A 所示，图 8A 中，80 表示原始视频帧的外框，50 表示目标对象的坐标框，坐标框 50 包括左边框 50a、右边框 50b、上边框 50c、下边框 50d， $\delta W/Width$ 小于 0.5，故而确定出裁剪宽度为 0.5 倍原始视频帧的宽度 $Width$ 。可以通过坐标框 50 确定出目标对象的中心点 82，该中心点 82 的坐标可以通过以下公式计算获得：

$$X \text{ 中心点} = (X_{\max} + X_{\min}) / 2$$

$$Y \text{ 中心点} = (Y_{\max} + Y_{\min}) / 2$$

其中，X 中心点指的中心点在 X 轴方向的坐标，Y 中心点指的是中心点在 Y 轴方向的坐标，可以仅仅确定出 X 中心点，也可以既确定出 X 中心点也确定出 Y 中心点。

在确定出中心点之后，将该中心点向左延伸第一预设宽度 $W1$ ，作垂直于 X 轴方向的直线确定出左裁剪边 81a，将该中心点向右延伸第二预设宽度 $W2$ ，作垂直于 X 轴方向的直线确定右裁剪边 81b，第一预设宽度 $W1$ 与第二预设宽度 $W2$ 之和为裁剪宽度，例如：第一预设宽度 $W1$ 与第二预设宽度 $W2$ 之和为 0.5 乘以 $Width$ 。第一预设宽度 $W1$ 与第二预设宽度 $W2$ 可以相等，则第一预设宽度 $W1$ 与第二预设宽度 $W2$ 均为裁剪宽度的 1/2，例如：1/4* $Width$ ；第一预设宽度 $W1$ 与第二预设宽度 $W2$ 也可以不等，本发明实施例不做限制。

其中，请参考图 8B，假设图 8B 中 $\delta W/Width=0.6$ ，则此时的裁剪宽度等于 δW ，在这种情况下，可以将坐标框 50 的左边框 50a 作为左裁剪边 81a，将坐标框 50 的右边框 50b 作为右裁剪边 81b。

请参考图 8C，假设图 8C 中 $\delta W/Width=0.85$ ，则此时的裁剪宽度 $Width$ ，则这种情况下，原始视频帧的左边框为左裁剪边 81a，原始视频帧的右边框作为右裁剪边 81b。

S640：基于坐标框 50 中目标对象的纵向坐标确定出上裁剪边 81c 和下裁剪边 81d。

示例来说，可以将上边框 50c 向上移动第一预设高度 $H1$ 作为上裁剪边 81c

(在具体实施过程中,也可以将上边框 50c 直接作为上裁剪边 81c);以上裁剪边 81c 向下延伸第二预设高度 H2 作为下裁剪边 81d,第一预设高度 H1 例如为:0.05 倍、0.01 倍(当然也可以为其他值)的原始图像帧的高度,第二预设高度 H2 例如为 0.5 倍、0.6 倍(当然也可以为其他值)的原始图像帧的高度。(见图 8A-8C 所示)。

在具体实施过程中,也可以以下边框 50d 直接作为下裁剪边 81d。还可以将下边框 50d 向下延伸一定距离作为下裁剪边 81d。

在具体实施过程中,也可以通过中心点 82 向上延伸预设高度确定出上裁剪边 81c、通过中心点 82 向下延伸预设高度确定出下裁剪边 81d(其确定方式与左裁剪边 81a、右裁剪边 81b 的确定方式类似,在此不再赘述。)

在具体实施过程中,也可以基于裁剪宽度占原始视频帧的比例确定出裁剪高度,基于裁剪高度确定出上裁剪边 81c 和下裁剪边 81d,其确定方式与基于裁剪宽度确定出左裁剪边 81a 和右裁剪边 81b 的方式类似,在此不再赘述。基于该方案能够保证对原始视频帧进行等比例裁剪,从而在原始视频帧符合视频帧的显示区域的宽高比的情况下,不需要再对裁剪框 81 进行调整,使其符合现实区域的宽高比。

S650:基于上裁剪边 81c、下裁剪边 81d、左裁剪边 81a、右裁剪边 81b 确定出裁剪框 81。从图 8A-图 8C 可知,基于目标对象的宽度 δW 不同,最终所确定出的裁剪框 81 的大小也不同,从而输出的视频帧的画面占原始视频帧的画面的大小也可能不同,而在进入主体跟踪模式之前,第一视频帧占原始视频帧的比例通常是固定的,例如:第一视频帧为原始图像帧的 100%、90%等等,故而基于目标对象的宽度 δW 不同,第二视频帧的画面占第一视频帧的画面的比例也不同。

其中,如果目标对象为一个人,则基于目标对象距离摄像头(电子设备)的距离不同,可能导致 δW 不同,从而导致第二视频帧的画面在第一视频帧所占的比例也不同;如果目标对象为多个人,则基于目标对象距离摄像头的距离不同、两个人之间的距离的不同,可能会导致 δW 不同,从而导致第二视频帧的画面在第一视频帧中所占的比例也不同。

可选的,为了保证各图像帧之间平滑过渡,可以选择当前帧的前预设帧、后预设帧共同确定出目标对象的中心点 82。例如:确定当前帧、前预设帧(例如:10 帧、15 帧等等)、后预设帧(例如:15 帧、20 帧等等)中每一帧的中心点,然后对所有帧的中心点取平均值,获得当前帧的中心点 82。

第二种,请参考图 9,在确定出目标对象的坐标框 50 之后,将左边框 50a 向左移动第三预设宽度 W3,获得左裁剪边 81a(在具体实施过程中,也可以直接将左边框 50a 作为左裁剪边 81a);将右边框 50b 向右移动第四预设宽度 W4 获得右裁剪边 81b(在具体实施过程中,也可以直接以右边框 50b 作为右裁剪边 81b);将上边框 50c 向上移动第三预设高度 H3 作为上裁剪边 81c(在具体实施过程中,也可以将上边框 50c 作为上裁剪边 81c);以下边框 50d 作为下裁剪边

81d (在具体实施过程中,也可以将下边框 50d 下移动第四预设高度 H4,获得下裁剪边 81d),通过上裁剪边 81c、下裁剪边 81d、左裁剪边 81a、右裁剪边 81b 确定出裁剪原始视频帧的裁剪框 81。

前述第三预设宽度 W3、第四预设宽度 W4、第三预设高度 H3、第四预设高度 H4 可以全部相同、部分相同、全部不同,其例如为:100 像素、200 像素、300 像素等等;其例如为原始视频帧的宽度的 0.1 倍、0.2 倍等等;其例如为原始视频帧的高度的 0.05 倍、0.15 倍等等。

通过上述方案,能够保证在视频采集过程中,以目标对象为中心对目标对象进行跟踪。

第三种,可以通过以下方式确定出下裁剪边 81d:确定距离下边框 50d 最近的一个预设关节点,确定预设关节点对应的裁剪位置作为下裁剪边 81d。该预设关节点例如为:踝关节、膝关节、髌关节等等,其中可以通过关键点检测算法确定出人物的关节点,该关键点识别技术例如为:Pictorial Structure 算法、自上而下的关键点检测算法、自下而上的人体关键点检测算法等等。确定的关节点如图 10 所示。

在具体实施过程中,该裁剪位置通常为预设关节点向上移动预设距离,预设距离例如为:①固定值,例如:30 像素、40 像素等等;②人体总高度的特定比例,例如:人体总高度的 1/20、1/30 等等;③视频帧的总高度的特定比例,例如:1/40、1/50 等等。④当前关节点与上一特定关节点的距离的特定比例,例如:1/3、1/4 等等,如图 10 所示,如果当前关节点为踝关节,则上一特定关节点例如为:膝关节;如果当前关节点为膝关节,则上一特定关节点例如为:髌关节;如果当前关节点为髌关节,则上一特定关节点例如为肘关节等等。

其他裁剪边的确定方式可以参考第一种、第二种方式,在此不再赘述。

第四种,可以通过以下方式确定出下裁剪边 81d:基于用户针对视频(和/或图像)的历史操作记录确定出裁剪框 81。例如:根据用户针对视频(和/或图像)的历史采集记录或历史裁剪记录,确定出用户最喜欢采用的裁剪框,例如:该裁剪框中目标对象距离视频帧的各个边框的距离。

可选的,在确定出裁剪框 81 之后,还可以对裁剪框 81 进行调整,在具体实施过程中,可以采用多种调整方式,下面列举其中的几种进行介绍,当然,在具体实施过程中,不限于以下几种情况。

第一种,确定距离下裁剪边 81d 最近的一个预设关节点,确定该预设关节点对应的裁剪位置;将下裁剪边 81d 移动至该裁剪位置。该预设关节点例如为:踝关节、膝关节、髌关节等等,其中可以通过关键点检测算法确定出人物的关节点,该关键点识别技术例如为:Pictorial Structure 算法、自上而下的关键点检测算法、自下而上的人体关键点检测算法等等。确定的关节点如图 9 所示。(主要针对上述第一种、第二种方式确定出裁剪框 81)

在一种实现方式中,该裁剪位置通常为预设关节点向上移动预设距离,预

设距离例如为：①固定值，例如：30 像素、40 像素等等；②人体总高度的特定比例，例如：人体总高度的 1/20、1/30 等等；③视频帧的总高度的特定比例，例如：1/40、1/50 等等。④当前关节与上一特定关节的距离的特定比例，例如：1/3、1/4 等等，如图 11 所示，如果当前关节为踝关节，则上一特定关节例如为：膝关节；如果当前关节为膝关节，则上一特定关节例如为：髌关节；如果当前关节为髌关节，则上一特定关节例如为肘关节等等。

第二种，基于用户针对视频（和/或图像）的历史操作数据确定出图像截止位置，基于图像截止位置对下载剪边 81d 进行调整。该历史操作数据可以包含历史视频（图像）采集数据、历史视频（图像）操作数据等等。

示例来说，可以预先基于人体关节将图像划分为多个人体截止区域，确定用户历史采集的视频帧中各个图像截止区域出现的时长，基于各个图像截止区域出现的时长确定出用户最喜欢的图像截止区域（出现时长最长的图像截止区域），基于该图像截止区域对下载剪边 81d 进行调整。

在一种实施例中，还可以将采集的视频拆分为多帧图像，然后结合电子设备中采集的其他图像，确定每张图像的图像截止区域，确定出出现次数最多的图像截止区域，基于该图像截止区域对下载剪边 81d 进行调整。

在一种实施例中，还可以将用户裁剪过的视频拆分为多帧图像，然后结合电子设备中用户裁剪的其他图像，确定出每张图像的图像截止区域，确定出出现次数最多的图像截止区域，基于该图像截止区域对下载剪边 81d 进行调整。

在基于图像截止区域对下载剪边 81d 进行调整时，可以采用多种方式，下面列举其中的两种进行介绍，当然，在具体实施过程中，不限于以下两种情况。

(1) 可以预先设定图像截止区域与下载剪边 81d 的对应关系，例如如表 1 所示：

图像截止区域	下载剪边 81d
踝关节以下区域	踝关节和膝关节之间的预设位置
踝关节以上、膝关节以下区域（不包含踝关节）	膝关节与髌关节之间的预设位置
膝关节以上、髌关节以下区域（不包含膝关节）	髌关节与肘关节之间的预设位置
髌关节以上、肘关节以下区域（不包含髌关节）	肘关节和颈关节之间的预设位置
肘关节以上（不包含肘关节）	不调整
人体坐标框 50 外（即脚部完整位于视频帧中）	不调整

表 1

在确定出图像截止区域之后，通过该对应关系确定出对应的下载剪边 81d，基于确定出的下载剪边 81d 对已确定出的裁剪框 81 的下载剪边 81d 进行调整。

(2) 在确定出图像截止区域之后，可以判断下载剪边 81d 是否位于该图像

截止区域内，如果位于，则不需要对下截止框进行调整，如果不位于，则可以将下截止框调整到该图像截止区域内。

在具体实施过程中，在基于图像截止区域对下裁剪边 81d 进行调整之前，可以先确定视频帧中的目标对象的数量，在目标对象的数量不大于预设阈值（例如 1、2 时），再通过图像截止区域对下裁剪边 81d 进行调整；如果目标对象的数量大于预设阈值，则不需要通过图像截止区域对下裁剪边 81d 进行调整。基于该方案，能够防止目标数量过多，采用该方案确定下裁剪边 81d 导致的部分目标对象被裁剪掉过多内容。

可选的，在基于图像截止区域对下裁剪边 81d 进行调整之前，还可以先判断当前视频帧相对于上一帧的运动量；在运动量小于预设运动量，通过图像截止区域对下裁剪边 81d 进行调整；如果运动量不小于预设运动量，则不通过图像截止区域点对下裁剪边 81d 进行裁剪。该预设运动量例如为：横坐标运动量小于视频帧宽度的预设比例（例如：0.02、0.025 等等）、纵坐标运动量小于视频帧高度的预设比例（例如：0.025、0.03 等等）。基于该方案，能够防止视频帧中的目标对象出现大幅度运动，基于图像截止区域对下裁剪边 81d 进行调整可能导致的视频过渡不平滑。

第三种，在确定出裁剪框 81 之后，还可以裁剪框 81 的宽高比是否符合预设比值（例如：16:9、4:3 等等），在不符合预设比值的情况下，还可以调整该裁剪框 81，从而使其符合预设比值，例如：如果宽高比小于该预设比值，则可以通过增加宽度使宽高比符合该预设比值，如果宽高比大于该预设比值，则可以通过增加高度使宽高比符合该预设比值，当然，还可以通过其他方式调整该裁剪框 81，使宽高比符合预设阈值，本发明实施例不再详细列举，并且不做限制。其中，如果该方案应用于视频通信过程，则电子设备可以获得对端电子设备的显示屏（或视频显示区域）的宽高比，基于该宽高比确定出预设比值。

第四种，在视频帧中包含多个目标对象时，如果下裁剪边 81d 为从上裁剪边 81c 往下延伸第二预设高度 H2 而确定，则该方法还包括：确定出每个目标对象的中心线，然后判断下裁剪边 81d 是否位于所有目标对象的中心线的下方，如果不位于，则将下裁剪边 81d 往下移动，直到下裁剪边 81d 位于所有的目标对象的中心线的下方，该中心线例如为：以每个目标对象的纵轴坐标的中点为基准所作的平行于 X 轴的线、以每个目标对象的髋关节为基准所作的平行于 X 轴的线等等。请继续参考图 5D，该图中包含两个目标对象，分别为目标对象 60、目标对象 61，目标对象 60 的中心线为 60a，目标对象的中心线为 61a，则确定出的下裁剪边 81d 应当位于目标对象 61 的中心线 61a 的下方。

在具体实施过程中，在基于各目标对象的中心线对下裁剪边调整时，该方法还包括：确定第一目标对象的第一预设关键点与所述第二目标对象的所述第一预设关键点之间的相对距离；判断所述相对距离是否大于预设阈值；如果所述相对距离大于所述预设阈值，确定所述第一中心线和所述第二中心线。

示例来说，第一预设关键点例如为：目标对象的头部、颈部等等；则第一目

标对象的第一预设关键点为第一目标对象的头部，第二目标对象的第一预设关键点为第二目标对象的颈部。其中，该预设阈值例如为：①固定值，例如：30 像素、40 像素、50 像素等等；②第一目标对象或第二目标对象的总高度的预设比例，例如：1/4、1/5 等等；③预裁剪框的预设比例，比如：1/4、1/6 等等。

第五种，判断第一预裁剪框的中心线是否位于第一目标对象中心线的下方，如果所述第一预裁剪框的中心线不位于第一目标对象中心线下方，则将所述第一预裁剪框的上裁剪边向上移动第二预设距离且将所述第一预裁剪框的下裁剪边向下移动第三预设距离，获得所述第一裁剪框。最终输出的视频帧为第一裁剪框中的内容。

在具体实施过程中，第二预设距离与第三预设距离可以相同或者不同，其例如为：①固定值，例如：30 像素、40 像素等等；②人体总高度的特定比例，例如：人体总高度的 1/20、1/30 等等；③视频帧的总高度的特定比例，例如：1/40、1/50 等等。④第一预裁剪框的特定比例，例如：1/3、1/4 等等。在基于上述主体跟踪模式，对目标对象进行跟踪过程中，还可以重新选择或切换被跟踪的目标对象，下面列举其中的几种切换方式进行介绍。

(1) 电子设备的触控显示单元接收到针对第一对象的点击操作，则控制第一对象作为目标对象，取消其他目标对象。以原始视频帧为图 8A 所示为例，之前的目标对象为人物 41、人物 42，在检测到针对人物 41 的点击操作之后，将人物 41 依然作为目标对象，将人物 42 取消作为目标对象。在这种情况下，电子设备输出第三视频帧，在第二视频帧中的目标对象未发生位移的情况下，第三视频帧中第一对象的位置与第二视频帧中第二对象的位置不同。如图 12 所示，第二视频帧与第三视频帧中，目标对象 41 并未发生位移，但是其在视频帧的画面中的相对位置发生变化。

电子设备检测到在触控显示单元上点击人物 41 所在区域的操作时，则以人物 41 作为目标对象对原始视频帧进行裁剪，在这种情况下，即使原始视频帧中包含人物 42，在裁剪时也不考虑 42，在这种情况下，则输出第三视频帧，第三视频帧中第一对象的位置与第二视频帧中第二对象的位置不同，虽然第三视频帧与第二视频帧的原始视频帧的内容相同。

(2) 在视频帧中第一对象产生跟踪操作；则判断第一对象是否为基于用户操作所确定的目标对象；如果第一对象为基于用户操作所确定的目标对象，则恢复默认主体跟踪模式；如果第一对象并非基于用户操作所确定的目标对象，则将第一对象作为目标对象，且取消其他目标对象。恢复默认主体跟踪模式例如为：恢复以电子设备自动确定的目标主体进行跟踪的主体跟踪模式。

示例来说，用户操作例如为：电子设备的用户的点击操作、语音指令等等，又或者，视频帧中的用户的预设手势、语音指令等等。基于该用户操作，能够将第一对象确定为目标对象，且取消其他目标对象。（如以上第（1）种情况所示）

接着前述第(1)种情况，目前人物 41 基于用户的点击操作确定为目标对象，而人物 42 被取消目标对象，在这种情况下，如果再次检测到人物 41 产生的跟踪

操作（例如：举手、比心）等等，则恢复默认主体跟踪模式，将人物 41、人物 42 都确定为目标对象。

另外，在检测到第一对象产生跟踪操作之后，如果再次检测到第一对象产生跟踪操作的间隔时长大于预设时长（例如 1 秒、2 秒等等），则认定该跟踪操作为有效操作，电子设备可以响应该跟踪操作，否则，认定该跟踪操作为无效操作，电子设备不响应给跟踪操作。

例如：在初始阶段，人物 41 并非由用户操作确定的目标对象，电子设备检测到视频帧中人物 41 的跟踪操作（例如：举手、比心等等），则将人物 41 确定为目标对象，此时及随后的原始视频帧以人物 41 为参照（例如：中心、黄金分割点等）进行裁剪，从而输出的视频帧为以人物 41 为参照输出的视频帧；在再次检测到人物 41 的跟踪操作之后，发现本次操作距离上次操作的时间只间隔了 0.3 秒，则认定本次操作为无效操作，依然以人物 41 为中心对原始视频帧进行裁剪；随后再次检测到人物 41 的跟踪操作，发现本次操作距离上次操作的时间间隔了 7 秒，则认定本次操作为有效操作，将人物 41、人物 42 共同作为目标对象，恢复默认的主体跟踪模式。

(3) 在第一对象作为用户操作确定的目标对象时，如果在电子设备的触控显示单元上检测到针对第一对象的点击操作，则保持第一对象作为目标对象（此时其他目标对象依然被取消）。

(4) 在第一对象作为用户操作确定的目标对象时，如果检测到原始视频帧中的第一对象产生跟踪操作，则取消第一对象作为目标对象。

可选的，如果触发第一对象作为目标对象的操作为跟踪操作，在再次检测到跟踪操作时，先判断该跟踪操作是否为有效操作，如果为有效操作，则恢复默认的主体跟踪模式；否则，保持第一对象作为目标对象。

以人物 42 为例，电子设备检测到在触控显示单元上点击人物 42 的操作，则控制人物 42 作为目标对象；随后，电子设备检测到原始视频帧中人物 42 产生跟踪操作，在这种情况下，恢复默认主体跟踪模式，将人物 41 和人物 42 作为目标对象。而如果初始阶段，电子设备检测到原始视频帧中人物 42 的跟踪操作，则控制人物 42 作为目标对象，且取消人物 41 作为目标对象；随后，电子设备再次检测到原始视频帧中人物 42 的跟踪操作，则先判断该跟踪操作是否为有效操作，在该跟踪操作为有效操作的情况下，恢复默认的主体跟踪模式；如果该跟踪操作并非有效操作，则保持人物 42 作为目标对象。

(5) 在第一对象作为基于用户操作确定的目标对象时，如果检测到控制第二对象作为目标对象的操作，则将目标对象由第一对象切换至第二对象。控制第二对象作为目标对象的操作，例如为：在电子设备的触控显示单元上点击第二对象所在区域的操作、视频帧中第二对象产生的跟踪操作等等。

示例来说，在人物 42（第一对象）作为目标对象（人物 41 并非目标对象）时，电子设备检测到在触控显示单元上点击人物 41（第二对象）所在区域的操作，则将人物 41 作为目标对象，且取消人物 2 作为目标对象。后续裁剪原始视

视频帧时，以人物 41 为基准（例如：中心、黄金分割点等等）进行裁剪，而不需要考虑人物 42 所在位置，从而输出的视频帧是以第二对象（人物 41）为中心或黄金分割点的视频帧。又例如，在人物 42 作为用户操作所确定的目标对象时，电子设备检测到原始视频帧中人物 41 存在跟踪操作（例如：举手动作），则取消人物 42 作为目标对象，而将人物 41 作为目标对象。

(6) 检测到针对触控显示单元的第一区域的操作之后，如果确定出该第一区域为空白区域，则恢复默认主体跟踪模式。其中，空白区域指的是没有目标对象的区域；或者，空白区域指的是没有目标对象也没有其他运动物体的区域。

上述 (1) - (6) 种切换过程中，针对的是重新选择或切换单个目标对象的场景，针对多个目标对象也可以进行选择或切换，下面列举其中的几种进行介绍，当然，在具体实施过程中，不限于以下几种情况。

(1) 接收在触控显示单元上的区域选择操作；响应该区域选择操作，确定出被选择的区域；将位于被选择区域内的对象作为目标对象。该区域选择操作例如为：绘制封闭区域（例如：圈、框等等），将位于该封闭区域内的对象作为目标对象；该区域选择操作又例如为，划线操作，将该划线操作的划线路径经过的对象作为目标对象等等。

(2) 检测原始视频帧中各个人产生的跟踪操作，如果多个人中任意相邻两个人产生跟踪操作的时间间隔小于预设时间间隔（例如：2 秒、3 秒）等等，则将这多个都确定为目标对象。

例如：原始视频帧中包含五个人，分别为人物甲、人物乙、人物丙、人物丁、人物戊，则先检测到原始视频帧人物甲的举手操作，则将人物甲作为目标对象；1 秒之后，检测到原始视频帧中人物丙的举手操作，则将人物丙与人物甲共同作为目标对象；又 1 秒之后，检测到原始视频帧中人物丁的举手操作，则将人物甲、人物丙与人物丁共同作为目标对象。

(3) 响应用户的语音指令，基于语音指令将多个人作为目标对象。

示例来说，原始视频帧中包含五个人，从左到右分别为人物甲、人物乙、人物丙、人物丁、人物戊，电子设备的用户产生如下语音指令“对从左到右第一个、第三个、第四人进行跟踪”，电子设备响应该语音指令，先识别出原始视频帧中包含的五个人，然后确定出从左到右第一个人为 人物甲、第三个人为 人物丙、第四个人为 人物丁，从而将人物甲、人物丙、人物丁设置为目标对象。

请继续参考图 2，该方法还进一步的包括以下步骤：

S340：继续采集获得原始视频帧，判断是否存在第一目标对象存在于上一原始视频帧且不存在于当前原始视频帧。

示例来说，针对每一帧都可以通过人体识别技术识别出其中所包含的人（目标对象），然后通过比对当前原始视频帧和上一原始视频帧中包含的人，来确定在上一原始视频帧采集到、但是当前原始视频帧未采集到的人作为第一目标对象。

例如：假设上一原始图像帧如图 13A 所示，图 13A 中包含两个目标对象，分别是目标对象 41 和目标对象 42，其中目标对象 42 位于原始图像帧的边缘。当

前的原始图像帧如图 13B 所示，人物 42 已不在原始图像帧，因此可以确定出人物 42 为第一目标对象。

S350：如果存在该第一目标对象，在第一预设时长内，持续检测第一目标对象是否再次出现在原始视频帧中；

该第一预设时长可以通过时间表示，例如：2 秒、3 秒等等，该第一预设时长也可以通过视频帧的数量表示，例如：20 帧、30 帧、40 帧等等。在检测出存在前述第一目标对象之后，电子设备就可以启动计时。

S360：如果第一目标对象在原始视频帧中再次出现，以第一目标对象和其他目标对象为中心对原始视频帧进行裁剪，输出以第一目标对象和其他目标对象为基准的视频帧。

请参考图 13C，假设在 1S 之后，又在视频帧中检测到第一目标对象 42，则继续以第一目标对象 42 和其他目标对象（例如：目标对象 41）为中心进行裁剪，所输出的视频帧如图 13D 所示。

S370：如果第一目标对象在原始视频帧中未再次出现，在第一预设时长内，以第一目标对象的原始位置和其他目标对象为基准裁剪原始视频帧，从而输出基于第一目标对象的原始位置和其他目标对象的位置所确定的视频帧；

示例来说，请参考图 13B，第一目标对象的原始位置例如为第一目标对象 42 最后一次在原始视频帧中出现时的位置 42a。在这种情况下则以人物 41 和第一目标对象 42 的原始位置 42a 为参考对原始视频帧进行裁剪，所输出的视频帧例如如图 13E 所示。

S380：在第一预设时长之后，在对原始视频帧进行裁剪时不再考虑第一目标对象的原始位置，而考虑原始视频帧中剩余的目标对象，所输出的视频帧为基于原始视频帧中剩余的目标对象所确定的视频帧，所输出的视频帧例如如图 13F 所示。

基于上述方案，能够在视频采集过程中，基于目标对象获得输出的视频帧时，在目标对象短暂退出摄像头的视野后又出现在摄像头的视野的情况，能够保证视频帧画面的流畅性。

第二方面，本发明实施例提供一种视频采集方法，请参考图 14，包括以下步骤：

S1400：采集获得视频帧；该采集与步骤 S300 类似，在此不再赘述。

S1410：在电子设备的显示单元上，输出第一视频帧；该步骤与步骤 S310 类似，在此不再赘述。

S1420：在视频采集模式下，进入主体跟踪模式。该步骤与 S320 类似，在此不再赘述。

S1430：在主体跟踪模式下，输出第二视频帧，第一视频帧和第二视频帧中都包含目标对象，第二视频帧中目标对象的显示比例与第一视频帧中目标对象的显示比例不同，且目标对象在第二视频帧中的相对位置与目标对象在第一视频帧

中的相对位置不同。该步骤与 S330 类似，在此不再赘述。

S1440：检测到画面切换操作，输出第三视频帧，其中，在目标对象未发生位移的情况下，第三视频帧中显示的画面相对于第二视频帧中显示的画面发生移动。

该画面切换操作可以为多种不同的操作，从而画面移动的方式也不同，下面列举其中四种进行介绍，当然在具体实施过程中，不限于以下四种情况。

第一种，请参考步骤 S1440A，响应于第一画面切换操作，将视频帧的画面向左切换，其中可以通过调整裁剪框 81 的方式实现画面切换。下面列举其中的两种调整方式，当然，在具体实施过程中，不限于以下两种情况。

①将视频帧的左边框 50a 作为左裁剪边 81a，将 X_{max} 的预设比例（例如：0.1、0.3、0.5 等等）作右裁剪点，以该右裁剪点做垂直于 X 轴的线获得右裁剪边 81b，将第二视频帧的上裁剪边 81c、下裁剪作为当前视频帧的上裁剪边 81c、下裁剪边 81d（或者上下波动第一预设距离，例如：20 像素、50 像素等等）；又或者，以上一原始视频帧的目标对象的中心点向上延伸 1/4 个原始视频帧的高度作平行线作为上裁剪边 81c，以上一原始视频帧的中心点向下延伸 1/4 个视频帧的高度作平行线作为下裁剪边 81d。假设原始视频帧如图 15A 所示，基于该方案使画面向左切换之后最终输出的视频帧例如如图 15B 所示。从而最终输出的视频帧为原始视频帧中左侧的部分。

② 将裁剪框 81 整体左移第二预设距离，例如：原始图像帧的宽度的 20%、30% 等等，可选的，还可以将裁剪框 81 上下浮动第一预设距离。从而输出的视频帧为第二视频帧向左移动第二预设距离确定的画面。

第一预设操作例如为：视频帧中的人物（可以为任何人，也可以仅仅为目标对象）的手臂向左指的操作、在触控显示单元上将视频帧向右拖动、产生语音指令等等。可以通过关键点识别技术识别出视频帧中的人物的关节点，然后确定出手的肘关节和腕关节的坐标判断是否存在手臂向左指的操作，例如：如果人面对摄像头，肘关节和腕关节的纵坐标值差别不大、肘关节的横坐标大于腕关节的横坐标，则可以认定存在手臂向左指的操作的操作。

第二种，请参考步骤 S1440B，响应于第二画面切换操作，将视频帧的画面向左切换。

其中可以通过调整裁剪框 81 的方式实现画面切换。下面列举其中的两种调整方式，当然，在具体实施过程中，不限于以下两种情况。

①将视频帧的右边框 50b 作为右裁剪边 81b，将 X_{max} 减去 X_{max} 的预设比例（例如：0.1、0.3、0.5 等等）作为左裁剪点，以该左裁剪点做垂直于 X 轴的线获得左裁剪边 81a，将第二视频帧的上裁剪边 81c、下裁剪作为当前视频帧的上裁剪边 81c、下裁剪边 81d（或者上下波动第一预设距离，例如：20 像素、50 像素等等）；又或者，以上一原始视频帧的目标对象的中心点向上延伸 1/4 个原始视频帧的高度作平行线作为上裁剪边 81c，以上一原始视频帧的中心点向下延伸 1/4 个视频帧的高度作平行线作为下裁剪边 81d。从而最终输出的视频帧为原

始图像帧的右边的部分。

假设原始视频帧如图 15A 所示，基于该方案使画面向右切换之后最终输出的视频帧例如如图 15C 所示。

② 将裁剪框 81 整体右移第三预设距离，例如：原始图像帧的宽度的 20%、30% 等等，可选的，还可以将裁剪框 81 上下浮动第一预设距离。从而最终输出的视频帧为第二视频帧的画面向右移动第三预设距离的画面。

第三种，请参考步骤 S1440C，响应于第三画面切换操作，将视频帧的画面向上切换。第三画面切换操作例如为在显示单元从上往下拖动的操作、手臂从上往下摆动的操作、语音指令等等。画面向上切换的实现方式与画面向左、向右切换的实现方式类似，在此不再赘述。

第四种，请参考步骤 S1440D，响应于第四画面切换操作，将视频帧的画面向下切换。第四画面切换操作例如为在显示单元从下往上拖动的操作、手臂从下往上摆动的操作、语音指令等等。画面向下切换的实现方式与画面向左、向右切换的实现方式类似，在此不再赘述。

S1450：在预设时间后，恢复主体跟踪模式，输出第四视频帧，在目标对象的位移小于第一预设位移的情况下，第四视频帧中目标对象的相对位置与第二视频帧中目标对象的相对位置的偏移量小于第二预设位移。也即是说，如果在输出第二视频帧和第四视频帧时，目标对象并未移动或者移动量较小，则第二视频帧和第四视频帧的差别也较小。

预设时间例如为：2 秒、3 秒等等，本发明实施例不做限制。如何基于目标对象对视频帧进行裁剪，前面已做介绍，故而在不再赘述。

在具体实施过程中，以上视频采集方法还可以包括以下步骤响应焦点集中操作，将焦点集中到该人物自身，例如：放大该人物在视频帧中的比例、对背景区域进行虚化处理、为该人物增加特效等等。第五预设操作例如为：双击视频帧中该人物所在区域的操作、产生特定手势的操作等等。

在基于 S1430 输出第二视频帧的之后，该方法还包括：响应用户操作，退出主体跟踪模式，输出第四视频帧，第四视频帧中目标对象的显示比例与第二视频帧中目标对象的显示比例不同，且目标对象在第四视频帧中的相对位置与目标对象在第二视频帧中的相对位置不同。例如：第四视频帧中目标对象的显示比例小于第二视频帧中目标对象的显示比例。在目标对象未发生位移的情况下，第四视频帧与第二视频帧类似，在此不再赘述。

在基于 S1430 输出第二视频帧之后，该方法还包括：响应放大操作，输出第五视频帧，第五视频帧中目标对象的显示尺寸大于第二视频帧中目标对象的显示尺寸。

示例来说，该放大操作例如为：预设手势（例如：手掌向外推、五指张开等等）、语音指令等等。如图 16 所示，为第二视频帧与第五视频帧的对比示意图，其中基于该放大操作可以逐渐放大目标对象的显示尺寸，从而输出多帧目标对象的显示尺寸越来越大的视频帧，如图 16 所示，在输出第二视频帧之后，输出第

五视频帧（一），然后输出第五视频帧（二），从而实现平滑过渡。

本发明第三实施例提供一种视频采集方法，请参考图 17，包括：

S1700：采集获得视频帧，该步骤与 S300 类似，在此不再赘述；

S1710：确定该视频帧中的特定对象；

在具体实施过程中，该特定对象可以为单个特定对象、也可以为至少两个特定对象，该特定对象可以为入、动物、其他移动物体等等。该目标物体可以通过多种方式确定，下面列举其中的几种进行介绍，当然，在具体实施过程中，不限于以下几种情况。

第一种，通过针对视频帧中的该特定对象的选择操作，从而确定该特定对象。

示例来说，请参考图 18A，为一视频通信过程中的视频帧的示意图，该视频帧中包含 5 个人物，分别为人物甲 18a、人物乙 18b、人物丙 18c、人物丁 18d、人物戊 18e，电子设备的用户希望将人物 18b 确定为特定对象，则针对视频帧中的人物乙 18b 产生选择操作（例如：点击、滑动等等），电子设备响应该选择操作，从而将人物乙 18b 确定为特定对象。

其中，用户可以选择一个特定对象，也可以选择多个特定对象，例如：用户可以通过多次选择操作选择出多个特定对象，用户也可以通过一个操作选择多个特定对象，例如：用户通过双指选择，每个手指对应一个目标，从而同时选择两个特定对象等等。

第二种，通过电子设备的麦克风对声源进行定位，确定出声源所在区域的人物为特定对象。

示例来说，该方案可以应用于多人讨论的场景、多人演唱的场景。还是以视频帧中包含人物甲 18a、人物乙 18b、人物丙 18c、人物丁 18d、人物戊 18e 为例，这五人在讨论问题，在第一时刻，由人物乙 18b 发言，则确定人物乙 18b 为特定对象，在第二时刻由人物丁 18d 发言，则确定人物丁 18d 为特定对象，依次类推。通过定位出特定对象，能够确定出当前发言的人。

第三种，通过人物识别技术确定出视频帧中包含的所有人物，将位于中间位置的人物确定为该特定对象。还是以视频帧中包含图 18A 所示的人物甲 18a、人物乙 18b、人物丙 18c、人物丁 18d、人物戊 18e 为例，则电子设备识别出这五个人物所在位置之后，确定出人物丙 18c 位于中间位置，故而将人物丙确定为特定对象。

第四种，接收用户的语音指令，通过语音指令确定出特定对象。例如：电子设备的用户说：“为中间的人设置聚光灯效果”，则确定特定对象为中间的人（例如：人物 18c），又例如，电子设备的用户说“为最高的人设置聚光灯效果”，则确定特定对象为视频帧中最高的人等等。

第五种，对视频帧中采集的人物进行手势识别；将使用预设手势的人确定为特定对象，该预设手势例如为：举手、摆手等等。

S1720：控制该特定对象进入聚光灯模式。

在具体实施过程中，可以在确定出特定对象之后，直接控制该特定对象进入

聚光灯模式；也可以在确定出特定对象之后，接收预设操作，电子设备响应该预设操作，控制特定对象进入聚光灯模式，在该聚光灯模式下对特定对象进行突出显示。

上述步骤 S1710、S1720 可以采用多种实现方式，下面列举其中的两种进行介绍，当然，在具体实施过程中，不限于以下两种情况。

第一种，电子设备的用户产生预设操作，该预设操作用于控制视频通信进入聚光灯模式，该预设操作例如为：点击表征聚光灯模式的预设按钮的操作、产生进入聚光灯模式的语音指令的操作、产生预设手势的操作等等；电子设备检测到该预设操作之后执行步骤 S1610，提示电子设备的用户选择特定对象或者自动确定特定对象；在确定出特定对象之后，电子设备自动执行步骤 S1720，也即：电子设备控制特定对象进行聚光灯模式。

第二种，电子设备基于用户操作确定出特定对象（步骤 S1710），其具体确定方式前面已做介绍，在此不再赘述。在确定出特定对象之后，电子设备的用户产生预设操作，电子设备响应该预设操作，控制特定对象进入聚光灯模式（步骤 S1720）。

在具体实施过程中，聚光灯模式指的是对特定对象突出显示的模式，可以采用多种方式对特定对象进行突出显示，例如：请参考图 18B，对特定对象设置聚光灯效果（设置高光）、控制特定对象彩色显示、将特定对象之外的其他内容黑白显示、将特定对象之外的其他内容虚化显示等等。

其他内容参考上文相关内容的描述，不再赘述。

可以理解的是，上述电子设备等为了实现上述功能，其包含了执行各个功能相应的硬件结构和/或软件模块。本领域技术人员应该很容易意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，本申请实施例能够以硬件或硬件和计算机软件的结合形式来实现。某个功能究竟以硬件还是计算机软件驱动硬件的方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明实施例的范围。

本申请实施例可以根据上述方法示例对上述电子设备等进行功能模块的划分，例如，可以对应各个功能划分各个功能模块，也可以将两个或两个以上的功能集成在一个处理模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。需要说明的是，本发明实施例中对模块的划分是示意性的，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。下面以采用对应各个功能划分各个功能模块为例进行说明：

本申请实施例提供的方法中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例描述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、网络设备、电

子设备或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（digital subscriber line, DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机可以存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质（例如，软盘、硬盘、磁带）、光介质（例如，数字视频光盘（digital video disc, DVD））、或者半导体介质（例如，SSD）等。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统、装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

以上，仅为本申请的具体实施方式，但本申请实施例的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请实施例揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请实施例的保护范围之内。因此，本申请实施例的保护范围应以权利要求的保护范围为准。

权 利 要 求 书

- 1、一种视频处理方法，应用于电子设备，包括：
获取第一视频帧，
确定所述第一视频帧中包含有第一目标对象；
以所述第一目标对象为中心，确定第一裁剪框，所述第一裁剪框的下裁剪边与所述第一目标对象的预设关节点之间存在第一预设距离；
获取所述第一裁剪框中的内容，作为第二视频帧进行输出。
- 2、如权利要求1所述的方法，其特征在于，所述确定第一裁剪框还包括：
按照预设规则确定第一预裁剪框；
判断所述第一预裁剪框的下裁剪边与所述预设关节点之间是否存在所述第一预设距离；
如果所述第一预裁剪框的下裁剪边与所述预设关节点之间不存在所述第一预设距离，将所述第一预裁剪框的下裁剪边向上移动至所述预设关节点对应的裁剪位置，获得所述第一裁剪框。
- 3、如权利要求1所述的方法，其特征在于，以所述第一目标对象为中心，确定第一裁剪框，还包括：
按照预设规则确定第一预裁剪框；
判断所述第一预裁剪框的中心线是否位于第一目标对象中心线下方；
如果所述第一预裁剪框的中心线不位于所述第一目标对象中心线下方，将所述第一预裁剪框的上裁剪边向上移动第二预设距离且将所述第一预裁剪框的下裁剪边向下移动第三预设距离，获得所述第一裁剪框。
- 4、如权利要求1所述的方法，其特征在于，在所述获取所述第一裁剪框中的内容，作为第二视频帧进行输出之后，所述方法还包括：
采集获得第三视频帧；
检测到所述第三视频帧中存在符合预设手势的内容，则确定所述预设手势对应的方向；
基于所述预设手势的方向确定第二裁剪框，所述第二裁剪框相对于所述第一裁剪框存在朝向所述预设手势对应的方向的移动；
获取所述第二裁剪框中的内容，作为第四视频帧输出。
- 5、如权利要求4所述的方法，其特征在于，所述基于所述预设手势的方向确定第二裁剪框，包括：
如果所述预设手势的方向为所述第三视频帧中的人物的手臂向左指的方向，所述第二裁剪框为相对于第一裁剪框向左移动的裁剪框；或，
如果所述预设手势的方向为所述第三视频帧中的人物的手臂向右指的方向，所述第二裁剪框为相对于所述第一裁剪框向右移动的裁剪框；或，
如果所述预设手势的方向为所述第三视频帧中的人物的手臂从上往下摆的方向，所述第二裁剪框为相对于所述第一裁剪框向上移动的裁剪框；或，

如果所述预设手势的方向为所述第三视频帧中的人物的手臂从下往上摆的方向，所述第二裁剪框为相对于所述第一裁剪框向下移动的裁剪框。

6、如权利要求1所述的方法，其特征在于，在所述获取所述第一裁剪框中的内容，作为第二视频帧进行显示之后，所述方法还包括：

采集获得第五视频帧，所述视频帧中包含所述第一目标对象和第二目标对象在内的至少两个目标对象；

以所述至少两个目标对象为中心，确定出第三裁剪框，所述第三裁剪框的下裁剪边位于所述第一目标对象的水平方向的中心线的下方，且位于所述第二目标对象的水平方向的中心线的下方；

获取所述第三裁剪框中的内容，作为第六视频帧输出。

7、如权利要求6所述的方法，其特征在于，所述确定第三裁剪框，包括：

按照预设规则确定出第三预裁剪框；

确定出所述第一目标对象的水平方向的第一中心线和所述第二目标对象的水平方向的第二中心线；

判断所述第三预裁剪框的下裁剪边是否位于所述第一中心线的下方且位于所述第二中心线的下方；

如果所述第三预裁剪框的下裁剪边不位于所述第一中心线的下方，或，所述第三预裁剪框的下裁剪边不位于所述第二中心线的下方，则将所述第三预裁剪框的下裁剪边往下移动，直到所述第三预裁剪框的下裁剪边位于所述第一中心线和所述第二中心线的下方，获得所述第三裁剪框。

8、如权利要求7所述的方法，其特征在于，所述确定出所述第一目标对象的水平方向的第一中心线和所述第二目标对象的水平方向的第二中心线，包括：

确定第一目标对象的第一预设关键点与所述第二目标对象的所述第一预设关键点之间的相对距离；

判断所述相对距离是否大于预设阈值；

如果所述相对距离大于所述预设阈值，确定所述第一中心线和所述第二中心线。

9、如权利要求1所述的方法，其特征在于，在所述获取所述第一裁剪框中的内容，作为第二视频帧进行显示之后，所述方法还包括：

采集获得第七视频帧，所述第七视频帧中包含所述第一目标对象和第四目标对象在内的至少两个目标对象；

判断是否存在聚焦于所述第四目标对象的聚焦操作；

如果存在所述聚焦操作，以所述第四目标对象为中心确定第四裁剪框；

获取所述第四裁剪框中的内容，作为第八视频帧输出。

10、如权利要求1-9任一所述的方法，其特征在于，在所述以所述第一目标对象为中心，确定第一裁剪框之前，所述方法还包括：

采集获得第九视频帧；输出未经裁剪的所述第九视频帧；

检测获得第一操作，响应于所述第一操作进入目标跟踪模式；

所述以所述第一目标对象为中心，确定第一裁剪框包括：在进入所述目标跟踪模式之后，以所述第一目标对象为中心，确定所述第一裁剪框。

11、如权利要求 10 所述的方法，其特征在于，在所述检测获得第一操作，响应于所述第一操作进入目标跟踪模式之后，所述方法还包括：

在所述目标跟踪模式下，检测到第二操作；

响应于所述第二操作，对所述电子设备输出的视频帧中的目标对象采用聚光灯效果。

12、如权利要求 1-9 任一的方法，其特征在于，所述方法应用于视频通话过程中，所述方法获取所述第一裁剪框中的内容，作为第二视频帧进行输出，包括：

将所述第二视频帧传输至对端电子设备；和/或，将所述第二视频帧显示于视频通信界面。

13、一种电子设备，其特征在于，包括：

一个或多个处理器；

一个或多个存储器；

多个应用程序；

以及一个或多个计算机程序，其中所述一个或多个计算机程序被存储在所述一个或多个存储器中，所述一个或多个计算机程序包括指令，当所述指令被所述电子设备的一个或多个处理器执行时，使得所述电子设备如权利要求 1-12 任一所述的方法。

14、一种计算机可读存储介质，包括指令，其特征在于，当所述指令在电子设备上运行时，使得所述电子设备执行如权利要求 1-12 中任一项所述的方法。

15、一种计算机程序产品，其特征在于，所述计算机程序产品包括软件代码，所述软件代码用于执行如权利要求 1-12 中任一项所述的方法。

16、一种包含指令的芯片，其特征在于，当所述芯片在电子设备上运行时，使得所述电子设备执行如权利要求 1-12 中任一项所述的方法。

电子设备100

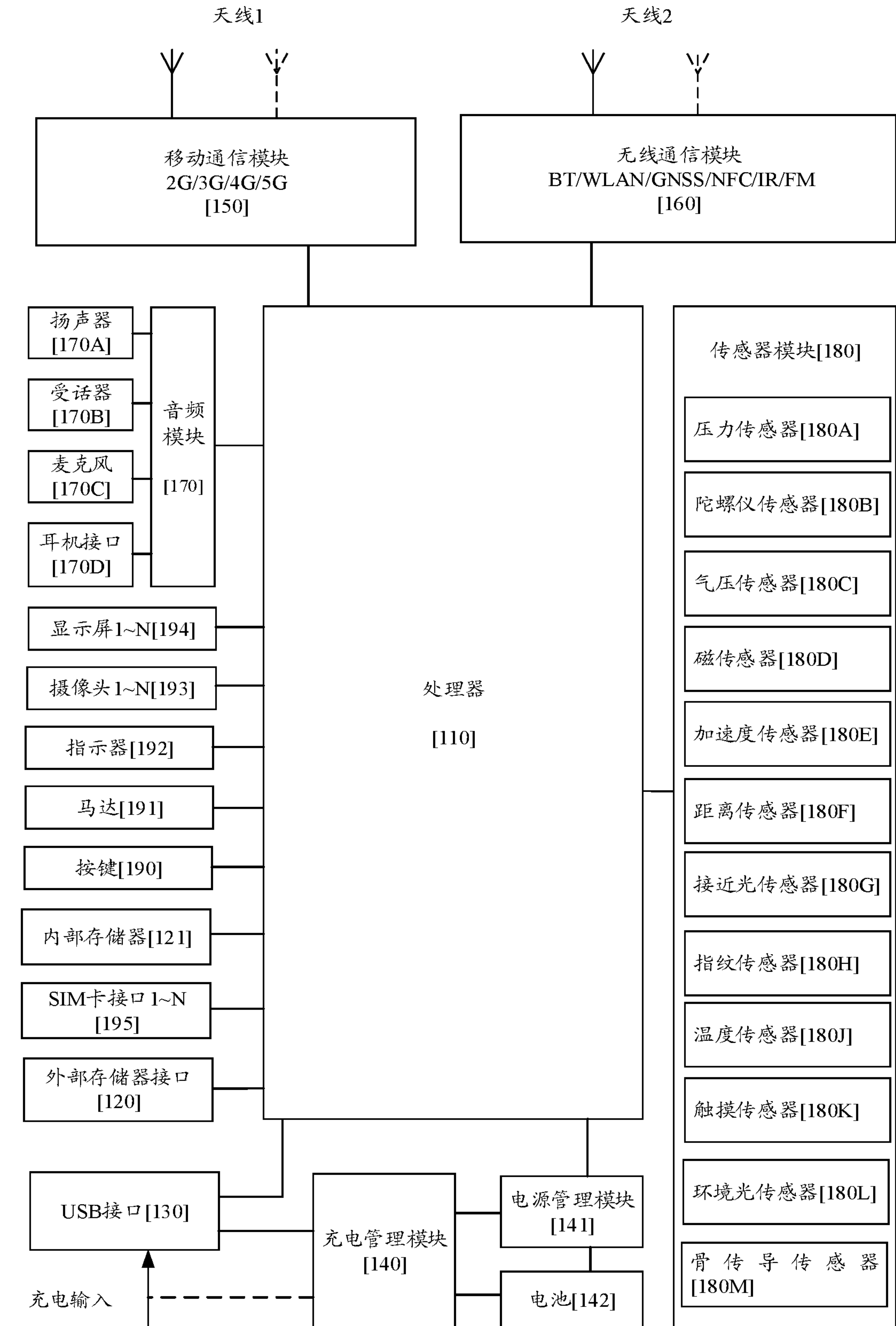


图 1A

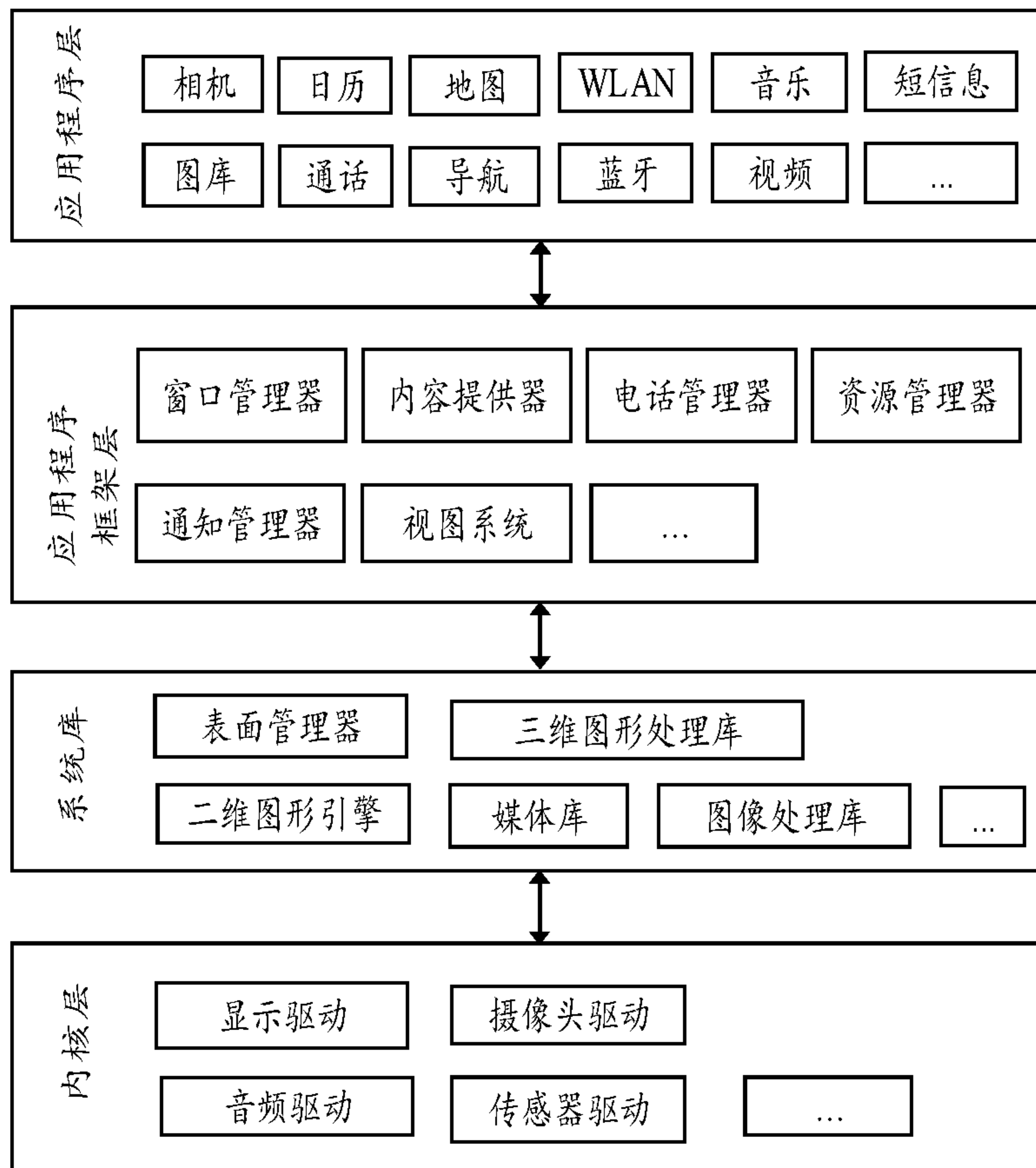


图 1B

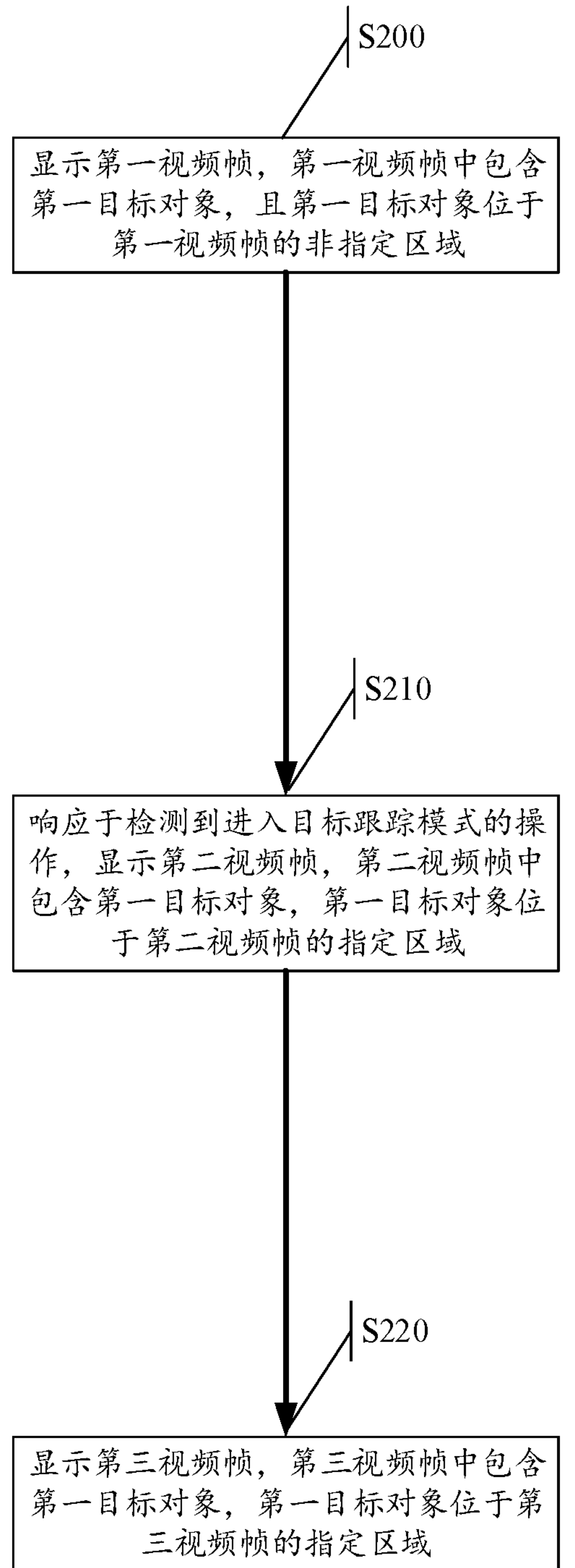
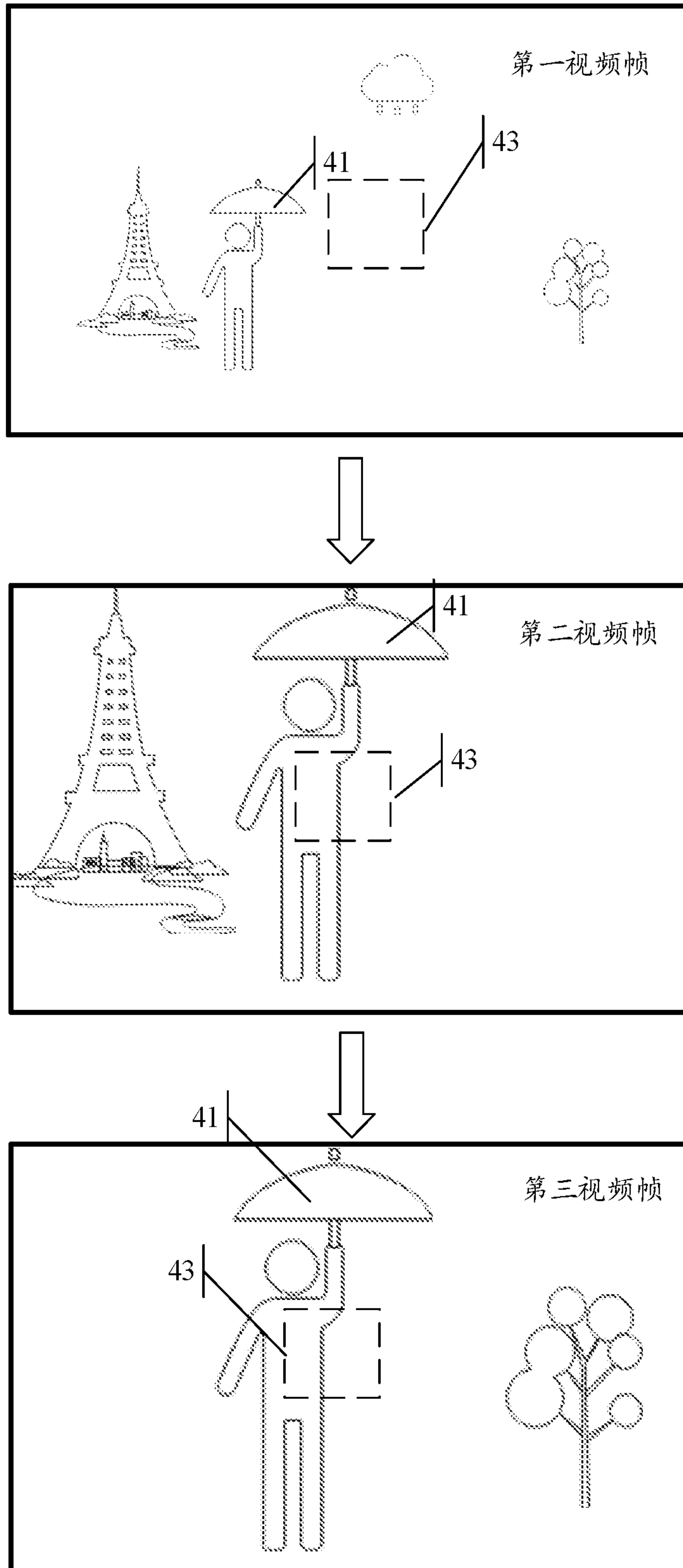


图 2

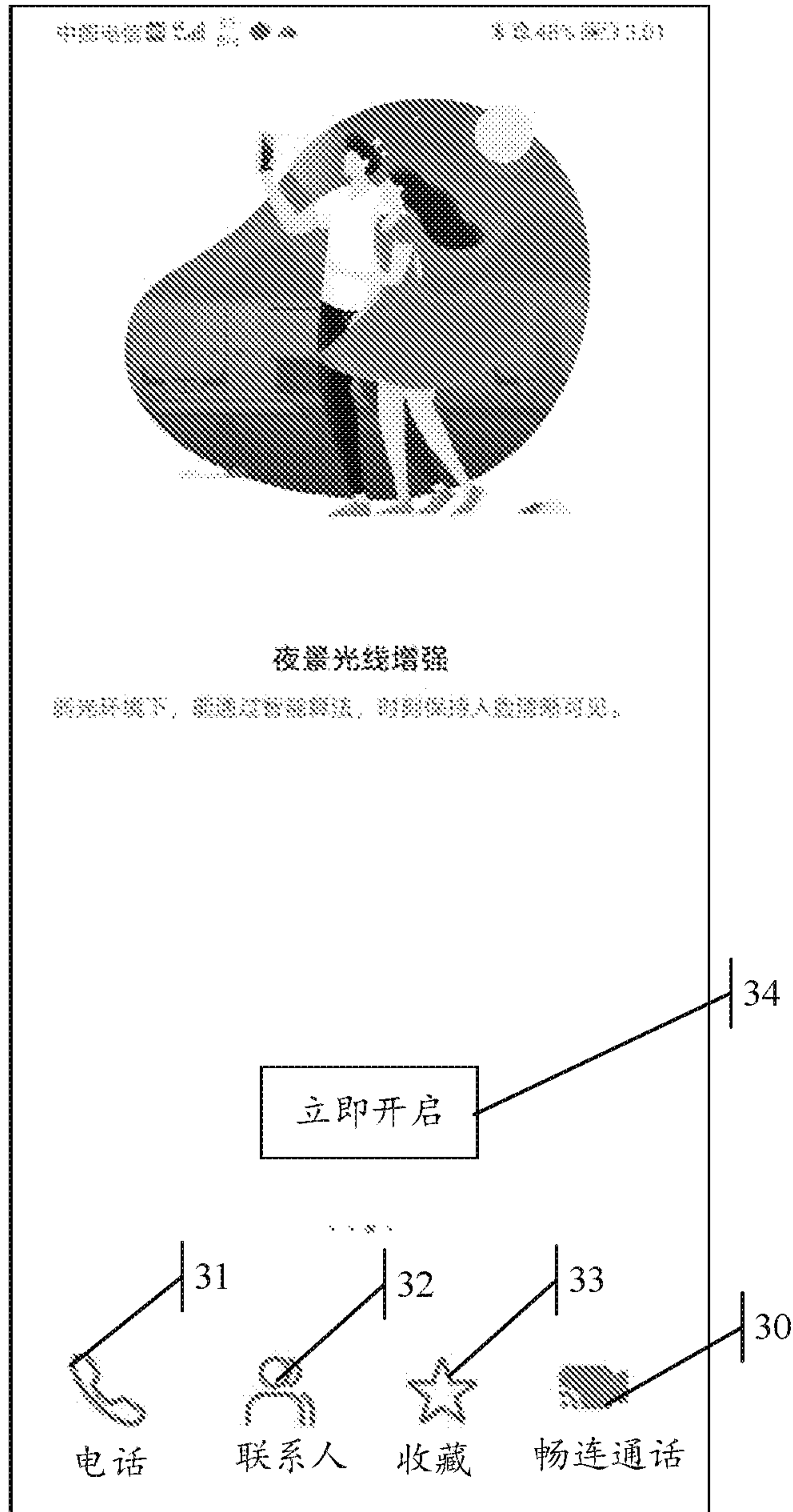


图 3

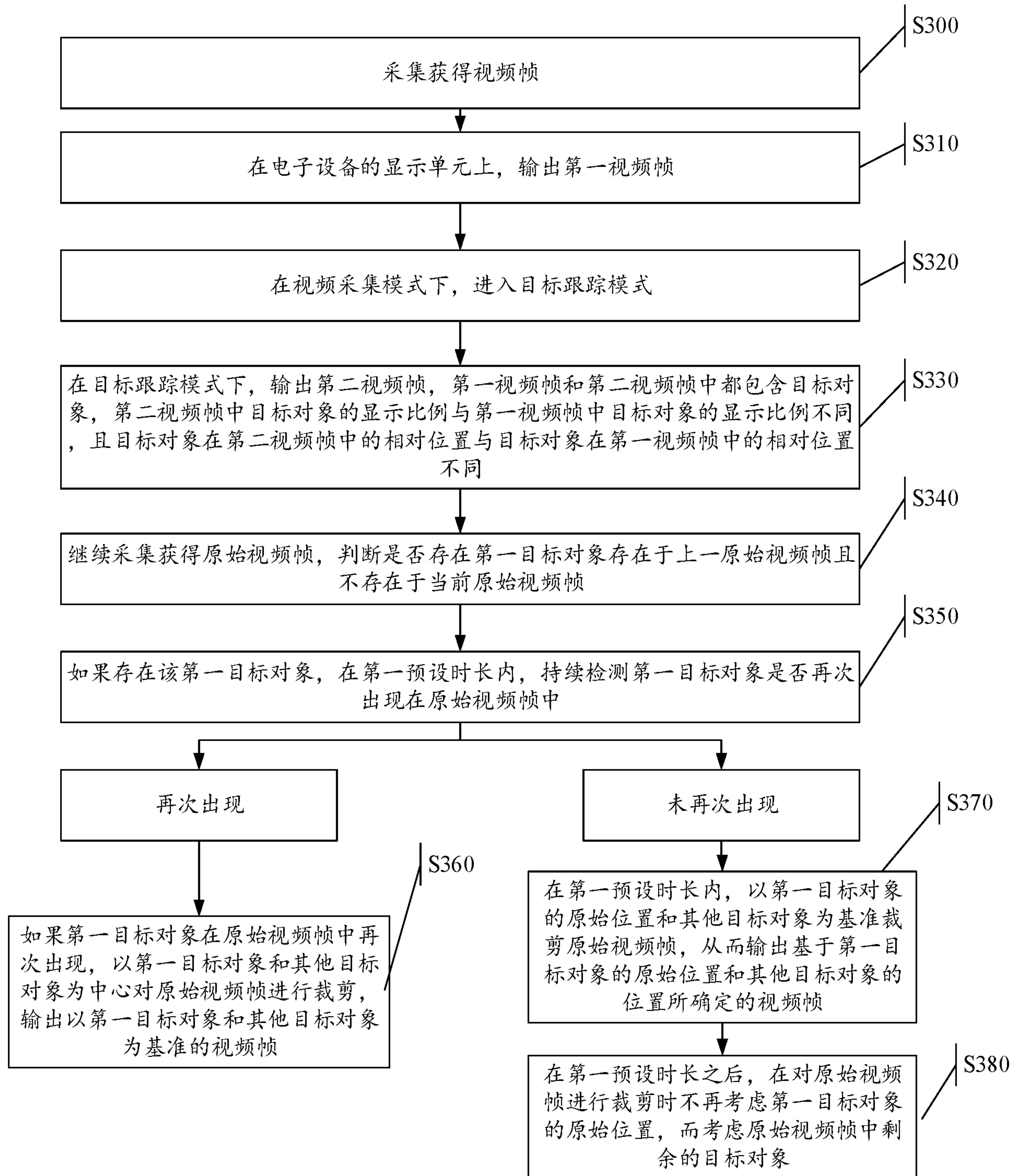


图 4

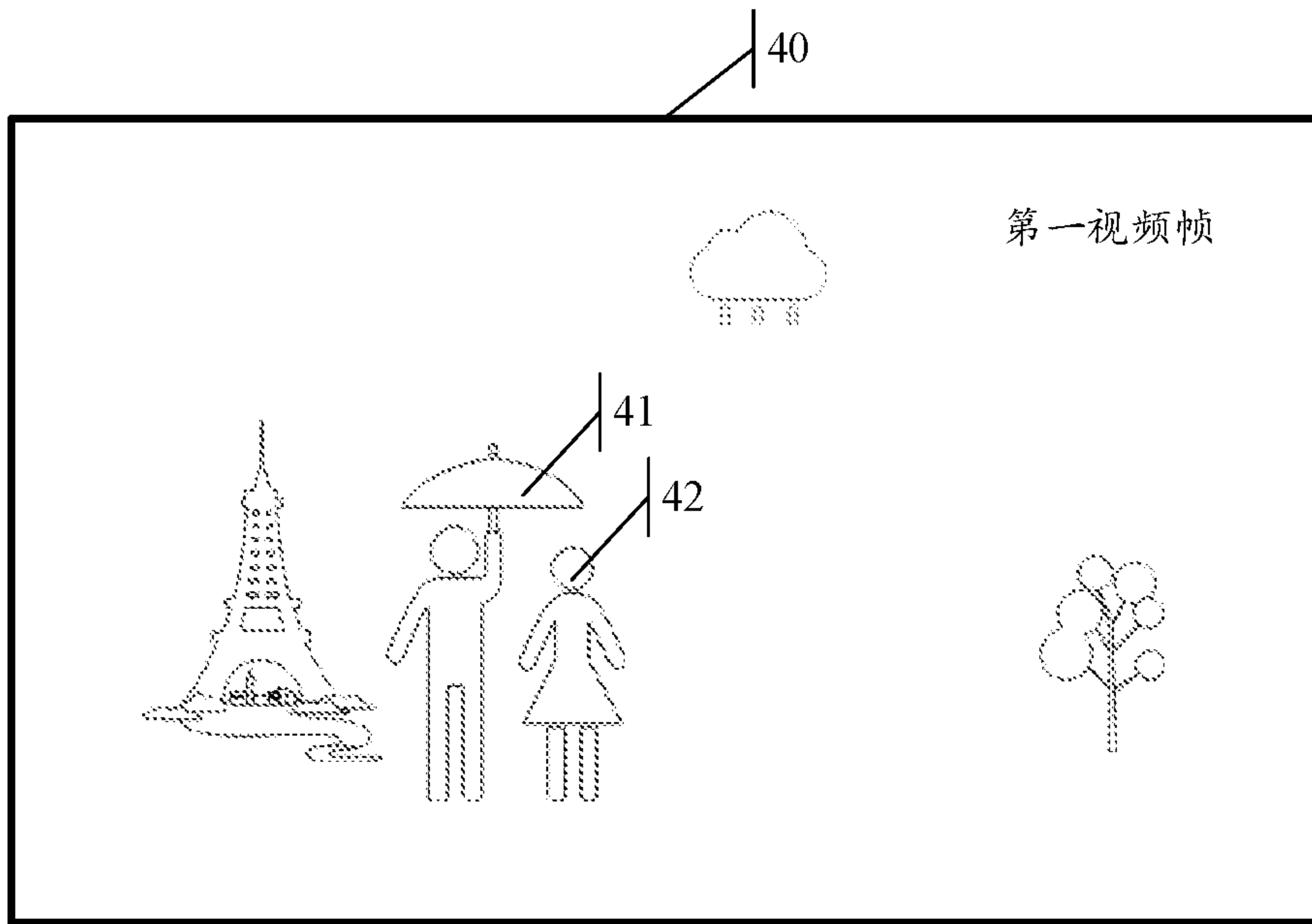


图 5A

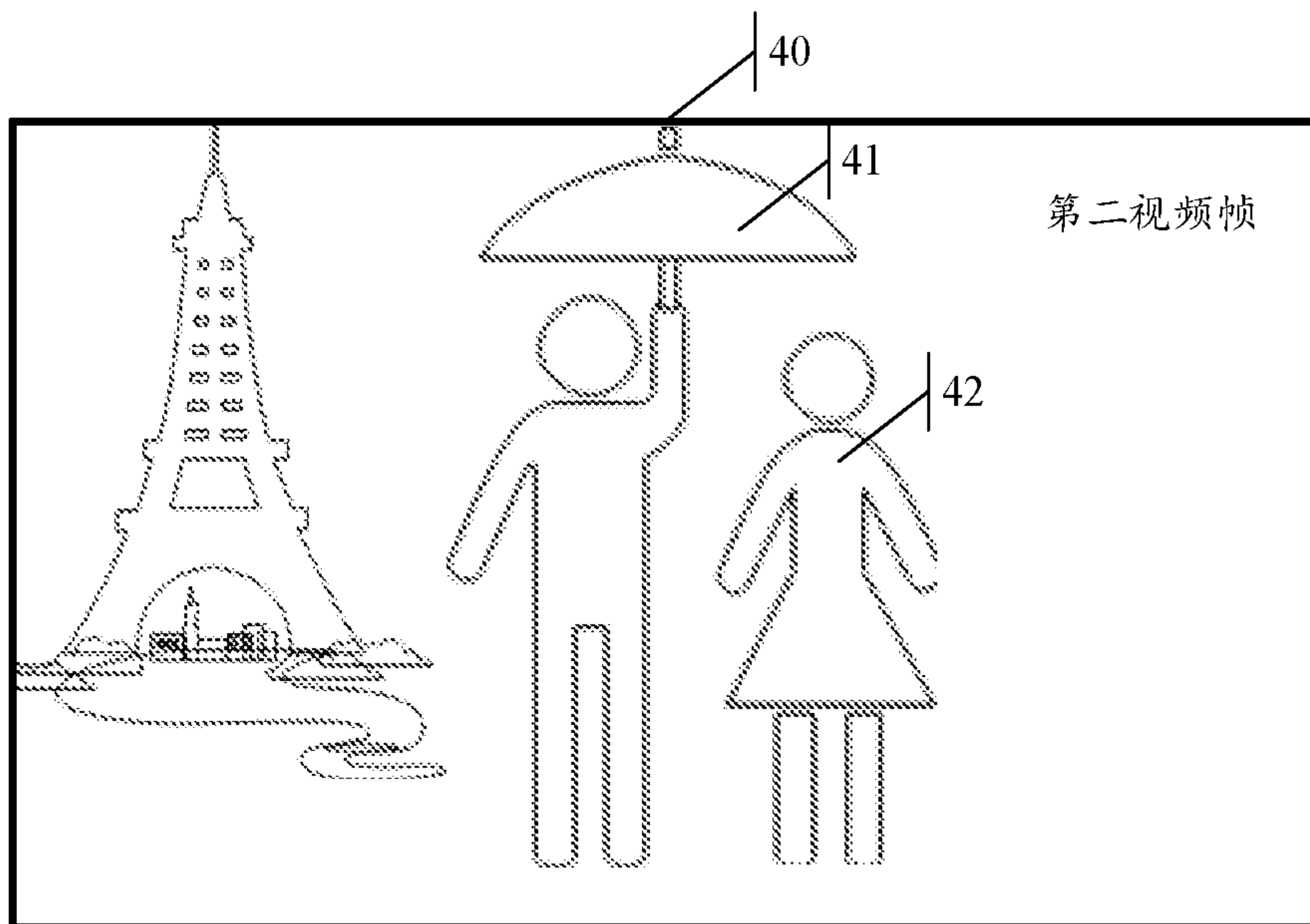


图 5B

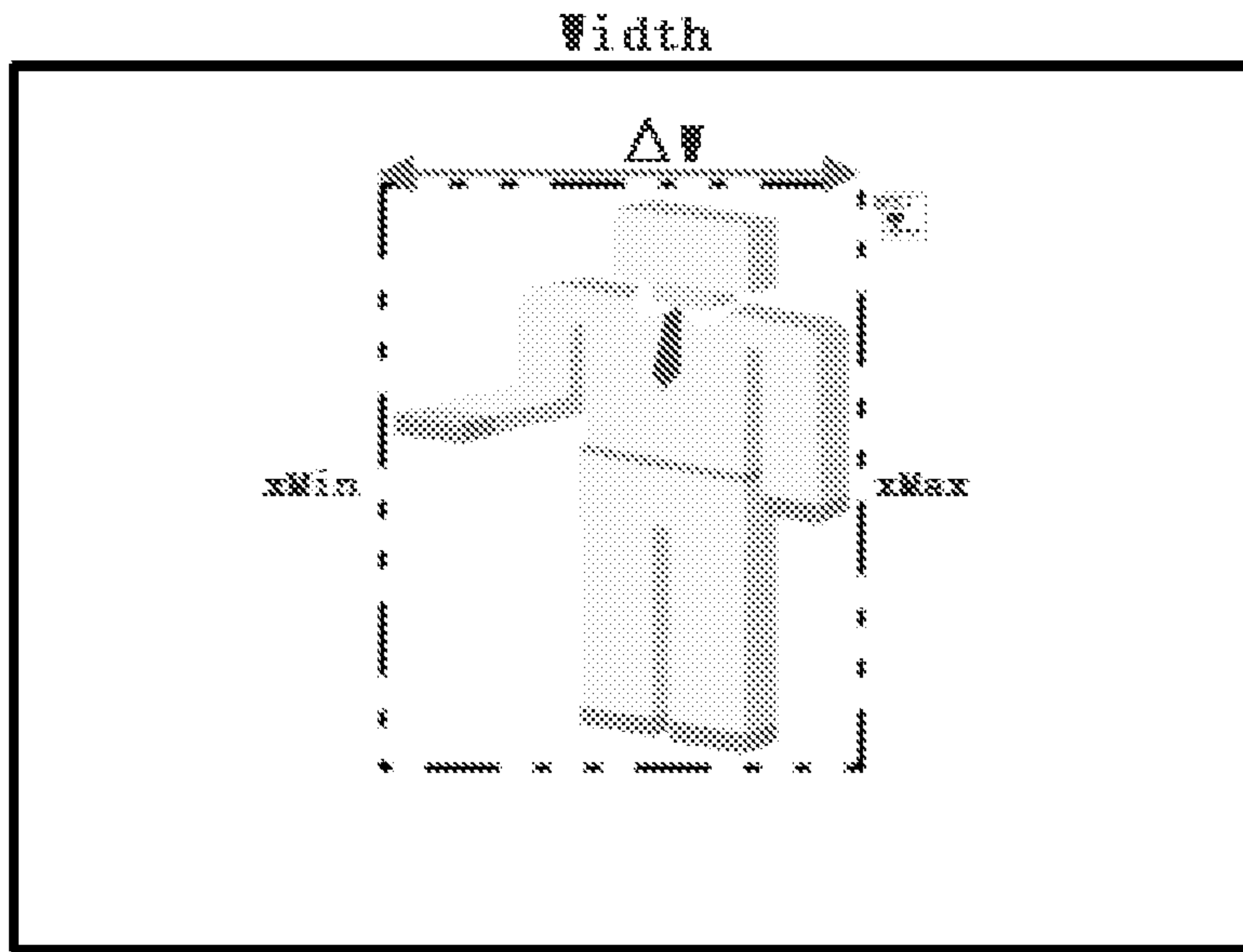


图 5C

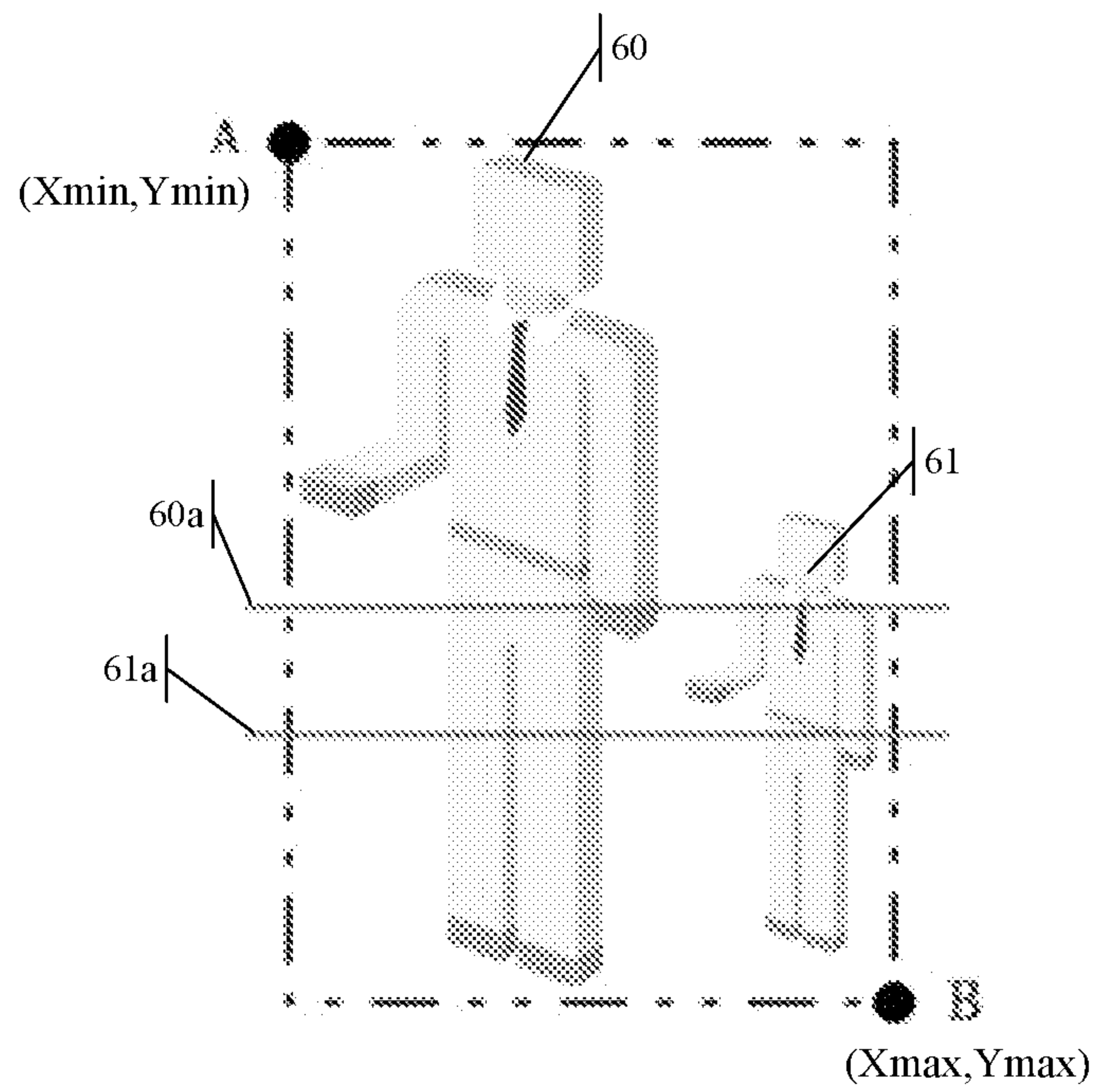


图 5D

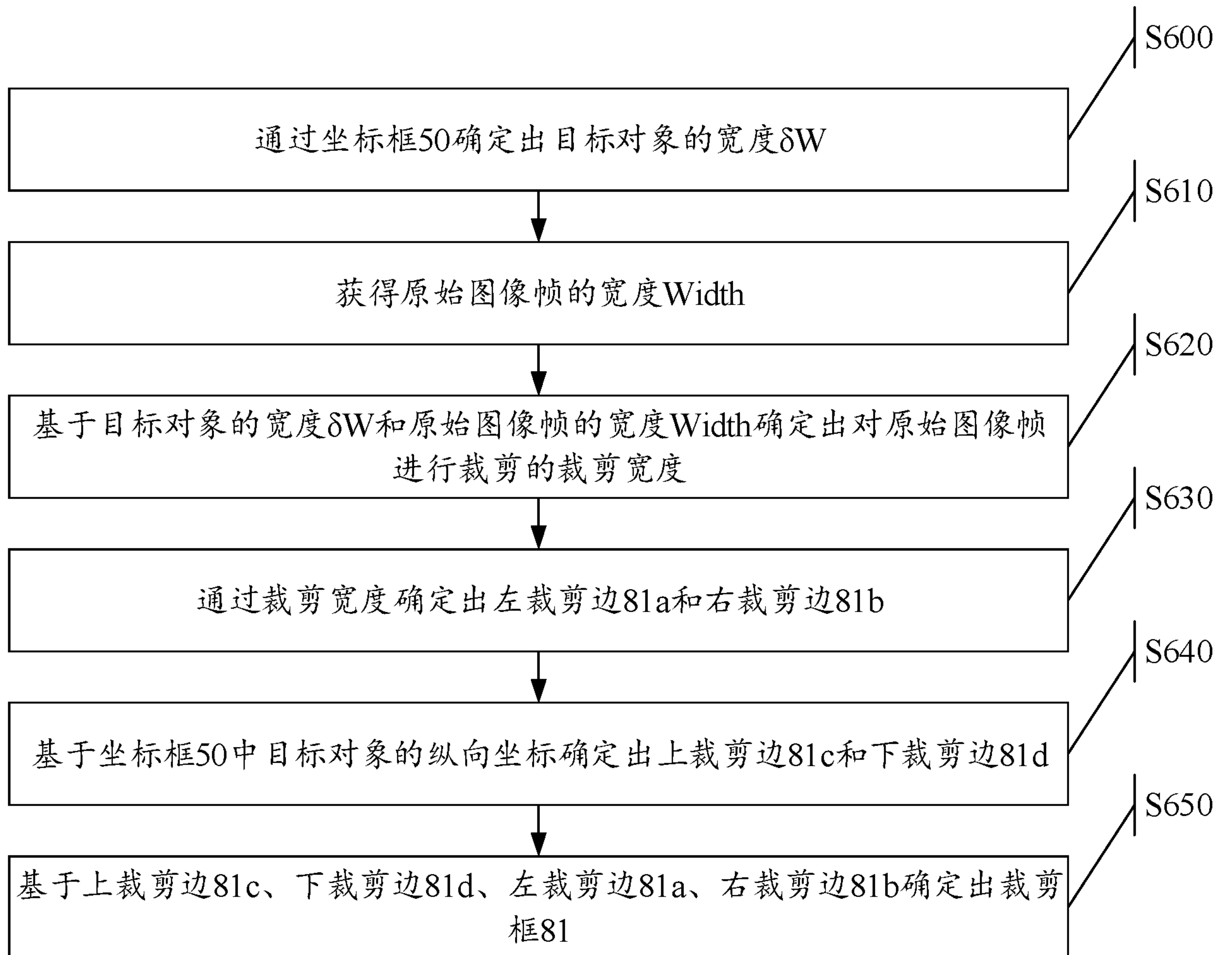


图 6

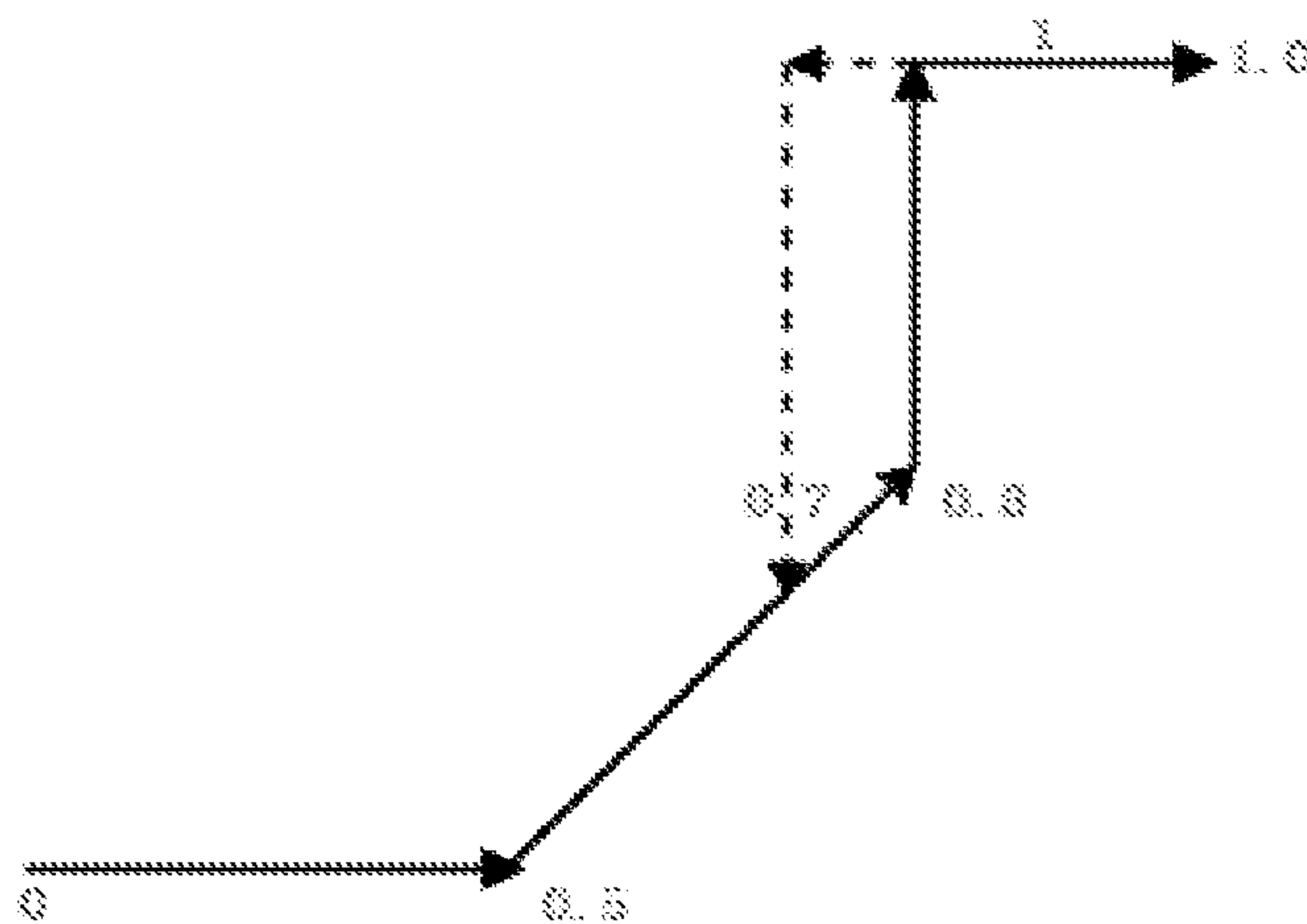


图 7

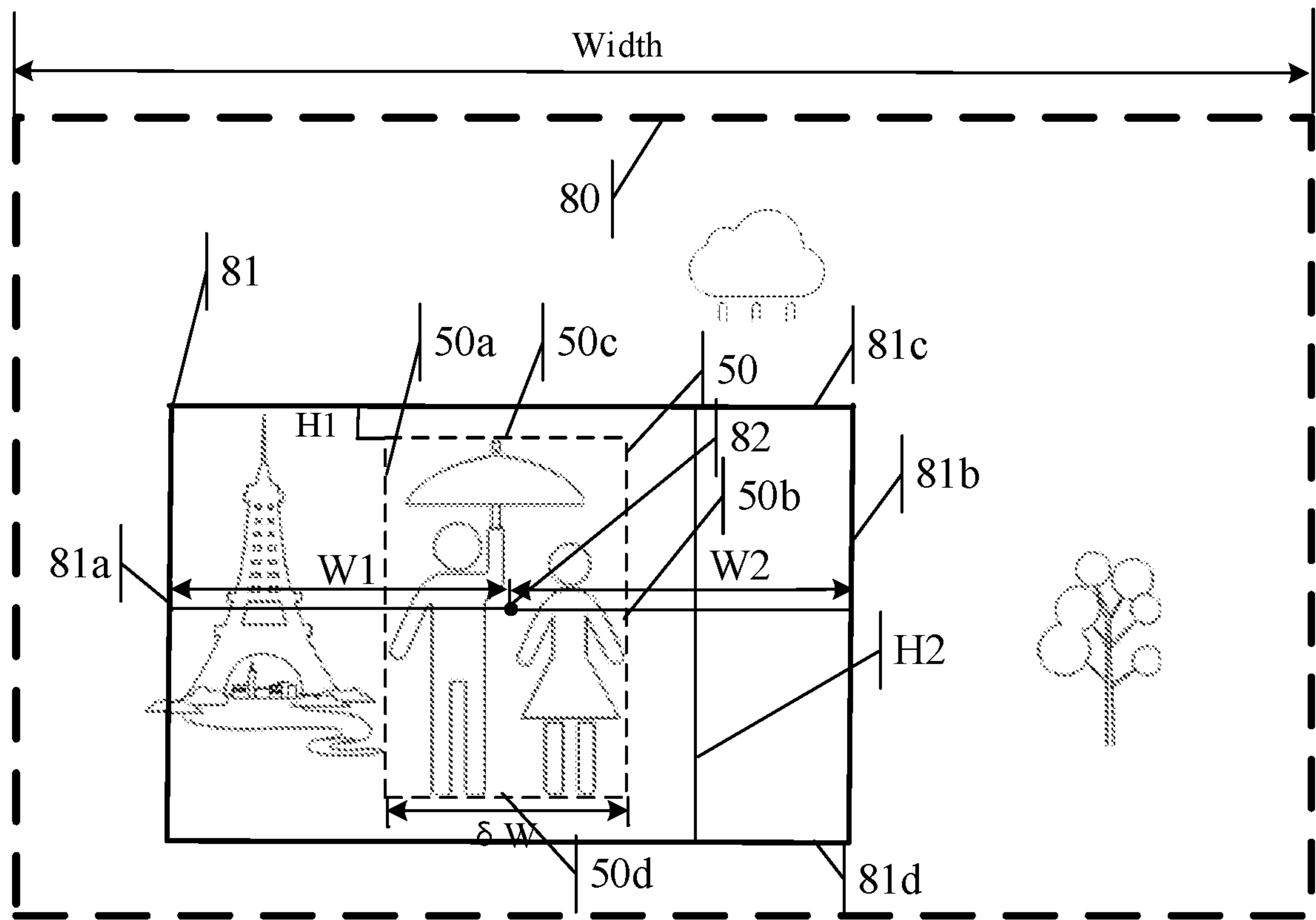


图 8A

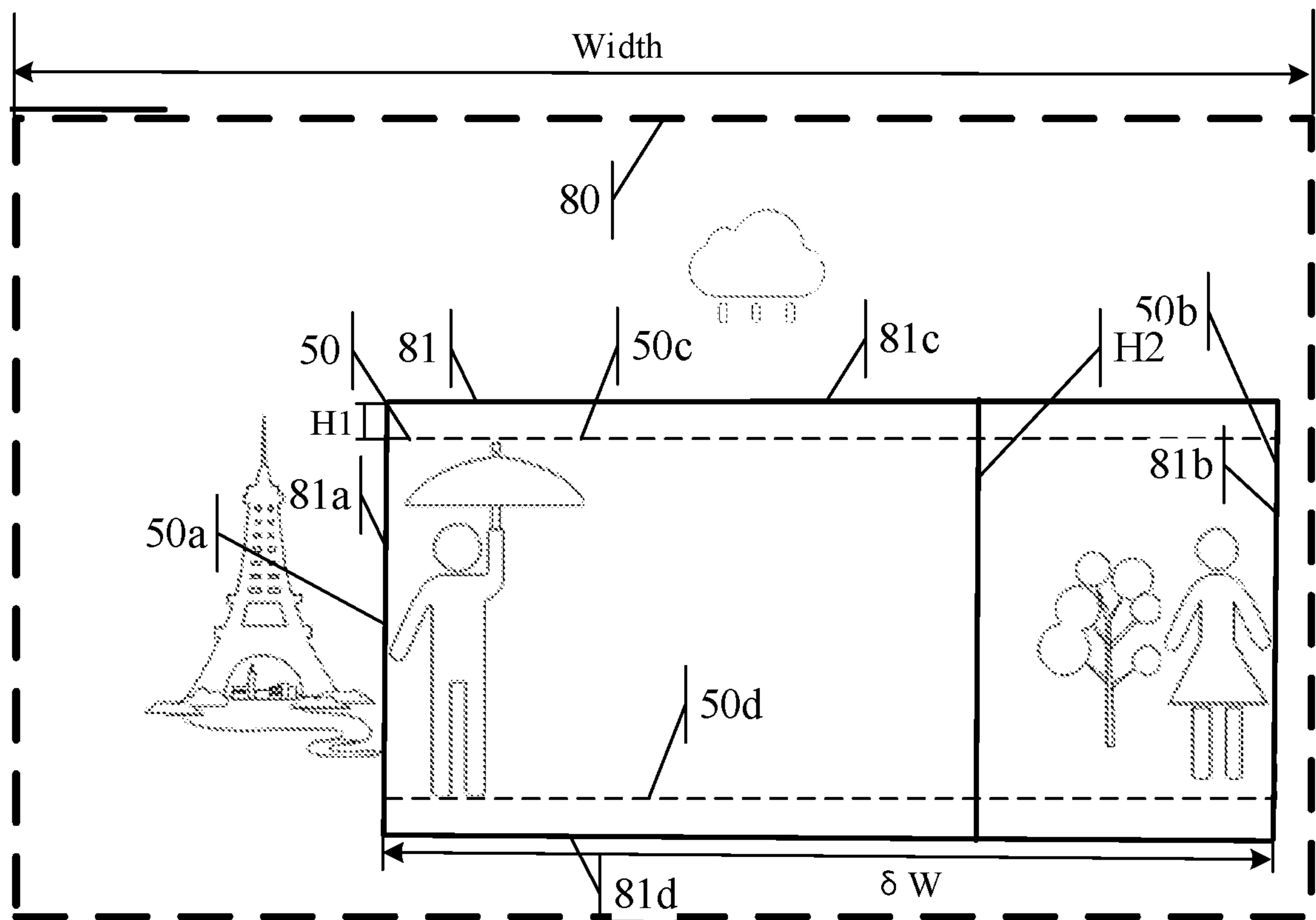


图 8B

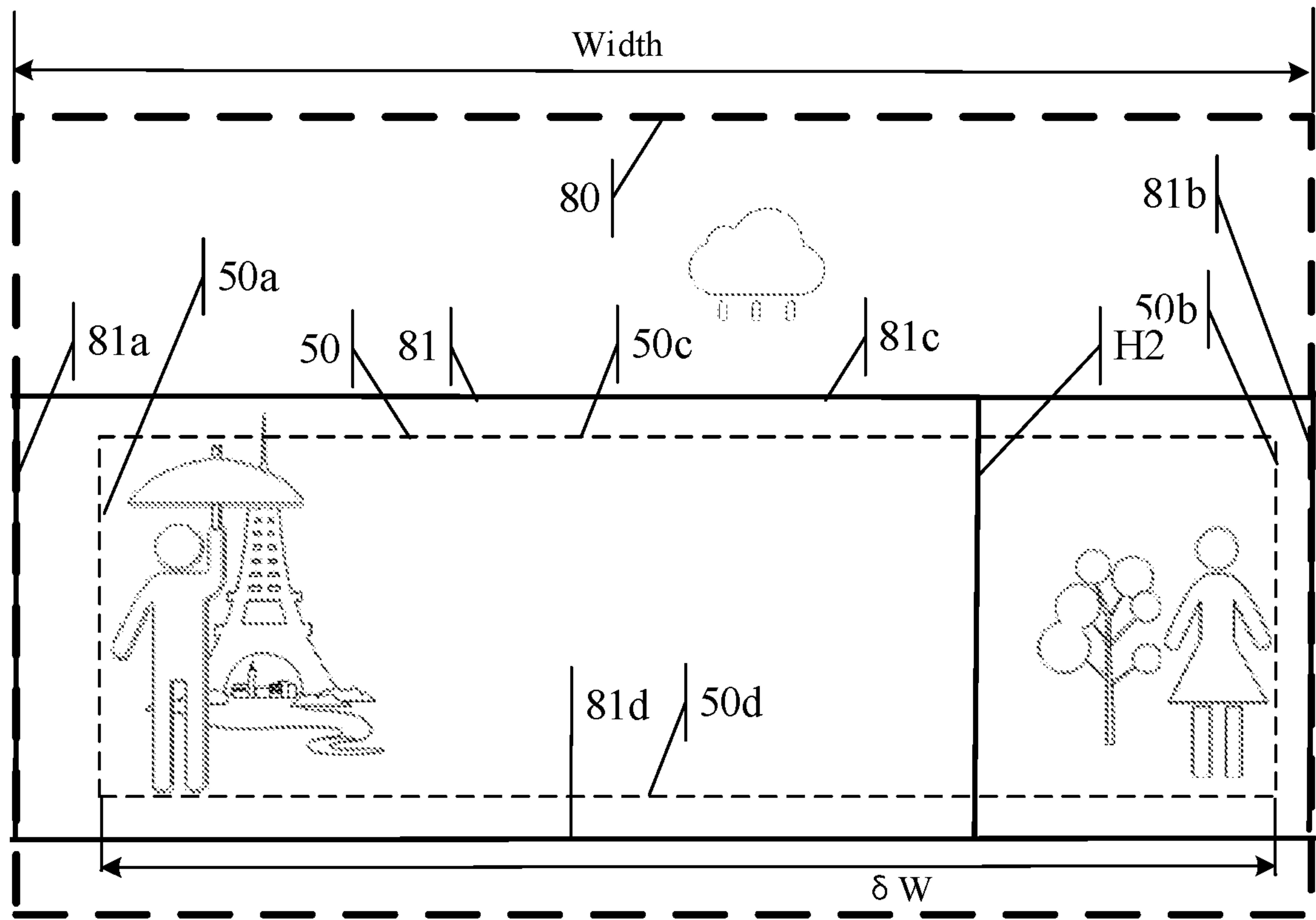


图 8C

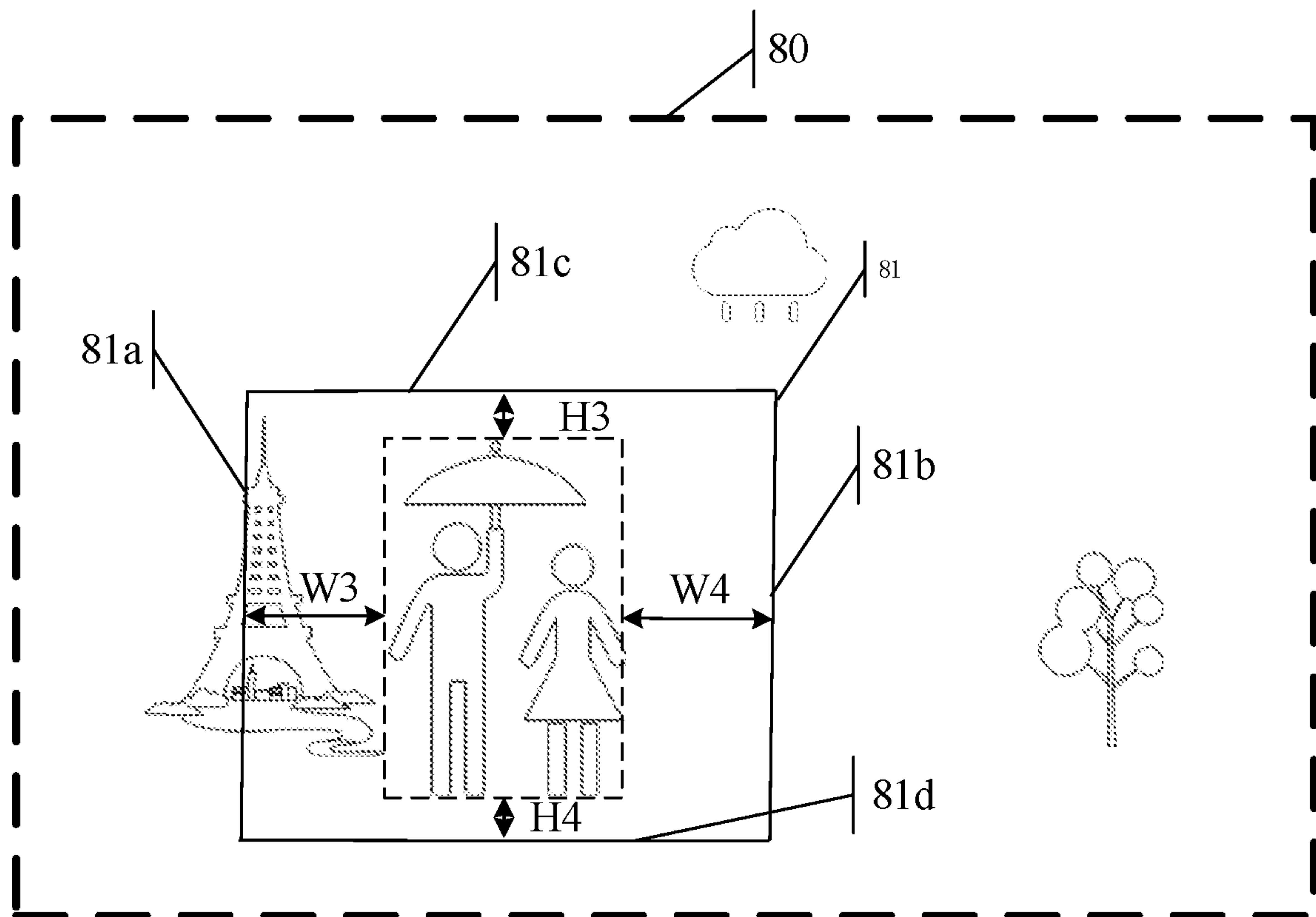


图 9

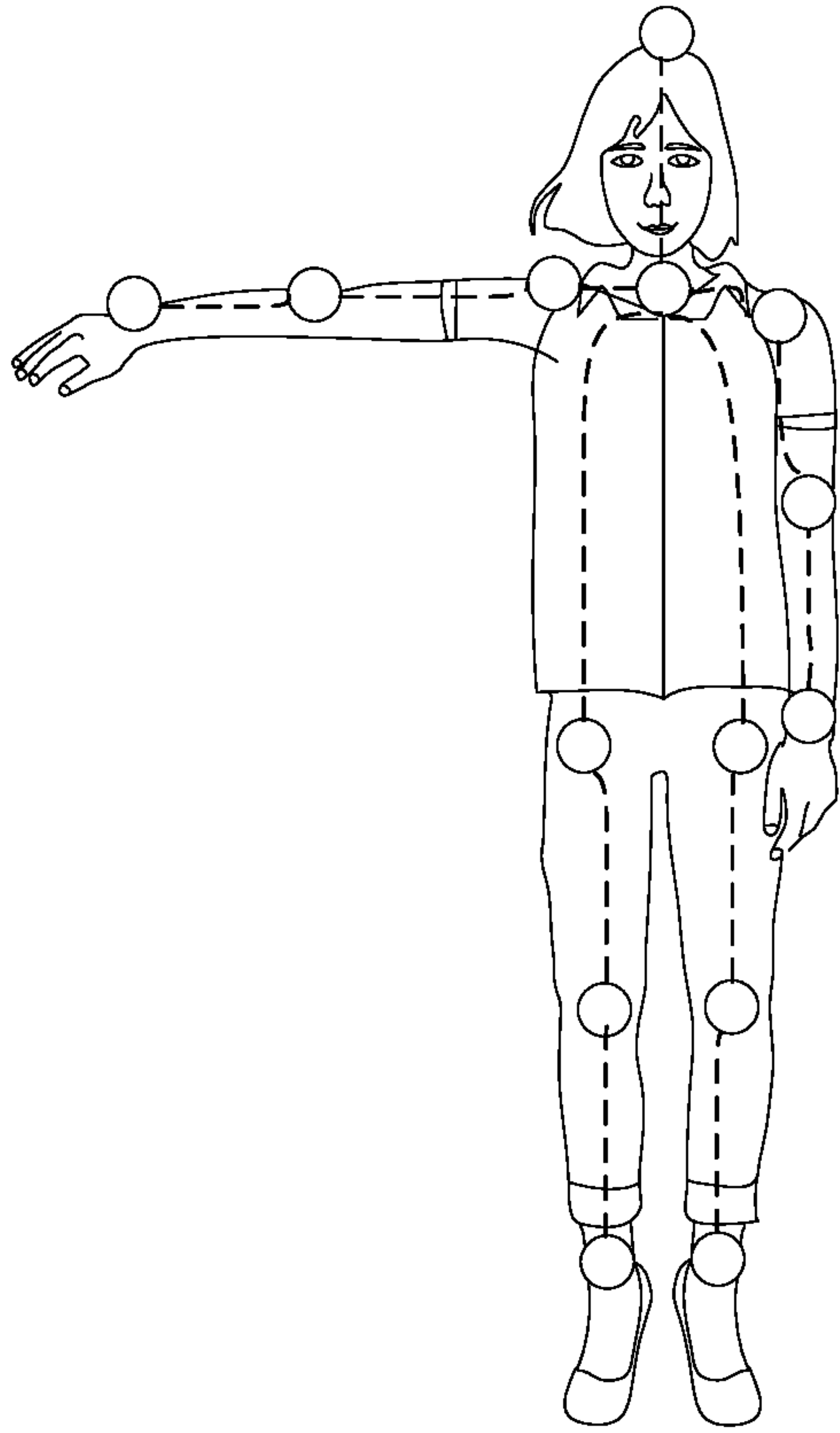


图 10

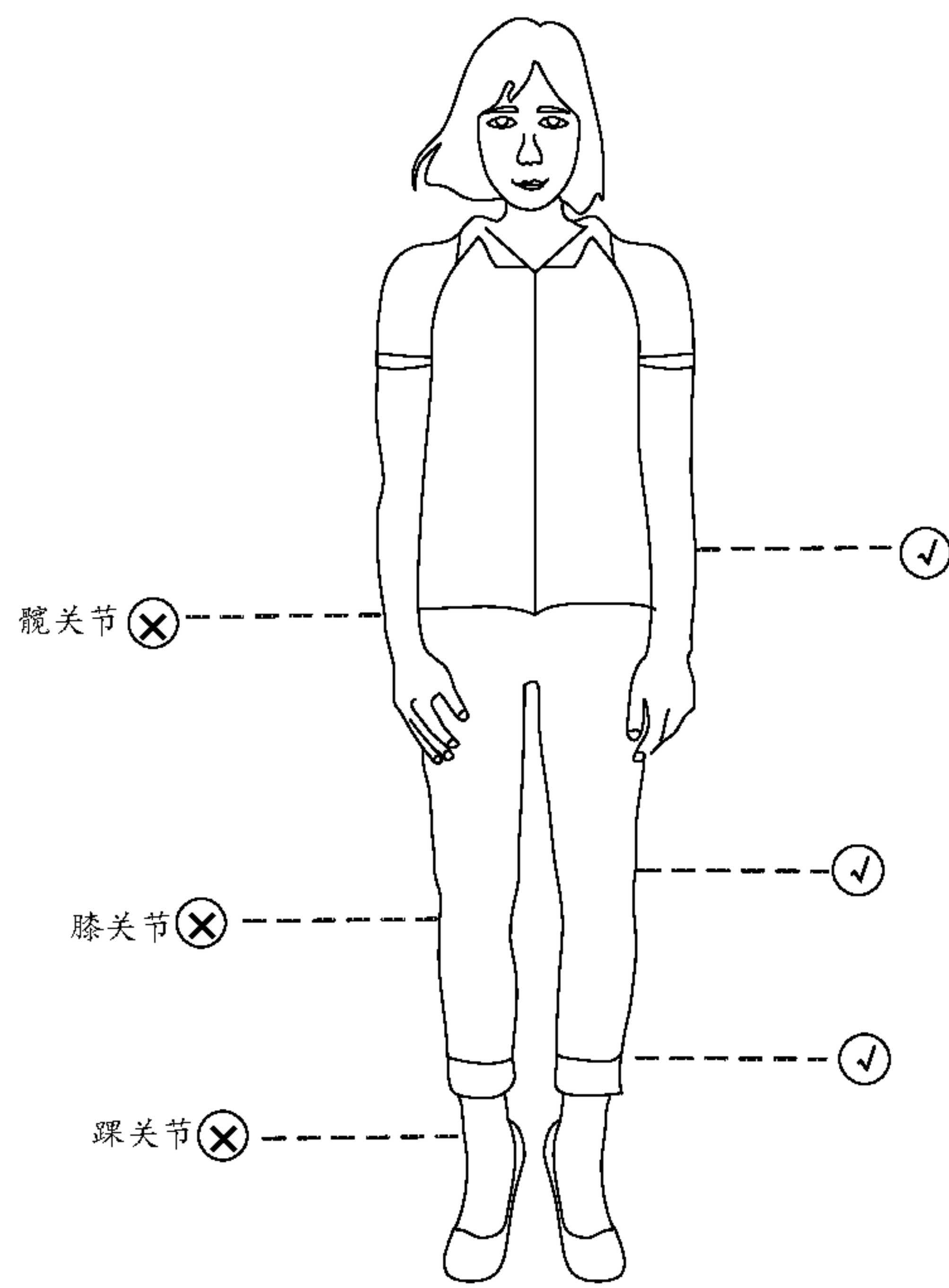


图 11

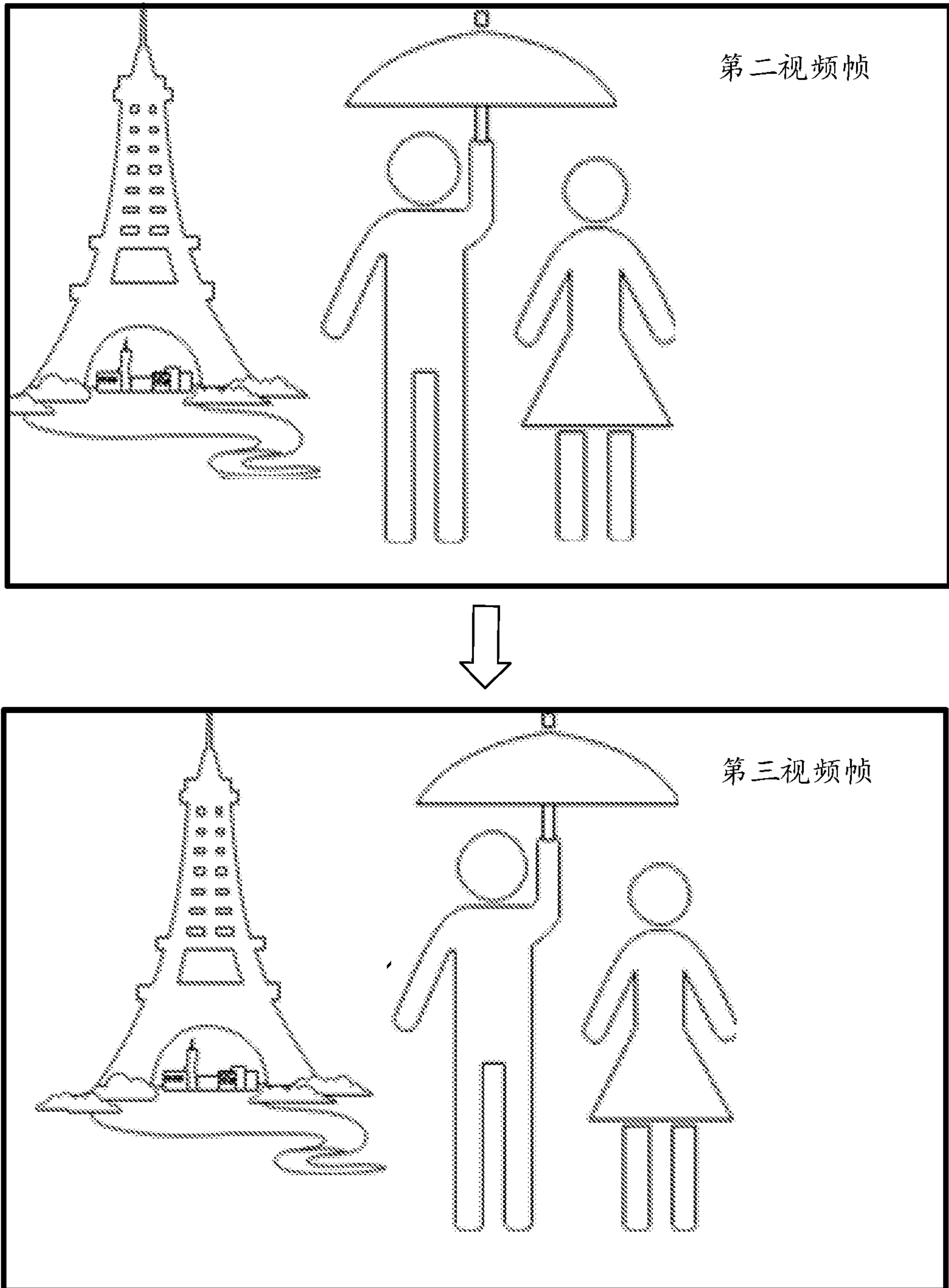


图 12

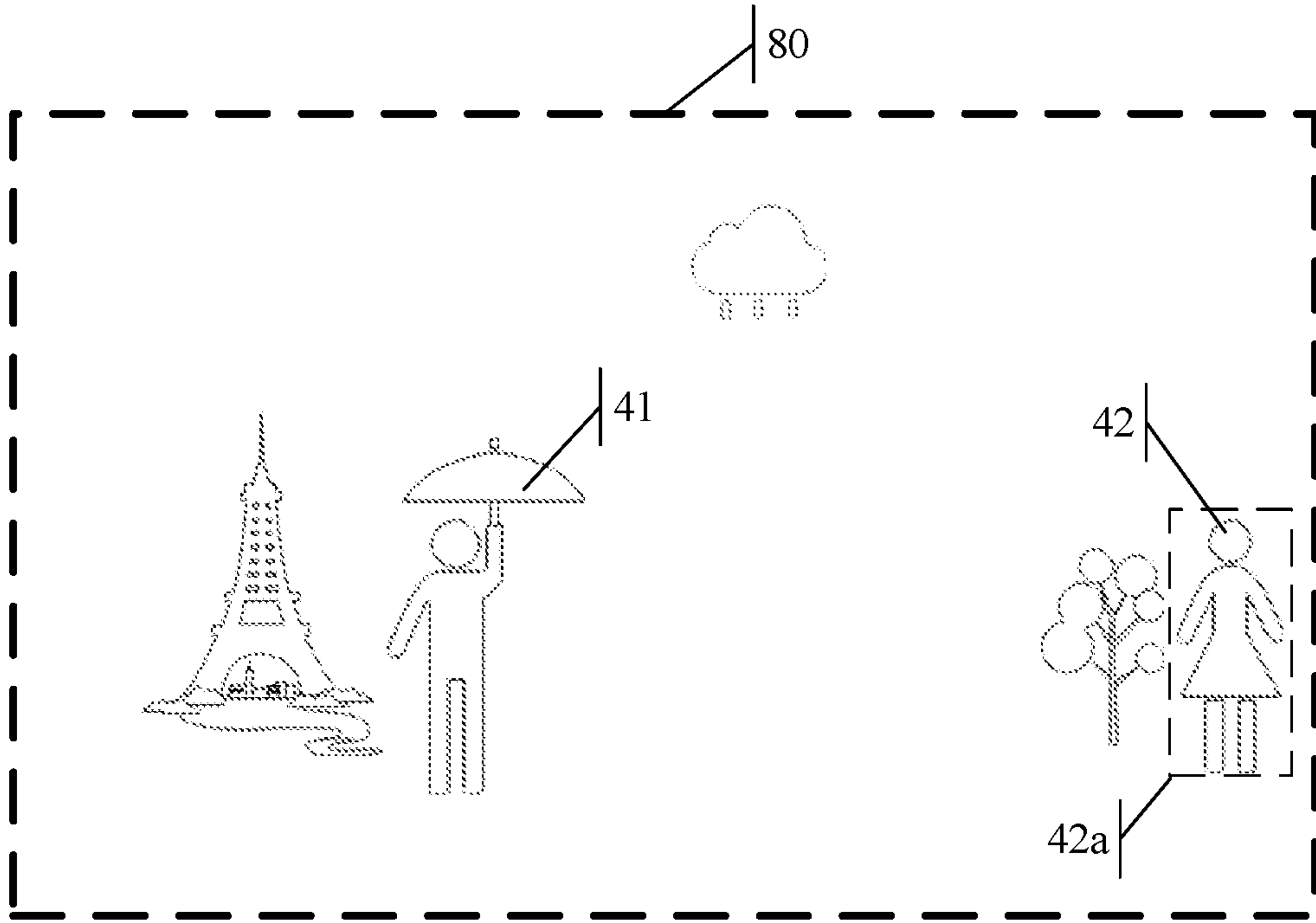


图 13A

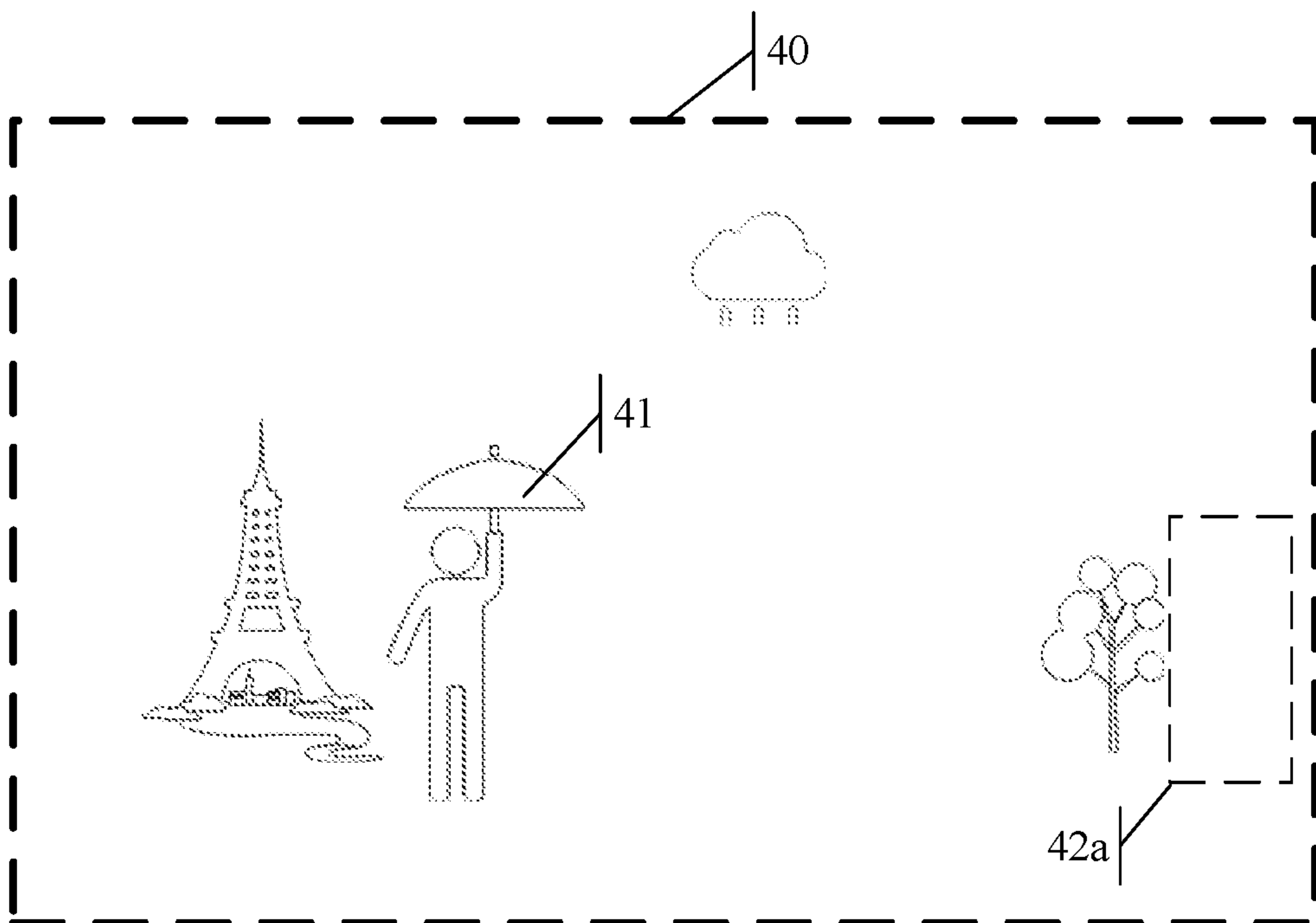


图 13B

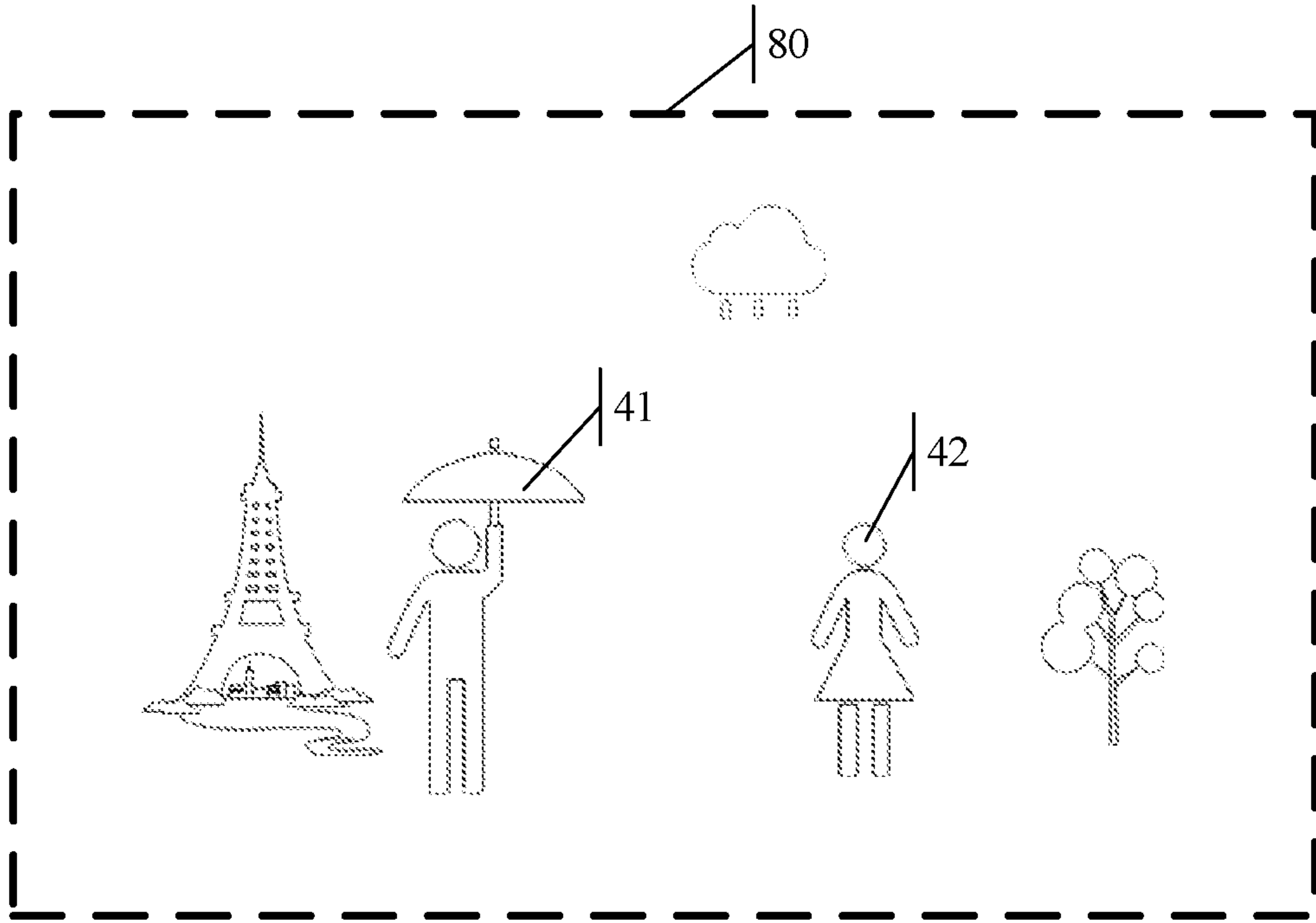


图 13C

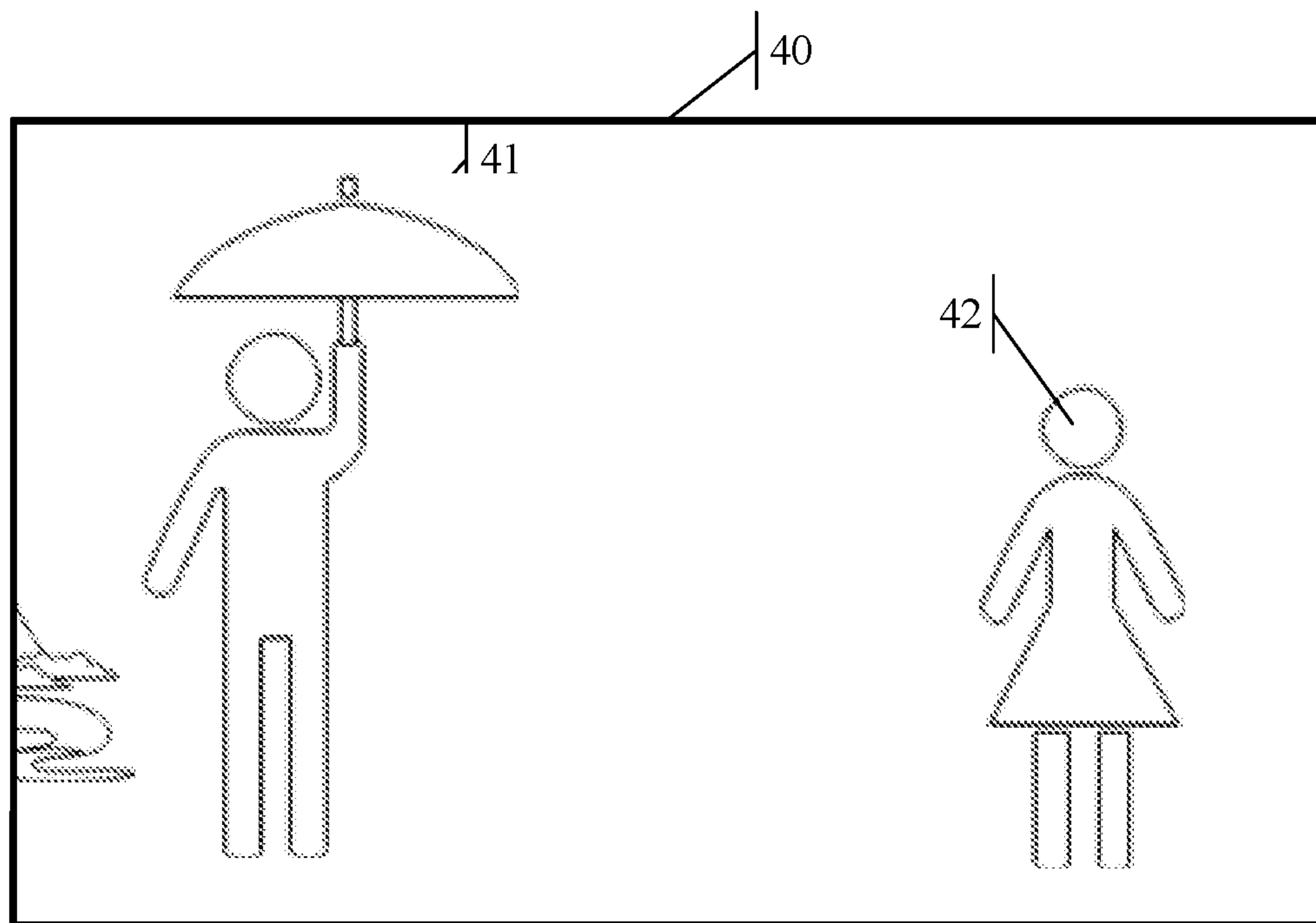


图 13D

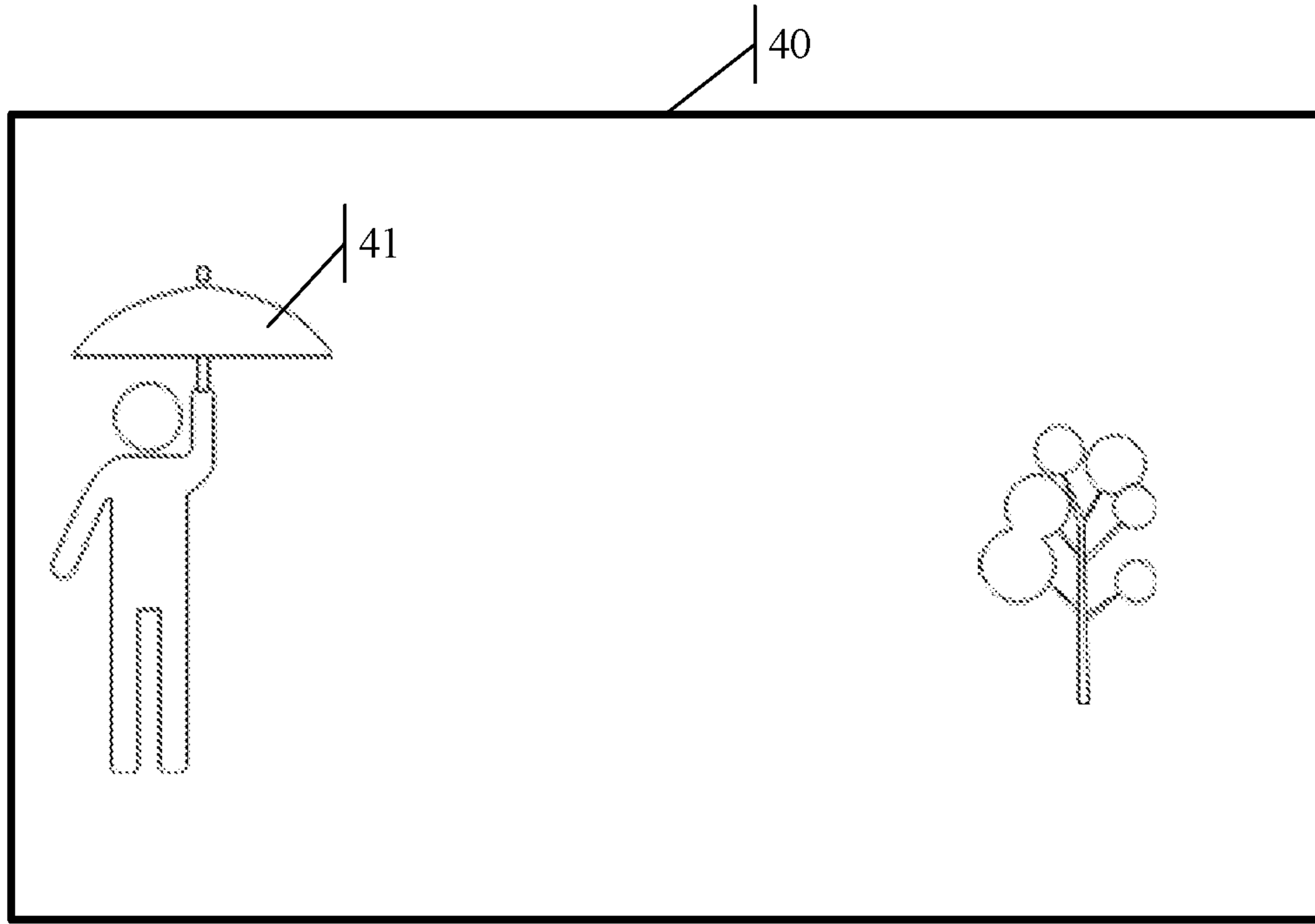


图 13E

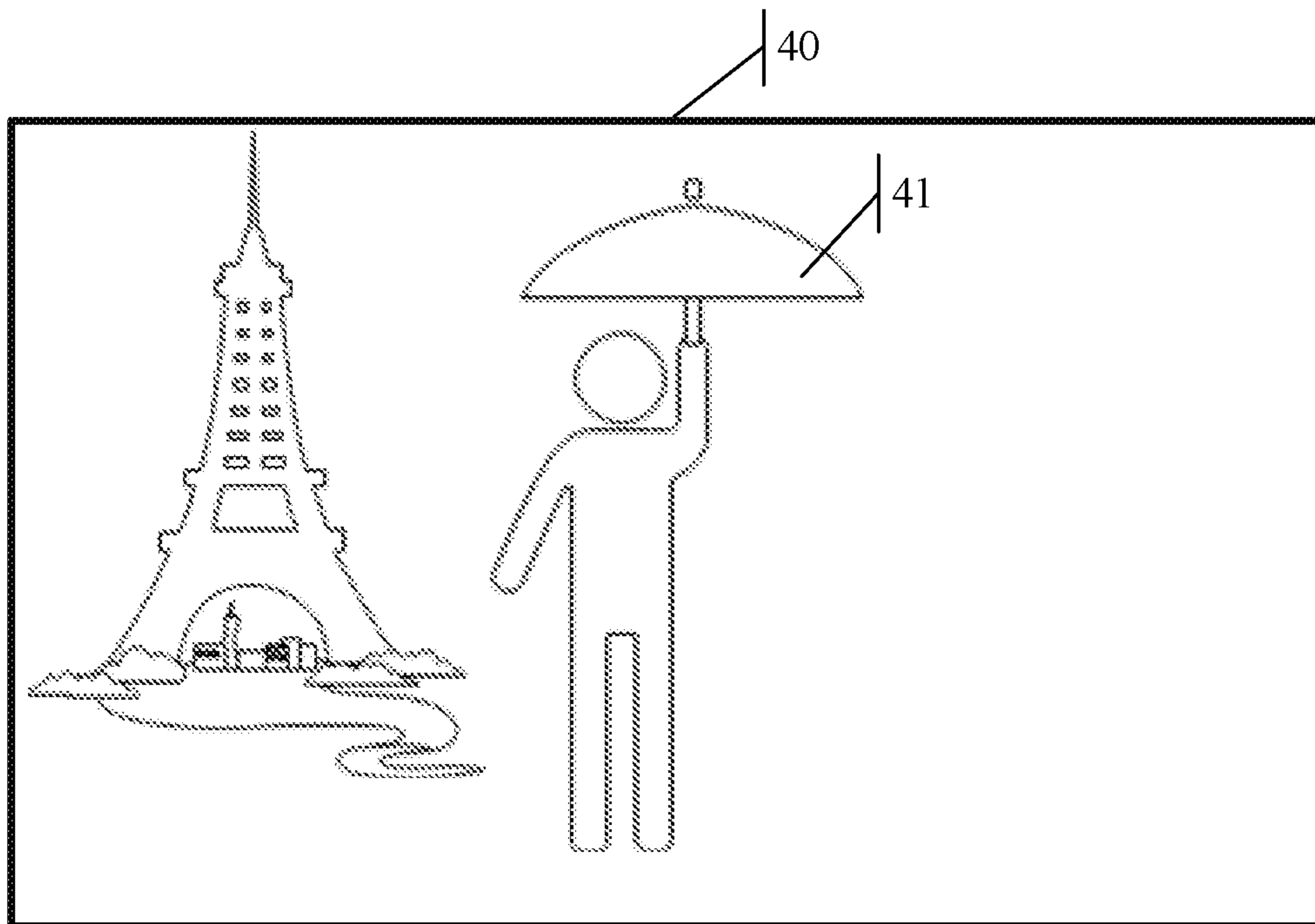


图 13F

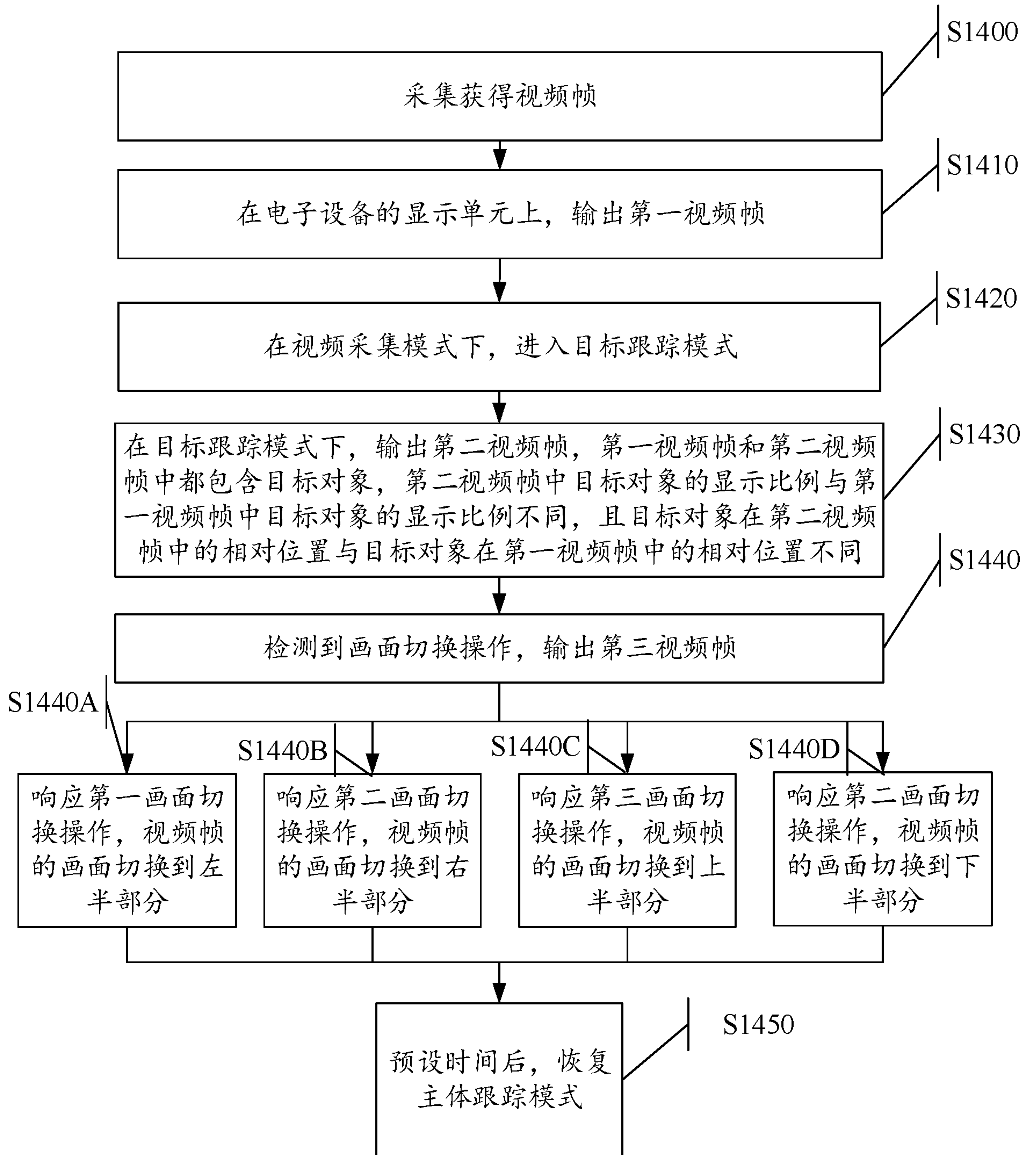


图 14

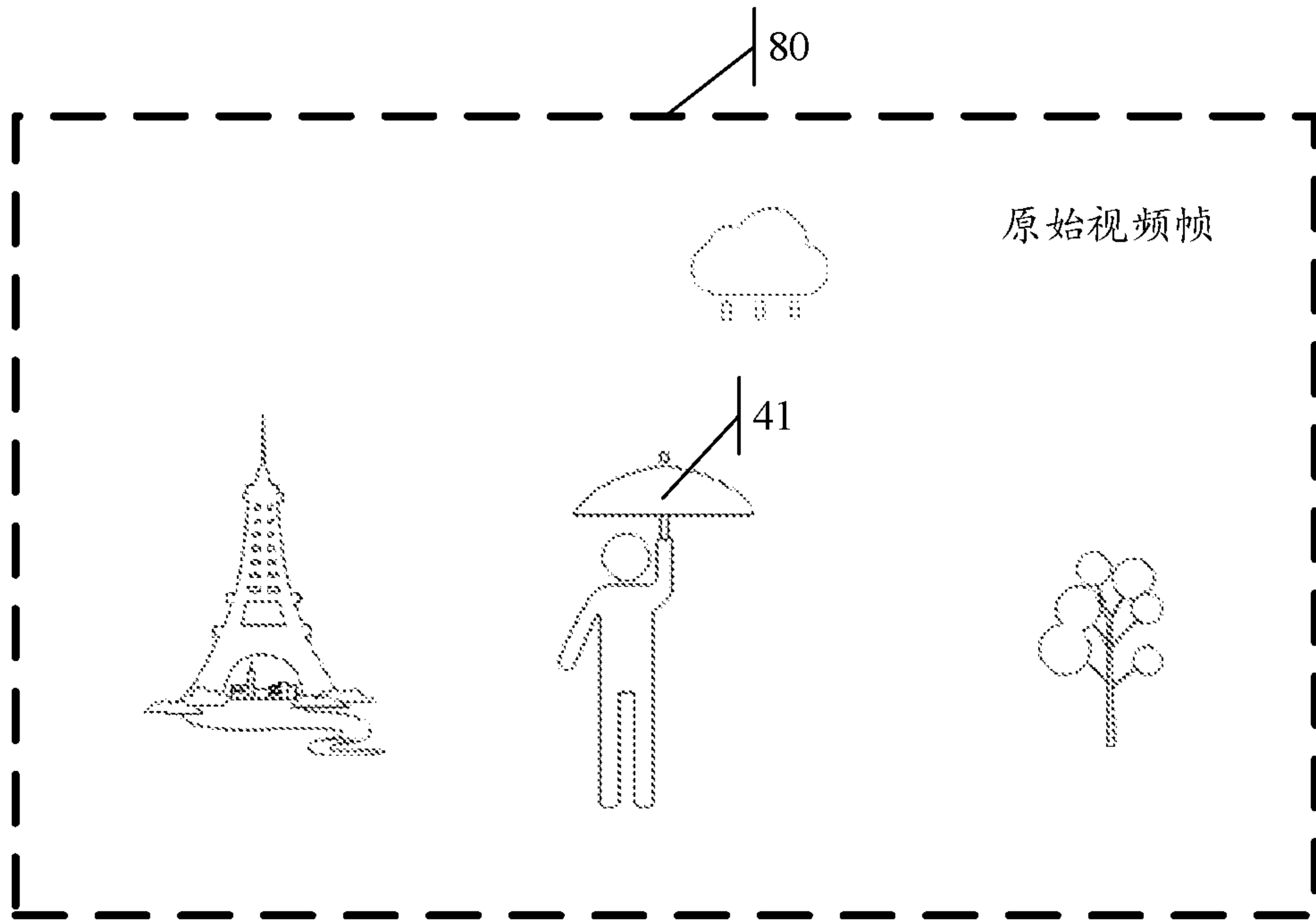


图 15A

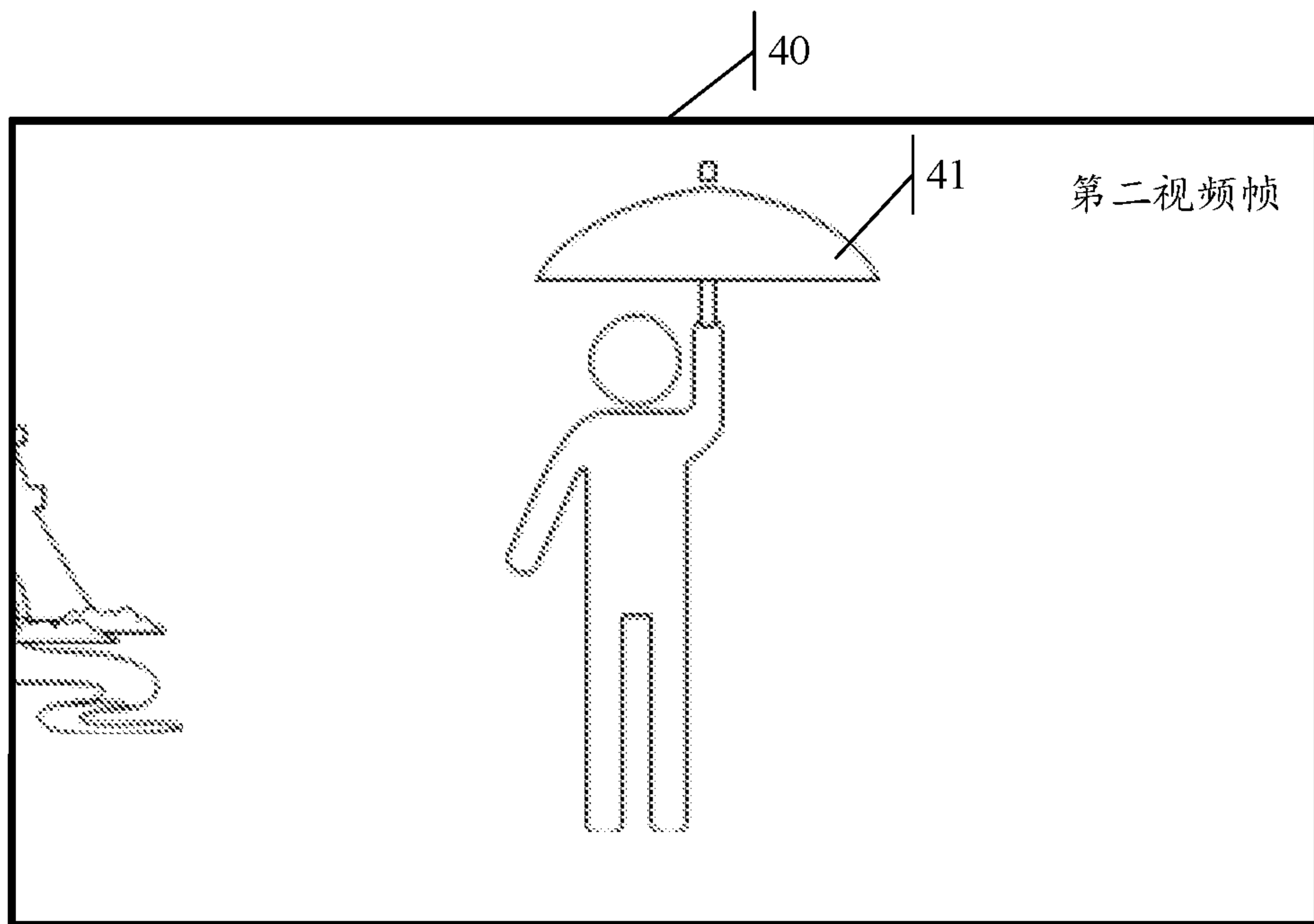


图 15B

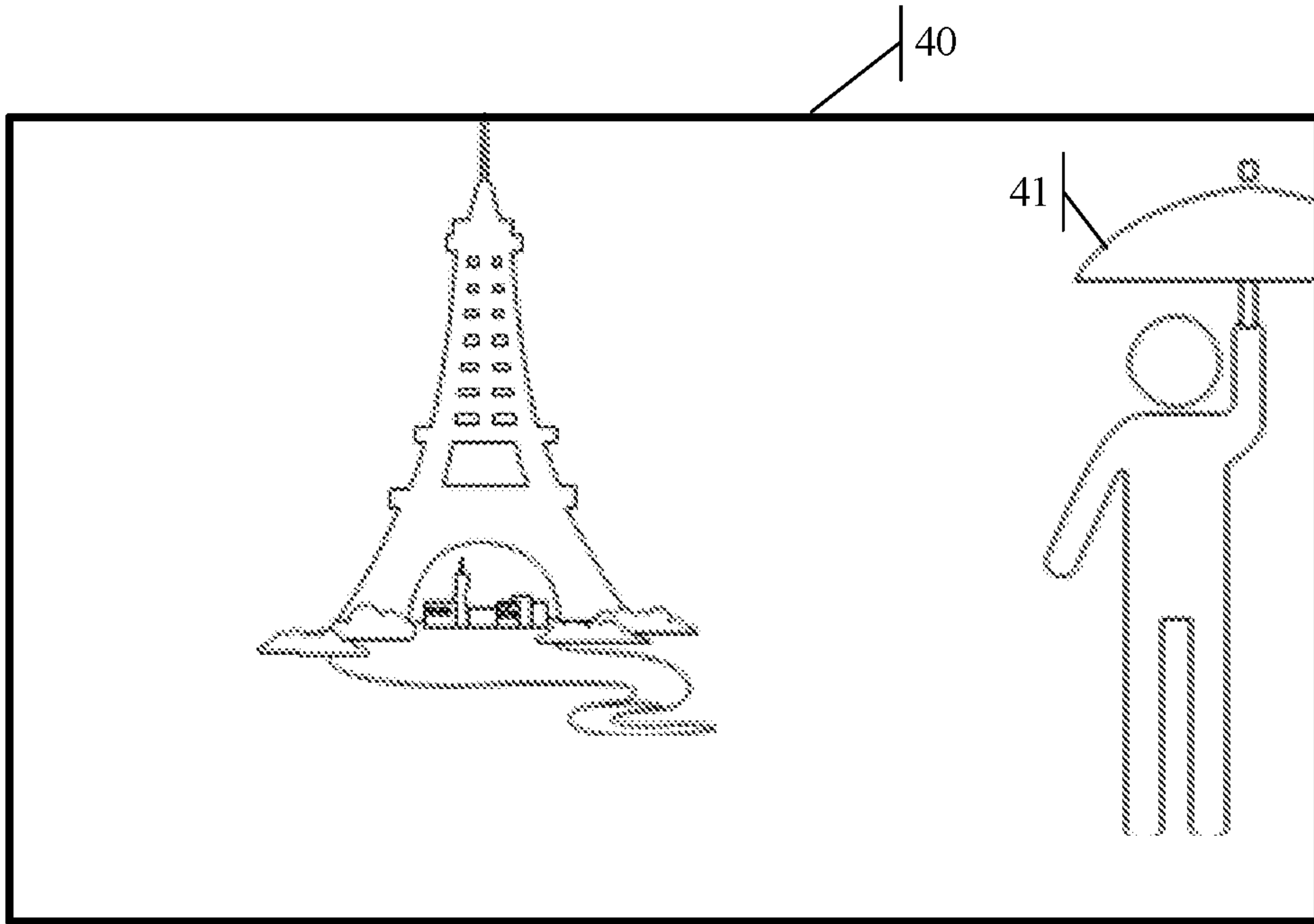


图 15C

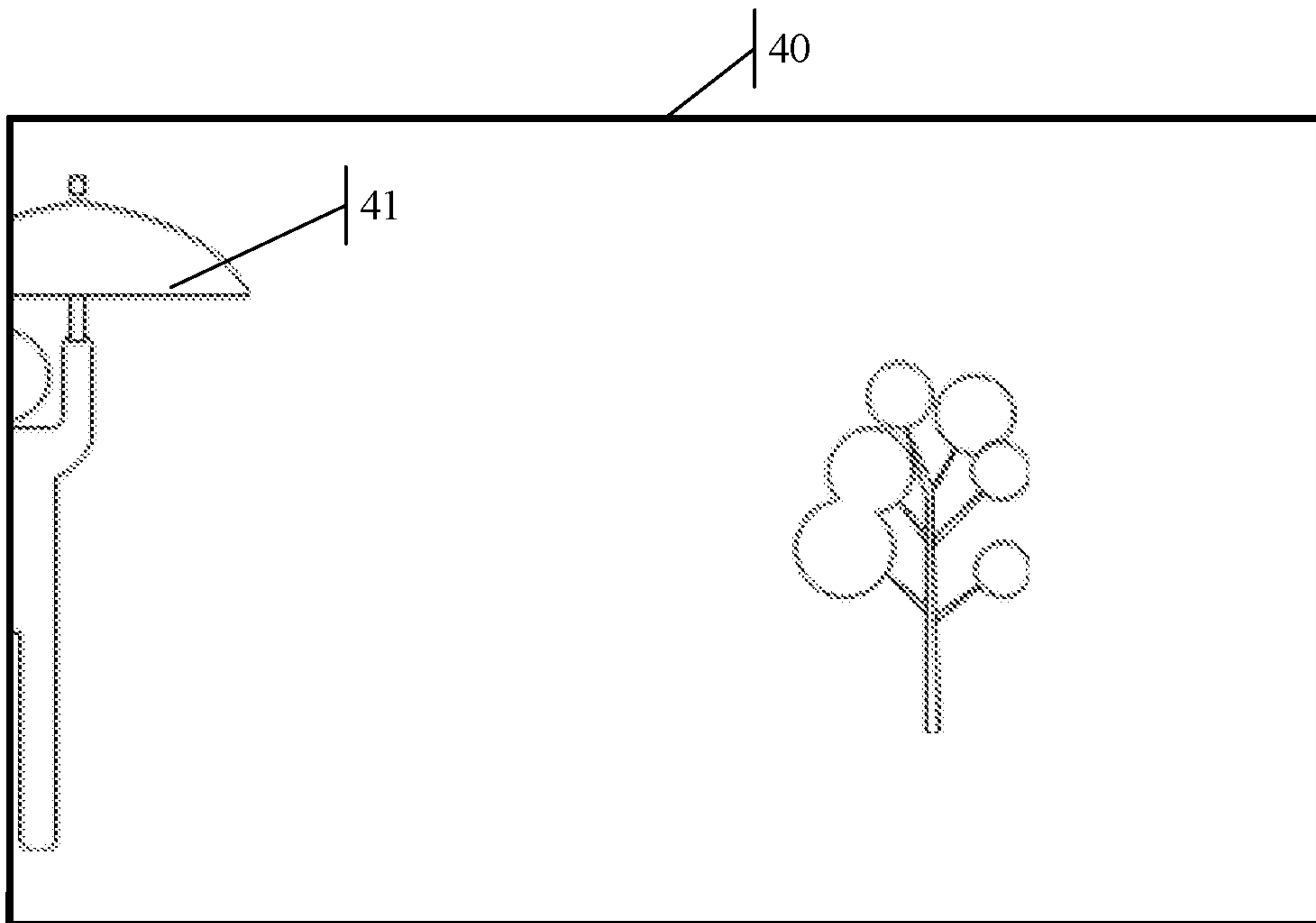


图 15D

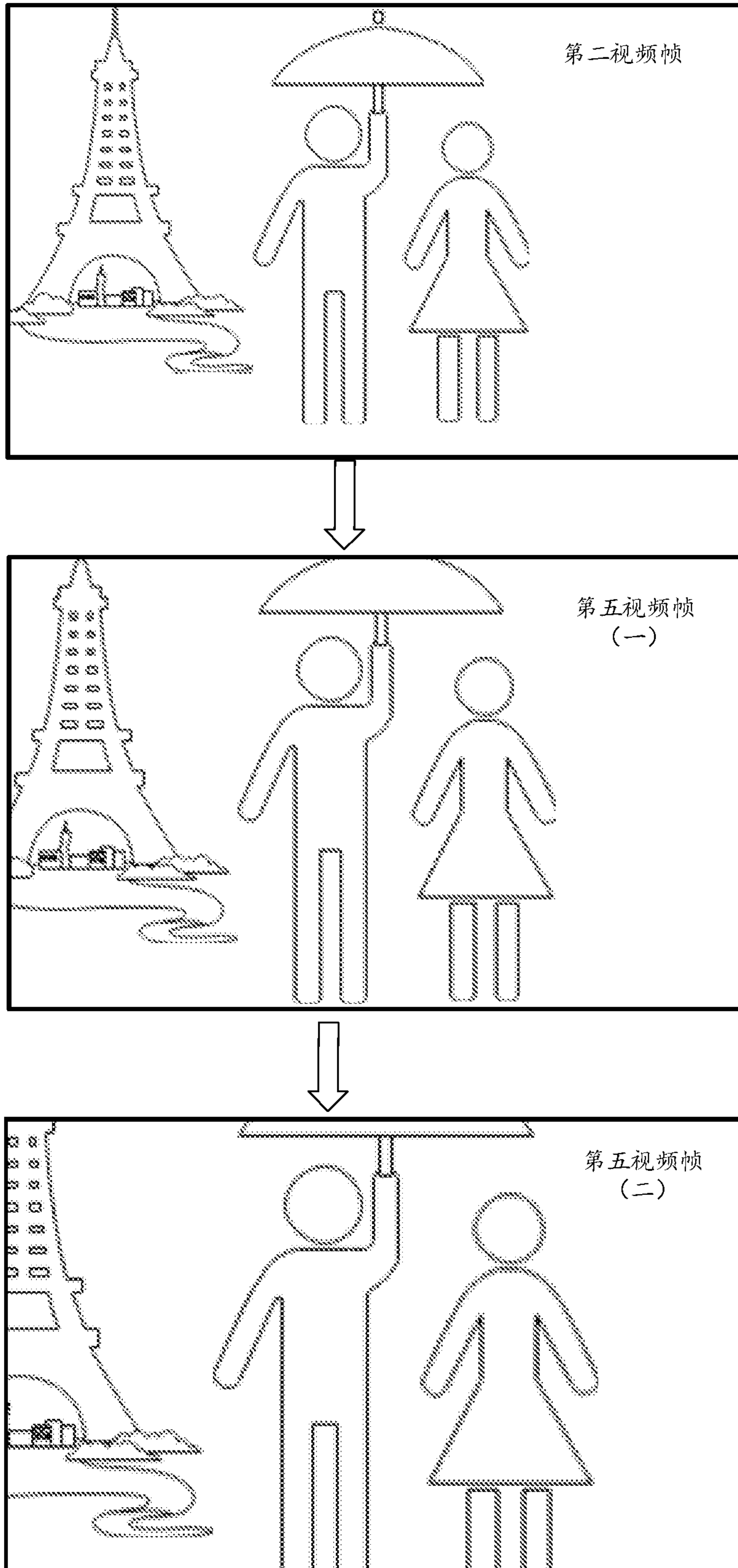


图16

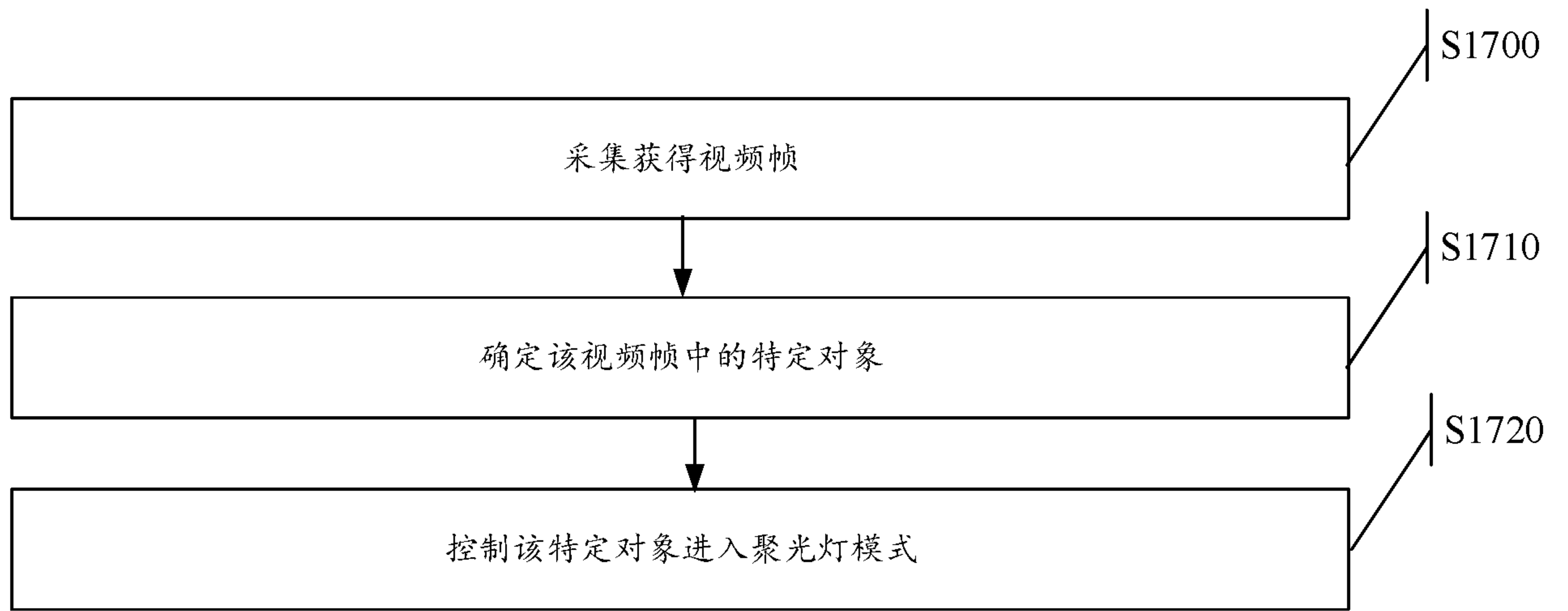


图 17

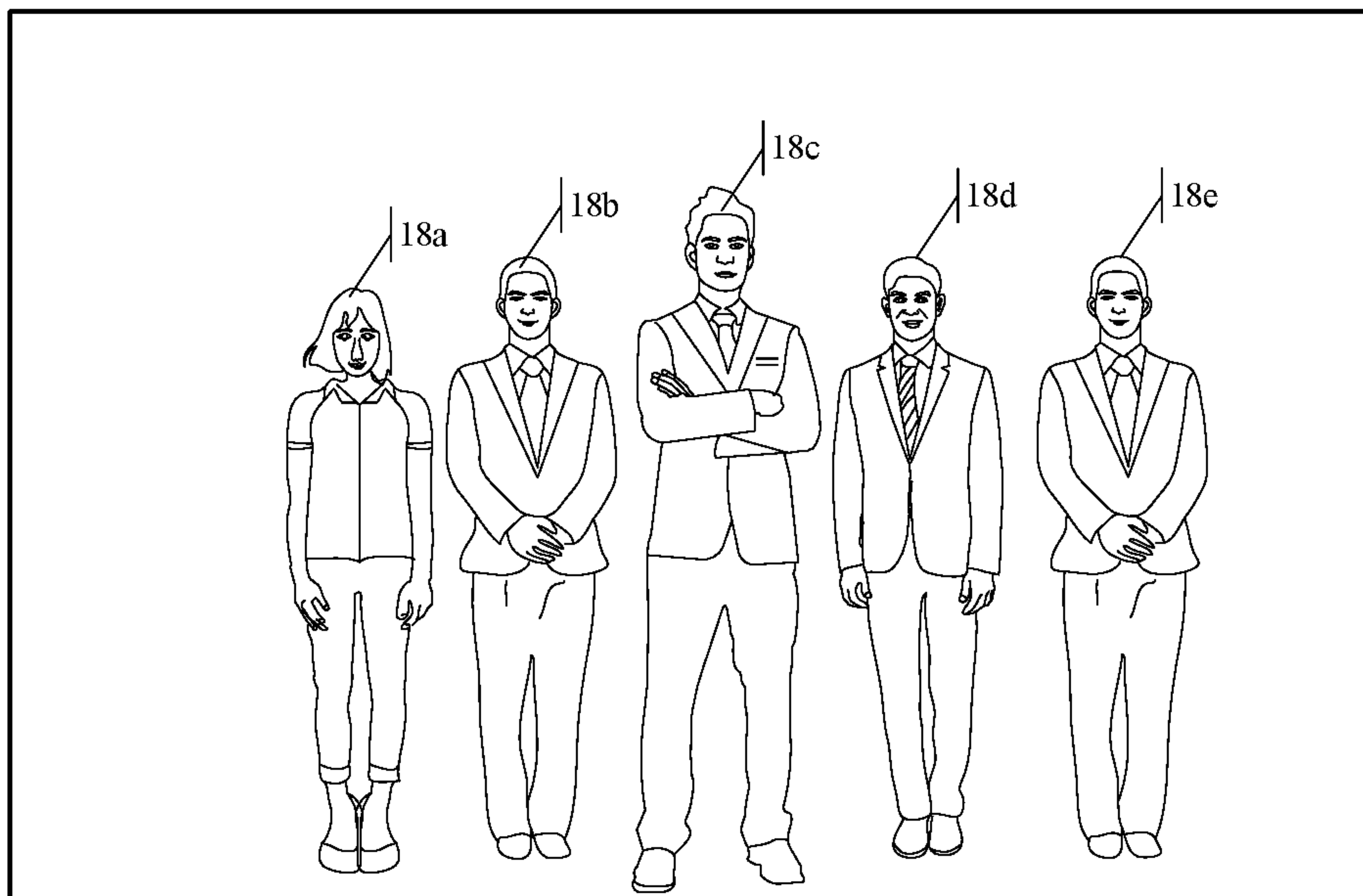


图 18A

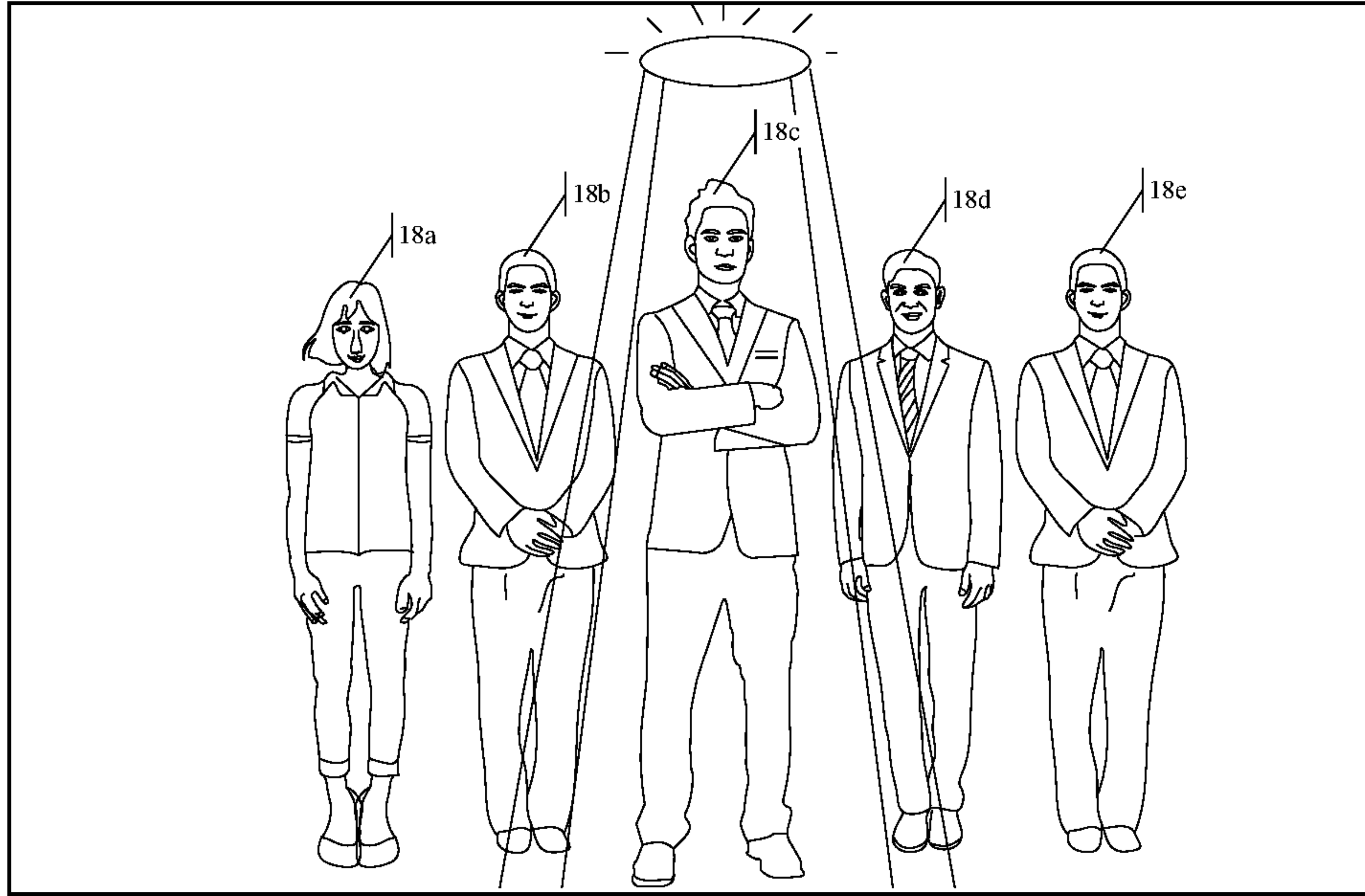


图 18B

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/137550

A. CLASSIFICATION OF SUBJECT MATTER H04N 5/232(2006.01)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) H04N 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) CNPAT, CNKI, WPI, EPODOC, IEEE: 图像, 帧, 视频, 裁剪, 截取, 主体, 对象, 识别, 框, 位置, 中心, 跟踪, 追踪, 关节, 手势, image, video, frame, clip+, object, recogni+, center, trace, gesture, joint, position		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 110347877 A (BEIJING QIYI CENTURY SCIENCE & TECHNOLOGY CO., LTD.) 18 October 2019 (2019-10-18) description, paragraphs 0069-0228, and figures 1-16	1-16
X	CN 108366303 A (NUBIA TECHNOLOGY CO., LTD.) 03 August 2018 (2018-08-03) description, paragraphs 0071-0196	1-16
X	CN 106034216 A (BEIJING ITSYNCO TECH CO., LTD.) 19 October 2016 (2016-10-19) description, paragraphs 0020-0056	1-16
A	CN 105979383 A (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) 28 September 2016 (2016-09-28) entire document	1-16
A	CN 103576848 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 12 February 2014 (2014-02-12) entire document	1-16
A	US 2011091108 A1 (CANON KABUSHIKI KAISHA) 21 April 2011 (2011-04-21) entire document	1-16
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>		
Date of the actual completion of the international search 23 February 2021		Date of mailing of the international search report 22 March 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		Authorized officer
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/137550

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110347877	A	18 October 2019	None			
CN	108366303	A	03 August 2018	None			
CN	106034216	A	19 October 2016	None			
CN	105979383	A	28 September 2016	CN	105979383	B	30 April 2019
CN	103576848	A	12 February 2014	US	2015153836	A1	04 June 2015
				RU	2613038	C2	14 March 2017
				CN	103576848	B	13 July 2016
				WO	2014023153	A1	13 February 2014
				US	10248216	B2	02 April 2019
				RU	2015107995	A	27 September 2016
US	2011091108	A1	21 April 2011	JP	5523053	B2	18 June 2014
				JP	2011086259	A	28 April 2011
				US	8666176	B2	04 March 2014

国际检索报告

国际申请号

PCT/CN2020/137550

<p>A. 主题的分类</p> <p>H04N 5/232(2006.01)i</p> <p>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</p>																							
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04N</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNPAT, CNKI, WPI, EPODOC, IEEE: 图像, 帧, 视频, 裁剪, 截取, 主体, 对象, 识别, 框, 位置, 中心, 跟踪, 追踪, 关节, 手势, image, video, frame, clip+, object, recogni+, center, trace, gesture, joint, position</p>																							
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 110347877 A (北京奇艺世纪科技有限公司) 2019年 10月 18日 (2019 - 10 - 18) 说明书第0069-0228段, 图1-16</td> <td>1-16</td> </tr> <tr> <td>X</td> <td>CN 108366303 A (努比亚技术有限公司) 2018年 8月 3日 (2018 - 08 - 03) 说明书第0071-0196段</td> <td>1-16</td> </tr> <tr> <td>X</td> <td>CN 106034216 A (北京同步科技有限公司) 2016年 10月 19日 (2016 - 10 - 19) 说明书第0020-0056段</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>CN 105979383 A (北京小米移动软件有限公司) 2016年 9月 28日 (2016 - 09 - 28) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>CN 103576848 A (腾讯科技深圳有限公司) 2014年 2月 12日 (2014 - 02 - 12) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>US 2011091108 A1 (CANON KABUSHIKI KAISHA) 2011年 4月 21日 (2011 - 04 - 21) 全文</td> <td>1-16</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 110347877 A (北京奇艺世纪科技有限公司) 2019年 10月 18日 (2019 - 10 - 18) 说明书第0069-0228段, 图1-16	1-16	X	CN 108366303 A (努比亚技术有限公司) 2018年 8月 3日 (2018 - 08 - 03) 说明书第0071-0196段	1-16	X	CN 106034216 A (北京同步科技有限公司) 2016年 10月 19日 (2016 - 10 - 19) 说明书第0020-0056段	1-16	A	CN 105979383 A (北京小米移动软件有限公司) 2016年 9月 28日 (2016 - 09 - 28) 全文	1-16	A	CN 103576848 A (腾讯科技深圳有限公司) 2014年 2月 12日 (2014 - 02 - 12) 全文	1-16	A	US 2011091108 A1 (CANON KABUSHIKI KAISHA) 2011年 4月 21日 (2011 - 04 - 21) 全文	1-16
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
X	CN 110347877 A (北京奇艺世纪科技有限公司) 2019年 10月 18日 (2019 - 10 - 18) 说明书第0069-0228段, 图1-16	1-16																					
X	CN 108366303 A (努比亚技术有限公司) 2018年 8月 3日 (2018 - 08 - 03) 说明书第0071-0196段	1-16																					
X	CN 106034216 A (北京同步科技有限公司) 2016年 10月 19日 (2016 - 10 - 19) 说明书第0020-0056段	1-16																					
A	CN 105979383 A (北京小米移动软件有限公司) 2016年 9月 28日 (2016 - 09 - 28) 全文	1-16																					
A	CN 103576848 A (腾讯科技深圳有限公司) 2014年 2月 12日 (2014 - 02 - 12) 全文	1-16																					
A	US 2011091108 A1 (CANON KABUSHIKI KAISHA) 2011年 4月 21日 (2011 - 04 - 21) 全文	1-16																					
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																							
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了解发明之理论或原理的在后文件</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																							
<p>国际检索实际完成的日期</p> <p>2021年 2月 23日</p>		<p>国际检索报告邮寄日期</p> <p>2021年 3月 22日</p>																					
<p>ISA/CN的名称和邮寄地址</p> <p>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</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>李萍</p> <p>电话号码 86-(10)-53961700</p>																					

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2020/137550

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	110347877	A	2019年 10月 18日	无			
CN	108366303	A	2018年 8月 3日	无			
CN	106034216	A	2016年 10月 19日	无			
CN	105979383	A	2016年 9月 28日	CN	105979383	B	2019年 4月 30日
CN	103576848	A	2014年 2月 12日	US	2015153836	A1	2015年 6月 4日
				RU	2613038	C2	2017年 3月 14日
				CN	103576848	B	2016年 7月 13日
				WO	2014023153	A1	2014年 2月 13日
				US	10248216	B2	2019年 4月 2日
				RU	2015107995	A	2016年 9月 27日
US	2011091108	A1	2011年 4月 21日	JP	5523053	B2	2014年 6月 18日
				JP	2011086259	A	2011年 4月 28日
				US	8666176	B2	2014年 3月 4日