

(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 1 月 13 日 (13.01.2022)



(10) 国际公布号
WO 2022/007565 A1

(51) 国际专利分类号:
G06T 19/00 (2011.01)

亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。

(21) 国际申请号: PCT/CN2021/098456

(22) 国际申请日: 2021 年 6 月 4 日 (04.06.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

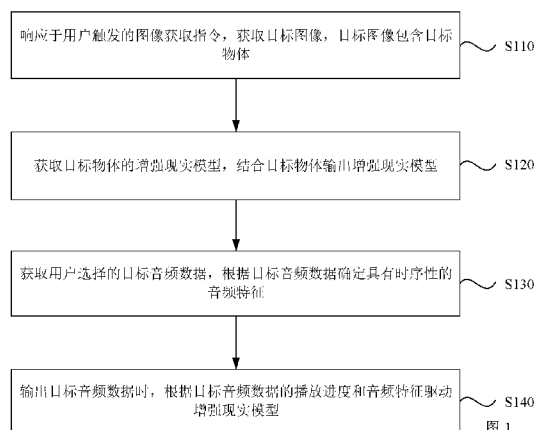
(30) 优先权:
202010662819.X 2020 年 7 月 10 日 (10.07.2020) CN

(71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市石景山区实兴大街 30 号院 3 号楼 2 层 B-0035 房间, Beijing 100041 (CN)。字节跳动有限公司 (BYTEDANCE INC.) [US/US]; 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。

(72) 发明人: 李云珠 (LI, Yunzhu); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。张璟聪 (ZHANG, Jingcong); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。宋旭晨 (SONG, Xuchen); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。杨建朝 (YANG, Jianchao); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。王国晖 (WANG, Guohui); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。陈志立 (CHEN, Zhili); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。罗琳捷 (LUO,

(54) Title: IMAGE PROCESSING METHOD AND APPARATUS FOR AUGMENTED REALITY, ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 增强现实的图像处理方法、装置、电子设备及存储介质



S110 In response to an image acquisition instruction triggered by a user, acquire a target image, wherein the target image comprises target objects
S120 Acquire augmented reality models for the target objects, and combining the target objects to output an augmented reality model
S130 Acquire target audio data selected by the user, and according to the target audio data, determine audio features that have a sequential nature
S140 When the target audio data is outputted, drive the augmented reality model according to the playback progress and audio features of the target audio data

(57) Abstract: An image processing method and apparatus for augmented reality, an electronic device and a storage medium. The method comprises: in response to an image acquisition instruction triggered by a user, acquiring a target image, wherein the target image comprises target objects; acquiring augmented reality models for the target objects, and combining the target objects to output an augmented reality model; acquiring target audio data selected by the user, and according to the target audio data, determining audio features that have a sequential nature; and when the target audio data is outputted, driving the augmented reality model according to the playback progress and audio features of the target audio data. When the augmented reality model is outputted, the audio features of the target audio data selected by the user can be combined to drive the output of an enhanced display model, so that the user can participate in the display process of the enhanced display model; and by means of selecting different target audio data, the augmented reality model is driven to display according to the audio features of the target audio data, thus improving the ease of use.

Linjie); 美国加利福尼亚州洛杉矶西杰弗逊大道12655号六层第137室, California 90066 (US)。

杨晓(YANG, Xiao); 美国加利福尼亚州洛杉矶西杰弗逊大道12655号六层第137室, California 90066 (US)。 李昊泽(LI, Haoze); 美国加利福尼亚州洛杉矶西杰弗逊大道12655号六层第137室, California 90066 (US)。 刘晶(LIU, Jing); 美国加利福尼亚州洛杉矶西杰弗逊大道12655号六层第137室, California 90066 (US)。

(74) 代理人: 上海光栅知识产权代理有限公司 (LINKER IP LLC); 中国上海市长宁区延安西路889号702室, Shanghai 201109 (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, 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))。

(57) 摘要: 一种增强现实的图像处理方法、装置、电子设备及存储介质。包括: 响应于用户触发的图像获取指令, 获取目标图像, 目标图像包含目标物体; 获取目标物体的增强现实模型, 结合目标物体输出增强现实模型; 获取用户选择的目标音频数据, 根据目标音频数据确定具有时序性的音频特征; 输出目标音频数据时, 根据目标音频数据的播放进度和音频特征驱动增强现实模型。能够在输出增强现实模型时, 结合用户选择的目标音频数据的音频特征, 驱动增强显示模型的输出, 使得用户能够参与到增强显示模型的展示过程中, 通过选择不同的目标音频数据, 驱动增强现实模型根据目标音频数据的音频特征进行展示, 提高易用性。

增强现实的图像处理方法、装置、电子设备及存储介质

相关申请交叉引用

本申请要求于 2020 年 7 月 10 日提交中国专利局、申请号为 202010662819.X、发明名称为“增强现实的图像处理方法、装置、电子设备及存储介质”的中国专利申请的优先权，其全部内容通过引用并入本文。

技术领域

本发明实施例涉及虚拟现实技术，尤其涉及一种增强现实的图像处理方法、装置、电子设备及存储介质。

背景技术

增强现实(Augmented Reality, 简称 AR)是一种现实信息与虚拟信息相叠加的技术。首先，计算机系统处理现实信息，根据现实信息生成匹配且包含虚拟物体、声音或文字等形式的虚拟信息；然后，将虚拟信息叠加到显示现实信息的人机交互界面中，从而增强用户对现实世界的感知。

目前，增强现实模型只能按照预先设置的固定方式展示，展示方式单一，缺乏互动，易用性差。

发明内容

本发明提供一种增强现实的图像处理方法、装置、电子设备及存储介质，以实现提高增强现实模型的互动性，提高易用性。

第一方面，本发明实施例提供了一种增强现实的图像处理方法，包括：

响应于用户触发的图像获取指令，获取目标图像，目标图像包含目标物体；

获取目标物体的增强现实模型，结合目标物体输出增强现实模型；

获取用户选择的目标音频数据，根据目标音频数据确定具有时序性的音频特征；

输出目标音频数据时，根据目标音频数据的播放进度和音频特征驱动增强现实模型。

第二方面，本发明实施例还提供了一种增强现实的图像处理装置，包括：

目标图像获取模块，用于响应于用户触发的图像获取指令，获取目标图像，目标图像包含目标物体；

增强现实模型获取模块，用于获取目标物体的增强现实模型，结合目标物体输出增强现实模型；

目标音频获取模块，用于获取用户选择的目标音频数据；

音频特征确定模块，用于根据目标音频数据确定具有时序性的音频特征；

5 输出模块，用于输出目标音频数据时，根据目标音频数据的播放进度和音频特征驱动增强现实模型。

第三方面，本发明实施例还提供了一种电子设备，电子设备包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序，

10 当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现如本公开实施例所示的增强现实的图像处理方法。

第四方面，本发明实施例还提供了一种包含计算机可执行指令的存储介质，计算机可执行指令在由计算机处理器执行时用于执行如本公开实施例所示的增强现实的图像处理方法。

15 第五方面，本发明实施例还提供了一种计算机程序产品，所述计算机程序产品包括：计算机程序，所述计算机程序存储在可读存储介质中，电子设备的一个或多个处理器可以从所述可读存储介质读取所述计算机程序，所述一个或多个处理器执行所述计算机程序，使得所述电子设备执行如本公开实施例所示的增强现实的图像处理方法。

第六方面，本公开实施例还提供了一种计算机程序，所述计算机程序存储在可读存储
20 介质中，设备的一个或多个处理器可以从所述可读存储介质中读取所述计算机程序，所述一个或多个处理器执行所述计算机程序，使得所述电子设备执行如本公开实施例所示的增强现实的图像处理方法。

本公开实施例公开的增强现实的图像处理方案，能够响应于用户触发的图像获取指令，获取目标图像，目标图像包含目标物体；获取目标物体的增强现实模型，结合目标物体输出增强现实模型；获取用户选择的目标音频数据，根据目标音频数据确定具有时序性的音频特征；输出目标音频数据时，根据目标音频数据的播放进度和音频特征驱动增强现实模型。相对于目前增强现实模型缺乏互动性，易用性差。本公开实施例公开的增强现实的图像处理方案，能够在输出增强现实模型时，结合用户选择的目标音频数据的音频特征，驱动增强显示模型的输出，使得用户能够参与到增强显示模型的展示过程中，通过选择不同的目标音频数据，驱动增强现实模型根据目标音频数据的音频特征进行展示，提高易用性。
25
30

附图说明

图 1 是本发明实施例一中的增强现实的图像处理方法的流程图；

图 2 是本发明实施例二中的增强现实的图像处理装置的结构示意图；

图 3 是本发明实施例三中的电子设备结构示意图。

5

具体实施方式

下面结合附图和实施例对本发明作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释本发明，而非对本发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与本发明相关的部分而非全部结构。

10 实施例一

图 1 为本发明实施例一提供的增强现实的图像处理方法的流程图，本实施例可适用于展示增强现实模型的情况，该方法可以由实现增强现实的电子设备来执行，该电子设备可以为智能手机、平板电脑等，具体包括如下步骤：

步骤 110、响应于用户触发的图像获取指令，获取目标图像，目标图像包含目标物体。

15 用户在预设应用中可以出发图像获取指令，当用户出发图像获取指令时，电子设备通过摄像头获取目标图像。用户可以在使用增强现实模型的意向的前提下，开启预设应用。使用电子设备拍摄目标物体的画面。目标物体可以为地标建筑等具备增强现实模型的物体。地标建筑物可以为某个区域中具有独特设计风格的建筑物。

预设应用可以为电子设备的拍照应用，或者具有增强现实功能的应用。电子设备的摄像头获取目标图像，电子设备将获取到的目标图像显示在预览页面中。预览页面可以为用户提供摄像头获取到的实时图像。

步骤 120、获取目标物体的增强现实模型，结合目标物体输出增强现实模型。

若识别到当前图像中的拍摄物存在相应的增强现实模型，则将增强现实模型映射到目标图像中的目标物体上，实现结合目标物体输出增强现实模型。

25 进一步的，可以为用户提供人工调整增强现实模型大小的接口，若机器结合效果不佳，如增强现实模型与目标物体无法准确结合，则可以通过该接口调整增强现实模型的大小，进而实现由用户调节增强现实模型大小，提高易用性。

进一步的，在一种实现方式中，当目标物体为地标建筑物时，获取目标物体的增强现实模型，可以通过下述方式实施：

30 当监测到图标图像中包含建筑物时，根据电子设备当前位置信息和拍摄角度确定目标

物体标识，将目标物体标识表示的增强现实模型确定为目标物体的增强现实模型。

获取电子设备的全球定位系统（Global Positioning System，GPS）定位信息。通过陀螺仪获取电子设备的朝向，将该朝向作为拍摄角度。根据定位信息以及拍摄角度，确定在一定拍照范围内是否具有地标建筑物。若具有地标建筑物则将该地标建筑物的增强现实模型作为目标物体的增强现实模型。

进一步的，在另一种实现方式中，获取目标物体的增强现实模型，可以通过下述方式实施：

从目标图像中确定目标物体的基础图像，根据该基础图像确定目标物体的增强现实模型。基础图像可以为目标图像中目标物体的图像。可以通过网络侧服务器对基础图像进行图像分析，进而确定该基础图像中的目标物体具体为何物。目标物体可以为建筑物、车辆、服装、鞋帽等物体。

上述两种实现方式可以用于查找目标物体的增强现实模型，也可以用于检验目标物体的增强现实模型是否准确。

步骤 130、获取用户选择的目标音频数据，根据目标音频数据确定具有时序性的音频特征。

目标音频数据可以为用户选择的歌曲或录音等音频数据。上述音频数据可以为本地存储的音频数据。也可以为用户在服务器提供的歌单中选择的音频数据。目标音频数据还可以为用户实时输入的音频数据。

可选的，对目标音频数据进行音频特征检测，得到目标音频数据的具有时序性的音频特征，音频特征包括重音、重拍或节拍中的一种或多种的组合。

音频特征检测用于解析目标音频数据中节拍规律、重音、重拍出现的位置或规律等。节拍为表示目标音频数据节奏的单位。目标音频数据中，具有一定强弱分别的一系列拍子在每隔一定时间重复出现。如 2/4 拍（四二拍）、4/4 拍（四四拍）、3/4 拍（四三拍）等。节拍随时间呈周期性变化。重音为目标音频中音量较高的音符。重拍指节拍中的强拍。

步骤 140、输出目标音频数据时，根据目标音频数据的播放进度和音频特征驱动增强现实模型。

当用户选择目标音频数据后，通过电子设备的音频输出模组输出目标音频数据。在输出目标音频数据的同时，根据当前的播放进度以及目标音频特征驱动增强现实模型。

示例性的，随着重音、轻拍或重拍的出现，可以驱动增强现实模型整体的颜色进行改变。重拍、轻拍或重音对应的颜色变化不同。

可选的，增强现实模型包括多个模型单元。根据目标音频数据的播放进度和音频特征驱动增强现实模型，包括：根据目标音频数据的播放进度和音频特征，驱动增强现实模型中的模型单元。

增强现实模型可以由多个模型单元组成，每个模型单元可以为一个立方体单元。拼接多个立方体单元形成目标物体的增强现实模型。可以通过着色器并行处理多个立方体。根据目标音频数据的播放进度和音频特征，驱动增强现实模型中的模型单元。

可选的，根据目标音频数据的播放进度和音频特征，驱动增强现实模型中的模型单元，包括：

步骤一、根据具有时序性的音频特征确定模型单元进行形态变化的目标时间和目标幅度。

可以将节拍中的轻拍、重拍和重音所在的时间确定为目标时间，根据轻拍、重拍和重音的程度确定目标幅度。

步骤二、若目标音频数据的播放进度为目标时间，则根据目标幅度驱动增强现实模型中的模型单元。

模型单元的驱动方式可以包括凸起动作、颜色变化或透明度变化等。

示例性的，根据目标幅度驱动增强现实模型中的模型单元，包括：

驱动增强现实模型中的预设模型单元进行凸起动作；或者，驱动增强现实模型中多个模型单元执行颜色变化；或者，驱动增强现实模型中多个模型单元执行透明度变化。

驱动增强现实模型中的预设模型单元进行凸起动作时，预设模型可以为随机选择的模型单元。根据目标幅度确定凸起动作的凸起幅度。驱动增强现实模型中多个模型单元执行颜色变化时，可以从全部模型单元中随机选出多个模型单元执行颜色变化或对全部模型单元执行颜色变化。透明度指模型单元纹理图形的透明度。纹理图像可以为纯色也可以为目标物体实际纹理图案。

进一步的，在获取用户选择的目标音频数据之后，还包括：

获取目标音频数据的情绪特征；根据情绪特征确定增强现实模型的第一形变特征，第一形变特征用于驱动增强现实模型的形状与情绪特征表达的情绪一致；输出目标音频数据时，根据第一形变特征驱动增强现实模型。

可以根据目标音频数据的歌曲名称确定情绪特征。或则根据目标音频中用户录入的文字内容确定情绪特征。情绪特征用于表示目标音频数据的音频内容的情绪偏向，例如欢快、忧郁等。可预先配置不同情绪特征对应的形变特征。例如欢快的情绪特征对应的第一形变

特征中模型单元的凸起幅度较高，凸起频率较快。例如忧郁的情绪特征对应的第一形变特征中模型单元的凸起幅度较低，凸起频率较慢。示例性的，在拍摄电视塔后，如若情绪特征为伤心，则电视塔模型呈弯腰形态。若建筑物模型出现形变，则通过裁剪方式删除建筑物原始图像，使用模型覆盖原始图像区域。

5 进一步的，在获取用户选择的目标音频数据之后，还包括：

获取用户的肢体动作或表情；根据肢体动作或表情确定增强现实模型的第二形变特征，第二形变特征用于驱动增强现实模型的形状与肢体动作或表情一致；输出目标音频数据时，根据第二形变特征驱动增强现实模型。

10 通过肢体检测可获取用户肢体动作。通过人脸识别可获取用户的表情。根据肢体动作或表情生成第二形变特征。示例性的，在拍摄电视塔后，如若人脸表情为悲伤，则第二形变特征为弯腰形态，此时电视塔模型呈弯腰形态。若建筑物模型出现形变，则通过裁剪方式删除建筑物原始图像，使用模型覆盖原始图像区域。

本公开实施例公开的增强现实的图像处理方法，能够响应于用户触发的图像获取指令，获取目标图像，目标图像包含目标物体；获取目标物体的增强现实模型，结合目标物体输出增强现实模型；获取用户选择的目标音频数据，根据目标音频数据确定具有时序性的音频特征；输出目标音频数据时，根据目标音频数据的播放进度和音频特征驱动增强现实模型。相对于目前增强现实模型缺乏互动性，易用性差。本公开实施例公开的增强现实的图像处理方法，能够在输出增强现实模型时，结合用户选择的目标音频数据的音频特征，驱动增强显示模型的输出，使得用户能够参与到增强显示模型的展示过程中，通过选择不同的目标音频数据，驱动增强现实模型根据目标音频数据的音频特征进行展示，提高易用性。

20 的目标音频数据，驱动增强现实模型根据目标音频数据的音频特征进行展示，提高易用性。

实施例二

图 2 为本公开实施例二提供的增强现实的图像处理装置的结构示意图，本实施例可适用于展示增强现实模型的情况，该方法可以由实现增强现实的电子设备来执行，该电子设备可以为智能手机、平板电脑等，该装置包括：目标图像获取模块 210、增强现实模型获取模块 220、目标音频获取模块 230、音频特征确定模块 240 或输出模块 250。

25 取模块 220、目标音频获取模块 230、音频特征确定模块 240 或输出模块 250。

目标图像获取模块 210，用于响应于用户触发的图像获取指令，获取目标图像，目标图像包含目标物体；

增强现实模型获取模块 220，用于获取目标物体的增强现实模型，结合目标物体输出增强现实模型；

30 目标音频获取模块 230，用于获取用户选择的目标音频数据；

音频特征确定模块 240，用于根据目标音频数据确定具有时序性的音频特征；

输出模块 250，用于输出目标音频数据时，根据目标音频数据的播放进度和音频特征驱动增强现实模型。

进一步的，音频特征确定模块 240 用于：

- 5 对目标音频数据进行音频特征检测，得到目标音频数据的具有时序性的音频特征，音频特征包括重音、重拍或节拍中的一种或多种的组合。

进一步的，增强现实模型包括多个模型单元；输出模块 250 用于：

根据目标音频数据的播放进度和音频特征，驱动增强现实模型中的模型单元。

进一步的，输出模块 250 用于：

- 10 根据具有时序性的音频特征确定模型单元进行形态变化的目标时间和目标幅度；
若目标音频数据的播放进度为目标时间，则根据目标幅度驱动增强现实模型中的模型单元。

进一步的，输出模块 250 用于：

驱动增强现实模型中的预设模型单元进行凸起动作；或者，

- 15 驱动增强现实模型中多个模型单元执行颜色变化；或者，
驱动增强现实模型中多个模型单元执行透明度变化。

进一步的，还包括第一形变特征获取模块。第一形变特征获取模块用于：

获取目标音频数据的情绪特征；

根据情绪特征确定增强现实模型的第一形变特征，第一形变特征用于驱动增强现实模

- 20 型的形状与情绪特征表达的情绪一致；

输出目标音频数据时，根据第一形变特征驱动增强现实模型。

进一步的，还包括第二形变特征获取模块。第二形变特征获取模块用于：

获取用户的肢体动作或表情；

根据肢体动作或表情确定增强现实模型的第二形变特征，第二形变特征用于驱动增强

- 25 现实模型的形状与肢体动作或表情一致；

输出目标音频数据时，根据第二形变特征驱动增强现实模型。

本公开实施例公开的增强现实的图像处理装置，目标图像获取模块 210 响应于用户触发的图像获取指令，获取目标图像，目标图像包含目标物体；增强现实模型获取模块 220 获取目标物体的增强现实模型，结合目标物体输出增强现实模型；目标音频获取模块 230

- 30 获取用户选择的目标音频数据，音频特征确定模块 240 根据目标音频数据确定具有时序性

的音频特征；输出模块 250 输出目标音频数据时，根据目标音频数据的播放进度和音频特征驱动增强现实模型。相对于目前增强现实模型缺乏互动性，易用性差。本公开实施例公开的增强现实的图像处理装置，能够在输出增强现实模型时，结合用户选择的目标音频数据的音频特征，驱动增强显示模型的输出，使得用户能够参与到增强显示模型的展示过程中，通过选择不同的目标音频数据，驱动增强现实模型根据目标音频数据的音频特征进行展示，提高易用性。

本发明实施例所提供的增强现实的图像处理装置可执行本发明任意实施例所提供的增强现实的图像处理方法，具备执行方法相应的功能模块和有益效果。

实施例三

下面参考图 3，其示出了适于用来实现本公开实施例三的电子设备 800 的结构示意图。本公开实施例中的终端设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、Personal digital assistant (PDA，个人数字助理)、Portable android device (PAD，平板电脑)、Portable media player (PMP，便携式多媒体播放器)、车载终端（例如车载导航终端）等等的移动终端以及诸如数字 TV (Television)、台式计算机等等的固定终端。图 3 示出的电子设备仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

如图 3 所示，电子设备 800 可以包括处理装置（例如中央处理器、图形处理器等）801，其可以根据存储在只读存储器（Read only memory, ROM）802 中的程序或者从存储装置 808 加载到随机访问存储器（Random access memory, RAM）803 中的程序而执行各种适当的动作和处理。在 RAM 803 中，还存储有电子设备 800 操作所需的各种程序和数据。处理装置 801、ROM 802 以及 RAM 803 通过总线 804 彼此相连。输入/输出（Input/output, I/O）接口 805 也连接至总线 804。

通常，以下装置可以连接至 I/O 接口 805：包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 806；包括例如液晶显示器（Liquid crystal display, LCD）、扬声器、振动器等输出装置 807；包括例如磁带、硬盘等的存储装置 808；以及通信装置 809。通信装置 809 可以允许电子设备 800 与其他设备进行无线或有线通信以交换数据。虽然图 3 示出了具有各种装置的电子设备 800，但是应理解的是，并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在计算机可读介质上

的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信装置 809 从网络上被下载和安装，或者从存储装置 808 被安装，或者从 ROM 802 被安装。在该计算机程序被处理装置 801 执行时，执行本公开实施例的方法中的步骤，以实现其限定的上述功能。

5 根据本公开的一个或多个实施例提供的一种计算机程序产品，该计算机程序产品包括：计算机程序，该计算机程序存储在可读存储介质中，电子设备的一个或多个处理器可以从所述可读存储介质读取所述计算机程序，所述一个或多个处理器执行所述计算机程序，使得所述电子设备执行上述任一实施例提供的方案。

10 本公开的一个或多个实施例提供的一种计算机程序，该计算机程序存储在可读存储介质中，电子设备的一个或多个处理器可以从所述可读存储介质中读取所述计算机程序，所述一个或多个处理器执行所述计算机程序，使得所述电子设备执行上述任一实施例提供的方案。

15 需要说明的是，本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（Electrical programmable read only memory, EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（Compact disc read only memory, CD-ROM）、光存储器件、磁
20 存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是
25 计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、射频（Radio Frequency, RF）等等，或者上述的任意合适的组合。

30 上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被该电子设备执行时，使得该电子设备：获取至少两个网际协议地址；向节点评价设备发送包括所述至少两个网际协议地址的节点评价请求，其中，所述节点评价设备从所述至少两个网际协议地址中，选取网际协议地址并返回；接收所述节点评价设备返回的网际协议地址；其中，

5 所获取的网际协议地址指示内容分发网络中的边缘节点。

或者，上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被该电子设备执行时，使得该电子设备：接收包括至少两个网际协议地址的节点评价请求；从所述至少两个网际协议地址中，选取网际协议地址；返回选取出的网际协议地址；其中，接收到的网际协议地址指示内容分发网络中的边缘节点。

- 10 可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上述程序设计语言包括面向对象的程序设计语言—诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。
- 15 在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(Local area network, LAN)或广域网(Wide area network, WAN)—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

- 附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表
- 20 一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，
- 25 可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。其中，单元的名称在某种情况下并不构成对该单元本身的限定，例如，第一获取单元还可以被描述为“获取至少两个网际协议地址的单元”。

- 30 以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应

当理解，本公开中所涉及的公开范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述公开构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开中公开的（但不限于）具有类似功能的技术特征进行互相替换而形成的技术方案。

权利要求书

1. 一种增强现实的图像处理方法，其特征在于，包括：

响应于用户触发的图像获取指令，获取目标图像，所述目标图像包含目标物体；
获取所述目标物体的增强现实模型，结合所述目标物体输出所述增强现实模型；

5 获取所述用户选择的目标音频数据，根据所述目标音频数据确定具有时序性的音频特征；

输出所述目标音频数据时，根据所述目标音频数据的播放进度和所述音频特征驱动所述增强现实模型。

2. 根据权利要求 1 所述的方法，其特征在于，所述根据所述目标音频数据确定具有时序性的音频特征，包括：

对所述目标音频数据进行音频特征检测，得到所述目标音频数据的所述具有时序性的音频特征，所述音频特征包括重音、重拍或节拍中的一种或多种的组合。

3. 根据权利要求 1 或 2 所述的方法，其特征在于，所述增强现实模型包括多个模型单元；

15 所述根据所述目标音频数据的播放进度和所述音频特征驱动所述增强现实模型，包括：

根据所述目标音频数据的播放进度和所述音频特征，驱动所述增强现实模型中的模型单元。

4. 根据权利要求 3 所述的方法，其特征在于，根据所述目标音频数据的播放进度和所述音频特征，驱动所述增强现实模型中的模型单元，包括：

根据所述具有时序性的音频特征确定所述模型单元进行形态变化的目标时间和目标幅度；

若所述目标音频数据的播放进度为所述目标时间，则根据所述目标幅度驱动所述增强现实模型中的模型单元。

25 5. 根据权利要求 4 所述的方法，其特征在于，所述根据所述目标幅度驱动所述增强现实模型中的模型单元，包括：

驱动所述增强现实模型中的预设模型单元进行凸起动作；或者，

驱动所述增强现实模型中多个模型单元执行颜色变化；或者，

驱动所述增强现实模型中多个模型单元执行透明度变化。

30 6. 根据权利要求 1-5 中任一所述的方法，其特征在于，在获取所述用户选择的目标音频数据之后，还包括：

获取所述目标音频数据的情绪特征；

根据所述情绪特征确定所述增强现实模型的第一形变特征，所述第一形变特征用

于驱动所述增强现实模型的形状与所述情绪特征表达的情绪一致；

输出所述目标音频数据时，根据所述第一形变特征驱动所述增强现实模型。

7. 根据权利要求 1-5 中任一所述的方法，其特征在于，在获取所述用户选择的目标音频数据之后，还包括：

5 获取所述用户的肢体动作或表情；

根据所述肢体动作或表情确定所述增强现实模型的第二形变特征，所述第二形变特征用于驱动所述增强现实模型的形状与所述肢体动作或表情一致；

输出所述目标音频数据时，根据所述第二形变特征驱动所述增强现实模型。

8. 一种增强现实的图像处理装置，其特征在于，包括：

10 目标图像获取模块，用于响应于用户触发的图像获取指令，获取目标图像，所述目标图像包含目标物体；

增强现实模型获取模块，用于获取所述目标物体的增强现实模型，结合所述目标物体输出所述增强现实模型；

目标音频获取模块，用于获取所述用户选择的目标音频数据；

15 音频特征确定模块，用于根据所述目标音频数据确定具有时序性的音频特征；

输出模块，用于输出所述目标音频数据时，根据所述目标音频数据的播放进度和所述音频特征驱动所述增强现实模型。

9. 一种电子设备，其特征在于，所述电子设备包括：

一个或多个处理器；

20 存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-7 中任一所述的增强现实的图像处理方法。

10. 一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行如权利要求 1-7 中任一所述的增强现实的图像处理方法。

25 11. 一种计算机程序产品，其特征在于，包括计算机程序指令，所述计算机程序指令使得计算机执行如权利要求 1-7 中任一所述的增强现实的图像处理方法。

12. 一种计算机程序，其特征在于，所述计算机程序使得计算机执行如权利要求 1-7 中任一所述的增强现实的图像处理方法。

30

1/2

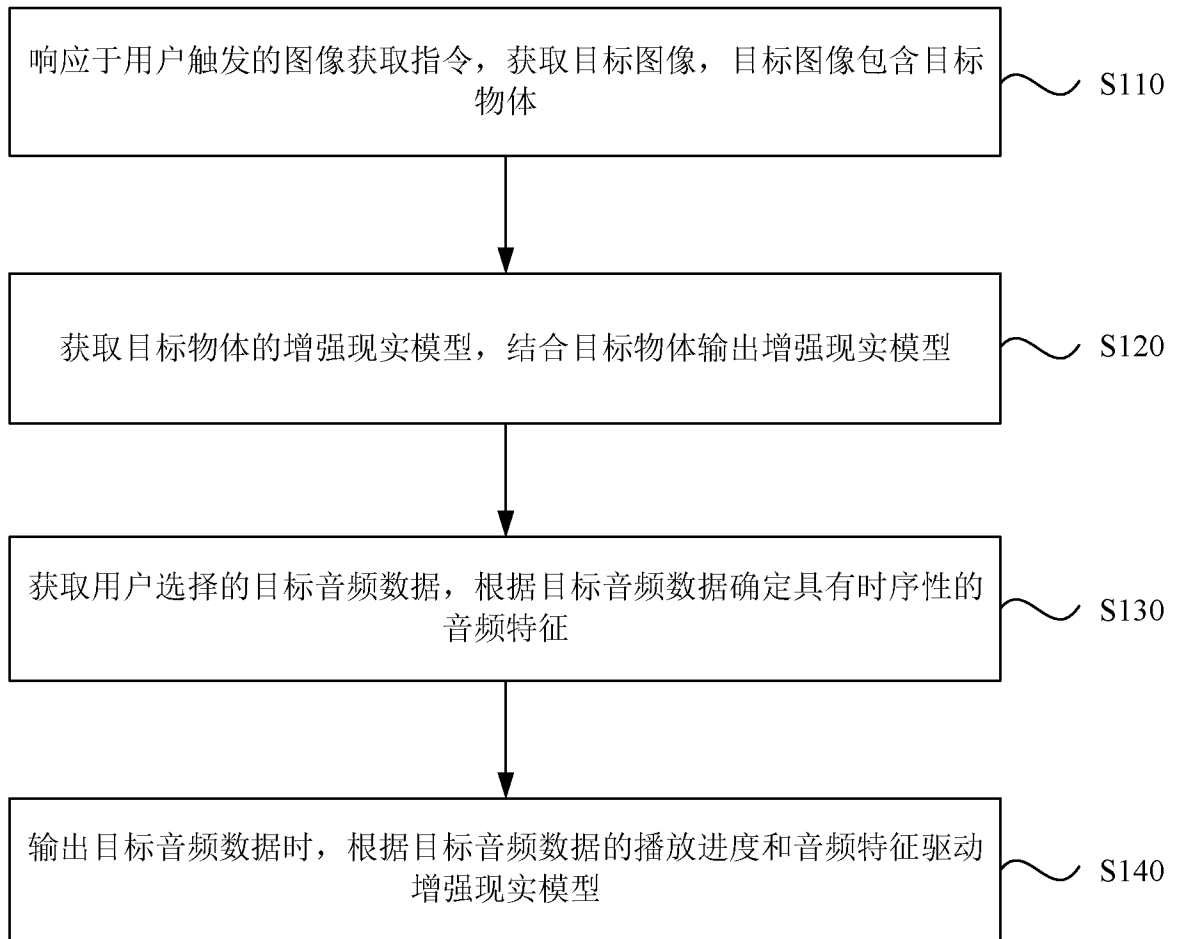


图 1

2/2

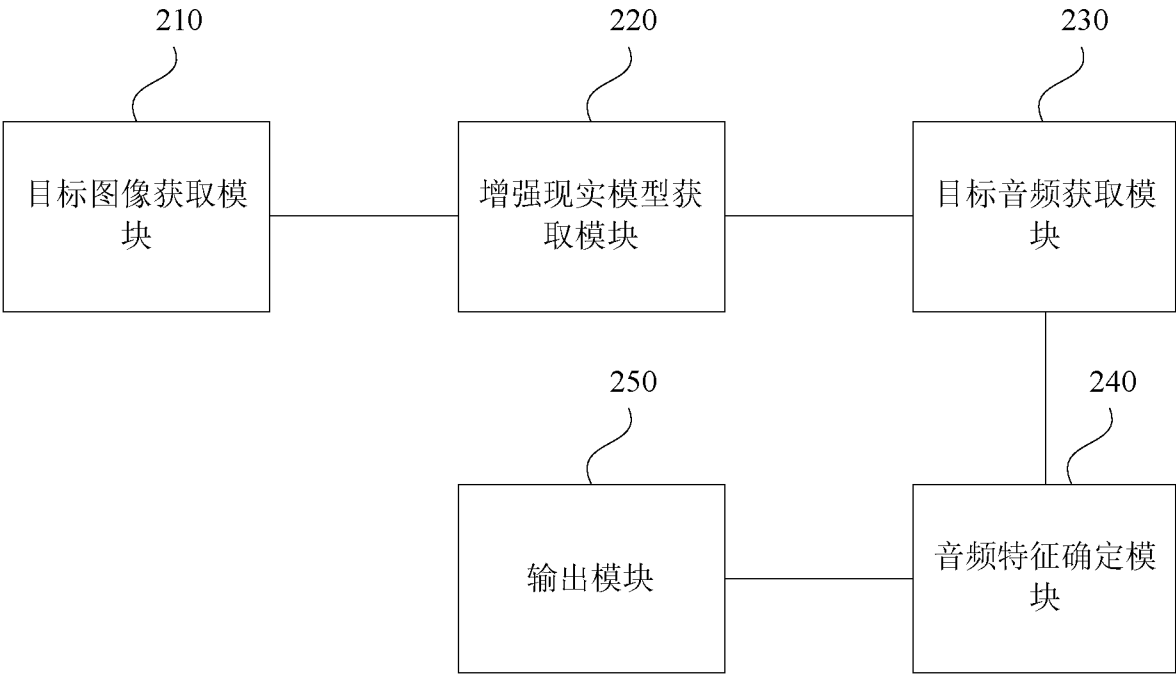


图 2

800

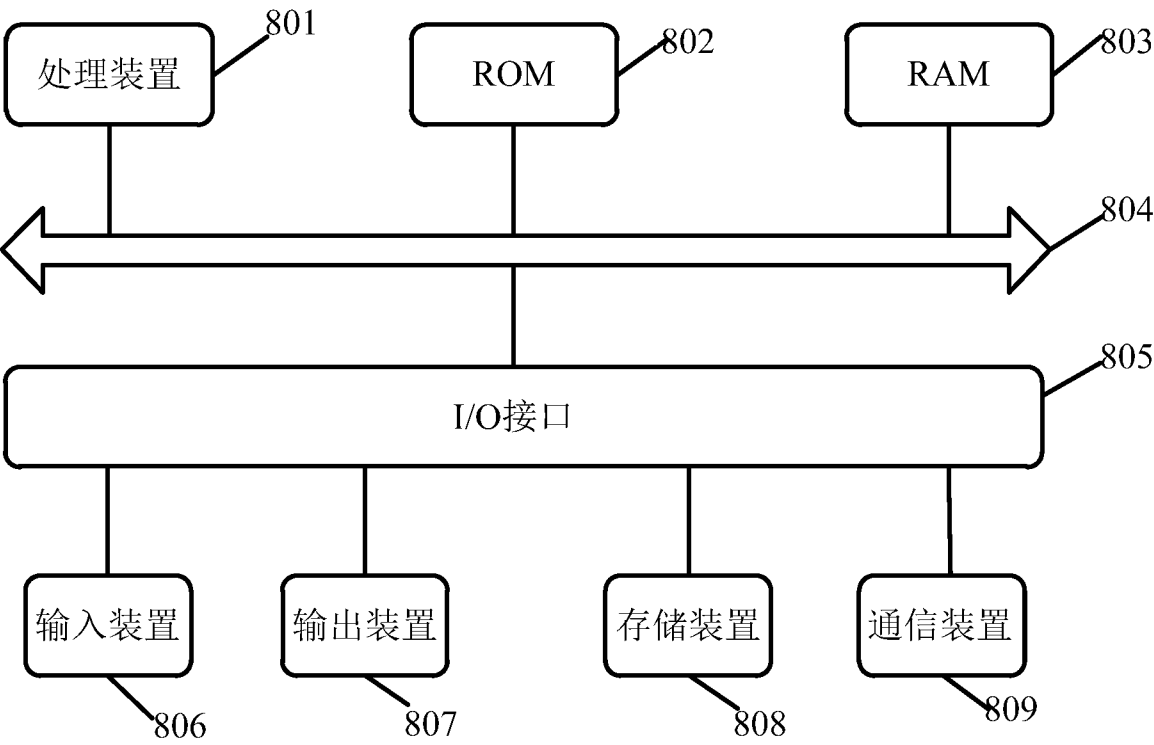


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/098456

A. CLASSIFICATION OF SUBJECT MATTER

G06T 19/00(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)

G06T

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, WPI, EPODOC, CNKI, IEEE: 增强现实, 节奏, 自动生成, 音乐, 配乐, ar, 角色, 配合, 模型, 音频, 虚拟角色, 驱动, 图像, 拍照, 摄像, 进度, 情绪, 动作, 表情, role, model, music, audio frequency, drive, picture, movement, mood

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111833460 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD. et al.) 27 October 2020 (2020-10-27) claims 1-10, description paragraphs [0096]-[0101]	1-12
A	CN 106506464 A (WUHAN SHOW BABY SOFTWARE CO., LTD.) 15 March 2017 (2017-03-15) description paragraphs [0070]-[0096]	1-12
A	CN 109407918 A (SUZHOU MXR SOFTWARE TECHNOLOGY CO., LTD.) 01 March 2019 (2019-03-01) entire document	1-12
A	CN 105608745 A (DALIAN NEW VISION MEDIA CO., LTD.) 25 May 2016 (2016-05-25) entire document	1-12
A	WO 2019029100 A1 (UNIV. SHANDONG SCIENCE & TECH.) 14 February 2019 (2019-02-14) entire document	1-12



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

20 August 2021

Date of mailing of the international search report

03 September 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/CN2021/098456

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111833460	A	27 October 2020	None			
CN	106506464	A	15 March 2017	None			
CN	109407918	A	01 March 2019	None			
CN	105608745	A	25 May 2016	WO	2017107758	A1	29 June 2017
WO	2019029100	A1	14 February 2019	CN	107515674	A	26 December 2017

国际检索报告

国际申请号

PCT/CN2021/098456

A. 主题的分类 G06T 19/00 (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) G06T 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)) CNPAT, WPI, EPDOC, CNKI, IEEE: 增强现实, 节奏, 自动生成, 音乐, 配乐, ar, 角色, 配合, 模型, 音频, 虚拟角色, 驱动, 图像, 拍照, 摄像, 进度, 情绪, 动作, 表情, role, model, music, audio frequency, drive, picture, movement, mood																				
C. 相关文件 <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>PX</td> <td>CN 111833460 A (北京字节跳动网络技术有限公司 等) 2020年 10月 27日 (2020 - 10 - 27) 权利要求1-10、说明书第[0096]-[0101]段</td> <td>1-12</td> </tr> <tr> <td>A</td> <td>CN 106506464 A (武汉秀宝软件有限公司) 2017年 3月 15日 (2017 - 03 - 15) 说明书第[0070]-[0096]段</td> <td>1-12</td> </tr> <tr> <td>A</td> <td>CN 109407918 A (苏州梦想人软件科技有限公司) 2019年 3月 1日 (2019 - 03 - 01) 全文</td> <td>1-12</td> </tr> <tr> <td>A</td> <td>CN 105608745 A (大连新锐天地传媒有限公司) 2016年 5月 25日 (2016 - 05 - 25) 全文</td> <td>1-12</td> </tr> <tr> <td>A</td> <td>WO 2019029100 A1 (UNIV. SHANDONG SCIENCE & TECH.) 2019年 2月 14日 (2019 - 02 - 14) 全文</td> <td>1-12</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	PX	CN 111833460 A (北京字节跳动网络技术有限公司 等) 2020年 10月 27日 (2020 - 10 - 27) 权利要求1-10、说明书第[0096]-[0101]段	1-12	A	CN 106506464 A (武汉秀宝软件有限公司) 2017年 3月 15日 (2017 - 03 - 15) 说明书第[0070]-[0096]段	1-12	A	CN 109407918 A (苏州梦想人软件科技有限公司) 2019年 3月 1日 (2019 - 03 - 01) 全文	1-12	A	CN 105608745 A (大连新锐天地传媒有限公司) 2016年 5月 25日 (2016 - 05 - 25) 全文	1-12	A	WO 2019029100 A1 (UNIV. SHANDONG SCIENCE & TECH.) 2019年 2月 14日 (2019 - 02 - 14) 全文	1-12
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 111833460 A (北京字节跳动网络技术有限公司 等) 2020年 10月 27日 (2020 - 10 - 27) 权利要求1-10、说明书第[0096]-[0101]段	1-12																		
A	CN 106506464 A (武汉秀宝软件有限公司) 2017年 3月 15日 (2017 - 03 - 15) 说明书第[0070]-[0096]段	1-12																		
A	CN 109407918 A (苏州梦想人软件科技有限公司) 2019年 3月 1日 (2019 - 03 - 01) 全文	1-12																		
A	CN 105608745 A (大连新锐天地传媒有限公司) 2016年 5月 25日 (2016 - 05 - 25) 全文	1-12																		
A	WO 2019029100 A1 (UNIV. SHANDONG SCIENCE & TECH.) 2019年 2月 14日 (2019 - 02 - 14) 全文	1-12																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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																
* 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 8月 20日		国际检索报告邮寄日期 2021年 9月 3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 李文浩 电话号码 86-(10)-53961341																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/098456

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	111833460	A	2020年 10月 27日	无	
CN	106506464	A	2017年 3月 15日	无	
CN	109407918	A	2019年 3月 1日	无	
CN	105608745	A	2016年 5月 25日	WO 2017107758	A1 2017年 6月 29日
WO	2019029100	A1	2019年 2月 14日	CN 107515674	A 2017年 12月 26日