

一种物体建模运动方法、装置与设备

5 技术领域

本发明涉及终端技术领域，尤其涉及一种物体建模运动的方法、装置与设备。

背景技术

随着信息和通信技术的发展，人们在生活和工作中接触到越来越多的图形图像。获取
10 图像的方法包括使用各种摄像机、照相机、扫描仪等，利用这些手段通常只能得到物体的
平面图像，即物体的二维信息。在许多领域，如机器视觉、面形检测、实物仿形、自动加
工、产品质量控制、生物医学等，物体的三维信息是必不可少的。因此 3D 扫描技术开始
应运而生，通常采用的设备是三维扫描仪(3D scanner)；它是一种科学仪器，用来侦测并分
析现实世界中物体或环境的形状（几何构造）与外观数据（如颜色、表面反照率等性质）。
15 三维扫描仪的用途是创建物体几何表面的点云（point cloud），这些点可用来插补成物体的
表面形状，越密集的点云可以创建更精确的模型（这个过程也称做三维重建）。若扫描仪
能够取得表面颜色，则可进一步在重建的表面上粘贴材质贴图，亦即所谓的材质映射
（texture mapping）。

然而现有技术中的三维扫描仪使用复杂，需要具有专业技能的用户才能玩转，并且应
20 用场景也比较受限；因此如何让大众用户能够玩转三维扫描技术是亟待解决的问题。

发明内容

本发明实施例提供了一种物体建模运动的方法、装置与设备，可以随时随地扫描感兴
趣的物体，并实现动态效果，增强趣味性和可玩性，提升用户粘性；引领时代潮流。

25 本发明实施例提供的具体技术方案如下：

第一方面，本发明实施例提供一种物体建模运动方法，该方法应用于移动终端，该移
动终端包括彩色相机和深度相机；且彩色相机和深度相机位于移动终端的同一侧，正面或
者反面；该方法具体包括：利用彩色相机和深度相机对目标物体进行全景扫描，得到目标
物体的 3D 模型；获取目标骨骼模型；将目标骨骼模型与目标物体的 3D 模型融合；获取目
30 标运动方式；根据目标运动方式控制上述骨骼模型，使目标物体的 3D 模型根据目标运动
方式进行运动。

第二方面，本发明实施例提供一种物体建模运动装置，该装置应用于移动终端，该移
动终端包括彩色相机和深度相机；且彩色相机和深度相机位于移动终端的同一侧，正面或
者反面；该装置包括：扫描模块，用于当彩色相机和深度相机对目标物体进行全景扫描时，
35 得到目标物体的 3D 模型；第一获取模块，用于获取目标骨骼模型；融合模块；用于将目
标骨骼模型与目标物体的 3D 模型融合；第二获取模块，用于获取目标运动方式；运动模

块，用于根据目标运动方式控制骨骼模型，使目标物体的 3D 模型根据目标运动方式进行运动。

根据本发明实施例提供的上述方法和装置的技术方案，移动终端可以实现物体从扫描、3D 重建、到骨骼装配，再到预设的动画展示的一体化设计。无需用户采用专业的、笨重的、复杂的设备进行专业扫描，也无需再跑到 PC 端做复杂的建模处理和动画处理，将这些功能集成在一起，提供给用户，使得用户在一个移动终端上能够轻松玩转这一系列的操作方法，使得用户身边任何一个“静态的物体（或近似静态的物体）”能够更鲜活，更具生命力。增加用户使用终端的趣味性，提升用户的体验度。

10 应理解，在媒体领域，“相机”与“摄像头”可以同义。

根据第一方面或者第二方面，在一种可能的设计中：深度相机可以采用 TOF 模组。

根据第一方面或者第二方面，在一种可能的设计中：深度相机可以采用结构光模组。

根据第一方面或者第二方面，在一种可能的设计中：深度相机的视场角范围为 40 度到 80 度。

15 根据第一方面或者第二方面，在一种可能的设计中：深度相机中红外线的发射功率的范围可以选择现在 50-400mw 之间；特殊应用下的超强光线，发生功率可以更高。

根据第一方面或者第二方面，在一种可能的设计中：扫描物体时的扫描距离为介于 20cm 到 80cm，扫描距离可以理解为深度相机到目标物体的距离。

20 根据第一方面或者第二方面，在一种可能的设计中：扫描过程中深度相机的拍摄帧率可以选择不低于 25fps。

根据第一方面或者第二方面，在一种可能的设计中：骨骼模型可以根据 3D 模型，通过一系列的算法计算得到。

25 根据第一方面或者第二方面，在一种可能的设计中：可以提供给用户骨骼模型制作库，如一些线段和点，其中，线段表示骨骼，点表示关节节点。接收用户的操作指令，如手势、滑动、快捷键等，将至少两个线段和至少一个点组合成骨骼模型，从而得到骨骼模型，进一步地，将该骨骼模型上传到云端或存储在本地。该方法可以由第一获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

30 根据第一方面或者第二方面，在一种可能的设计中：可以提供给用户更加开放的制作库，线段和点完全由用户自由设计，其中，线段表示骨骼，点表示关节节点。接收用户的操作指令，如手势、滑动、快捷键等，将至少两个线段和至少一个点组合成骨骼模型，从而得到骨骼模型，进一步地，将该骨骼模型上传到云端或存储在本地。该方法可以由第一获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

35 根据第一方面或者第二方面，在一种可能的设计中：在一种具体实现过程中，可以从至少一个预设骨骼模型中选择与所述目标物体的外形匹配度最高的骨骼模型作为目标骨骼模型。预设骨骼模型可以存储在网络、云端或者本地。例如本地存储了鸡的骨骼模型、狗的骨骼模型、鱼的骨骼模型，当目标物体为鸭子的时候，系统通过外形识别，将鸡的骨骼模型作为目标骨骼模型。相似的判定标准包括但不限于骨骼形态、骨骼的长度、骨骼的粗细、骨骼的数量、骨骼的组成方式等。该方法可以由第一获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

根据第一方面或者第二方面，在一种可能的设计中：可以接收用户的选择指令，该选择指令用于从至少一个预设骨骼模型中选择出目标骨骼模型，这些预设模型是本地存储或者是从云端或网络上调用的。该方法可以由第一获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

- 5 根据第一方面或者第二方面，在一种可能的设计中：可以获取第一物体的运动方式，将第一物体的运动方式作为目标运动方式。其中，第一物体可以是当前实时运动的物体；也可以是过去拍过且保存下来物体的运动方式；也可以是预设的某一特定物体的运动方式。该方法可以由第二获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

- 10 根据第一方面或者第二方面，在一种可能的设计中：可以从预设的目标运动方式选
出一种。该方法可以由第二获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。其中，预设的目标运动方式可以是一套完整的运动方式，也可以是与用户操作相对应的运动方式，如用户对着终端中显示的“复活的物体”招手，则物体可以按照预设的招手回应动画方式进行运动。更为通用的，用户可以对终端输入预设的交互动作，终端根据该交互动作获得对应的回应运动方式，并控制物体的 3D 模型按照回应运动方式进行运动。

- 15 根据第一方面或者第二方面，在一种可能的设计中：运动方式可以是用户自己使用动画制作软件制作的，当然这个软件可以是系统内嵌的工具集或者在扫描重建运动的 APP 中载入的工具集，或者来源于第三方动画设计工具；可以是当前制作也可以是历史制作的运动方式或动画。该方法可以由第二获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

- 20 根据第一方面或者第二方面，在一种可能的设计中：运动方式可以是根据物理属性在多个预先存储的运动方式中选择出属性匹配度最高的运动方式作为目标运动方式。该方法可以由第二获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

- 25 根据第一方面或者第二方面，在一种可能的设计中：运动方式还可以是基于目标物体的骨骼模型（可以采用上一步骤中的任一方法去获取）由系统或者用户对该骨骼模型进行自主设计，得到目标运动方式。这一种方式是对后续物体的 3D 模型实现动画的最契合的动画操作。该方法可以由第二获取模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

- 30 根据第一方面或者第二方面，在一种可能的设计中：用户扫描完物体的 3D 模型之后，可以将 3D 模型存储在本地或者云端，若干时间以后可以直接调用该 3D 模型，自由选择骨骼装配或自由选择动画方式，还可以选择动画背景，包括但不限于实时拍摄的图像、本地已经存储的其它图像、云端数据图像等。此外，在实现物体动画的同时，还可以显示出物体的影子，或者增加其它的音效、特效等等。动画可以由移动终端自动播放，还可以在通过用户输入操作指令进行播放的控制。

- 35 根据第一方面或者第二方面，在一种可能的设计中：在根据所述目标运动方式控制所述骨骼模型运动的过程中，对所述骨骼模型和所述目标物体的 3D 模型进行蒙皮操作；所述蒙皮操作用于根据骨骼模型的运动，确定所述 3D 模型表面上点的位置变化；使所述目标物体的 3D 模型跟随所述骨骼模型进行运动。该方法可以由运动模块完成；硬件上可以由处理器调用存储器中的程序指令实现。

通过上述可能的设计，加强了人机交互的程度，给用户更加自由的发挥空间，可以让用户在物体复活的过程中深度参与其中，开发想象增加乐趣。

更具体地，上述操作中涉及到的其它技术实现可以由处理器调用存储器中的程序与指令进行相应的处理，如使能相机，采集图像，生成 3D 模型，获取骨骼模型或动画、存储骨骼模型或动画、增加特效以及用户之间的交互操作等。

第三方面，本发明实施例提供一种终端设备，所述终端设备包含存储器、处理器、总线、深度相机和所述彩色相机；彩色相机和深度相机位于移动终端的同一侧；存储器、深度相机、彩色相机以及处理器通过总线相连；深度相机和彩色相机用于在处理器的控制下对目标物体进行全景扫描；存储器用于存储计算机程序和指令；处理器用于调用存储器中存储的计算机程序和指令，使终端设备执行如上述任何一种可能的设计方法。

根据第三方面，在一种可能的设计中，终端设备还包括天线系统、天线系统在处理器的控制下，收发无线通信信号实现与移动通信网络的无线通信；移动通信网络包括以下的一种或多种：GSM 网络、CDMA 网络、3G 网络、4G 网络、5G 网络、FDMA、TDMA、PDC、TACS、AMPS、WCDMA、TDSCDMA、WIFI 以及 LTE 网络。

应理解，发明内容可以包含权利要求中所能实现的所有方法，此处不予以列举。

对于上述任意一种可能的实现方法和步骤，在不违背自然规律的前提下，可以进行不同方法和步骤之间的自由组合，不同的方法和步骤还可以增加或减少一些可能的步骤。本发明不予以穷举和赘述。

本发明实现了物体从扫描、3D 重建、到骨骼装配，再到预设的动画展示都可以在一个终端内完成，实现了静态物体的复活，提升用户使用移动终端的趣味性。

附图说明

- 图 1 为本发明实施例中一种终端的结构示意图；
图 2 为本发明实施例中一种物体建模运动方法的流程图；
图 3 为本发明实施例中一种是对物体扫描到实现动画的主要过程；
图 4 为本发明实施例中一种结构光示意图；
图 5 为本发明实施例中一种 TOF 示意图；
图 6 为本发明实施例中一种网格化+纹理映射的方法流程图；
图 7 为本发明实施例中一种具体的网格化实现方案流程图；
图 8 为本发明实施例中一种具体的纹理映射实现方案流程图；
图 9 为本发明实施例中一种具体的网格化+纹理映射实例；
图 10 为本发明实施例中一种具体的骨骼装配方案流程图；
图 11 为本发明实施例中一种具体的动画流程图；
图 12 为本发明实施例中一种物体建模运动装置示意图。

具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，并不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

- 5 本发明实施例中，移动终端，可以是向用户提供拍照和/或数据连通性的设备，具有无线连接功能的手持式设备、或连接到无线调制解调器的其他处理设备，比如：数码相机、单反相机、智能手机，还可以是其它带有拍照功能以及显示功能的智能设备，如可穿戴设备、平板电脑、PDA (Personal Digital Assistant, 个人数字助理)、无人机、航拍器等。

图1示出了终端100的一种可选的硬件结构示意图。

- 10 参考图1所示，终端100可以包括射频单元110、存储器120、输入单元130、显示单元140、拍摄单元150、音频电路160、扬声器161、麦克风162、处理器170、外部接口180、电源190等部件。

- 射频单元110可用于收发信息或通话过程中信号的接收和发送，特别地，将基站的下行信息接收后，给处理器170处理；另外，将设计上的数据发送给基站。通常，RF电路
15 包括但不限于天线、至少一个放大器、收发信机、耦合器、低噪声放大器 (Low Noise Amplifier, LNA)、双工器等。此外，射频单元110还可以通过无线通信与网络设备和其他设备通信。所述无线通信可以使用任一通信标准或协议，包括但不限于全球移动通讯系统 (Global System of Mobile communication, GSM)、通用分组无线服务 (General Packet Radio Service, GPRS)、码分多址 (Code Division Multiple Access, CDMA)、宽带码分多址
20 (Wideband Code Division Multiple Access, WCDMA)、长期演进 (Long Term Evolution, LTE)、电子邮件、短消息服务 (Short Messaging Service, SMS)等。

- 存储器120可用于存储指令和数据，存储器120可主要包括存储指令区和存储数据区，存储数据区可存储关节触摸手势与应用程序功能的关联关系；存储指令区可存储操作系统、应用、至少一个功能所需的指令等软件单元，或者他们的子集、扩展集。还可以包括非易
25 失性随机存储器；向处理器170提供包括管理计算处理设备中的硬件、软件以及数据资源，支持控制软件和应用。还用于多媒体文件的存储，以及运行程序 and 应用的存储。

- 输入单元130可用于接收输入的数字或字符信息，以及产生与所述便携式多功能装置的用户设置以及功能控制有关的键信号输入。具体地，输入单元130可包括触摸屏131
30 以及其他输入设备132。所述触摸屏131可收集用户在其上或附近的触摸操作（比如用户使用手指、关节、触笔等任何适合的物体在触摸屏上或在触摸屏的附近操作），并根据预先设定的程序驱动相应的连接装置。触摸屏可以检测用户对触摸屏的触摸动作，将所述触摸动作转换为触摸信号发送给所述处理器170，并能接收所述处理器170发来的命令并加以执行；所述触摸信号至少包括触点坐标信息。所述触摸屏131可以提供所述终端100和用户之间的输入界面和输出界面。此外，可以采用电阻式、电容式、红外线以及表面声波等
35 多种类型实现触摸屏。除了触摸屏131，输入单元130还可以包括其他输入设备。具体地，其他输入设备132可以包括但不限于物理键盘、功能键（比如音量控制按键132、开关按键133等）、轨迹球、鼠标、操作杆等中的一种或多种。

进一步的，触摸屏131可覆盖显示面板141，当触摸屏131检测到在其上或附近的触摸操作后，传送给处理器170以确定触摸事件的类型，随后处理器170根据触摸事件的类

型在显示面板 141 上提供相应的视觉输出。在本实施例中，触摸屏与显示单元可以集成为一个部件而实现终端 100 的输入、输出、显示功能；为便于描述，本发明实施例以触摸显示屏代表触摸屏和显示单元的功能集合；在某些实施例中，触摸屏与显示单元也可以作为两个独立的部件。

5 显示单元 140 可用于显示由用户输入的信息或提供给用户的信息以及终端 100 的各种菜单。在本发明实施例中，显示单元还用于显示设备利用摄像头 150 获取到的图像，可以包括某些拍摄模式下的预览图像、拍摄的初始图像以及拍摄后经过一定算法处理后的目标图像。

10 拍摄单元 150，用于采集图像或视频，可以通过应用程序指令触发开启，实现拍照或者摄像功能。拍摄单元可以包括成像镜头，滤光片，图像传感器等部件。物体发出或反射的光线进入成像镜头，通过滤光片，最终汇聚在图像传感器上。成像镜头主要是用于对拍照视角中的物体（也可称为待拍摄对象或目标物体）发出或反射的光汇聚成像；滤光片主要是用于将光线中的多余光波（例如除可见光外的光波，如红外）滤去；图像传感器主要是用于对接收到的光信号进行光电转换，转换成电信号，并输入到处理 170 进行后续处理。

15 具体地，拍摄单元 150 还可以包括彩色相机（彩色摄像头）151 和深度相机（深度摄像头）152；彩色相机用于采集目标物体的彩色图像，包含当下流行的终端产品中通用的彩色相机。深度相机用于获取目标物体的深度信息，作为举例，深度相机的可以通过 TOF 技术和结构光技术实现。

20 其中，TOF 是飞行时间（Time of Flight）技术的缩写，即传感器发出经调制的近红外光，遇物体后反射，传感器通过计算光线发射和反射时间差或相位差，来换算被拍摄景物的距离，以产生深度信息。此外，再结合传统的彩色摄像头拍摄，就能将物体的三维轮廓以不同颜色代表不同距离的地形图方式呈现出来。

25 其中，结构光是一组由投影元件和摄像头组成的系统结构。用投影元件投射特定的光信息（如经过光栅衍射）到物体表面后及背景后，由摄像头采集。根据物体造成的光信号的变化（如光线粗细的变化与位移）来计算物体的位置和深度等信息；进而复原整个三维空间。

30 音频电路 160、扬声器 161、麦克风 162 可提供用户与终端 100 之间的音频接口。音频电路 160 可将接收到的音频数据转换后的电信号，传输到扬声器 161，由扬声器 161 转换为声音信号输出；另一方面，麦克风 162 用于收集声音信号，还可以将收集的声音信号转换为电信号，由音频电路 160 接收后转换为音频数据，再将音频数据输出处理器 170 处理后，经射频单元 110 以发送给比如另一终端，或者将音频数据输出至存储器 120 以便进一步处理，音频电路也可以包括耳机插孔 163，用于提供音频电路和耳机之间的连接接口。

35 处理器 170 是终端 100 的控制中心，利用各种接口和线路连接整个手机的各个部分，通过运行或执行存储在存储器 120 内的指令以及调用存储在存储器 120 内的数据，执行终端 100 的各种功能和处理数据，从而对手机进行整体监控。可选的，处理器 170 可包括一个或多个处理单元；优选的，处理器 170 可集成应用处理器和调制解调处理器，其中，应用处理器主要处理操作系统、用户界面和应用程序等，调制解调处理器主要处理无线通信。可以理解的是，上述调制解调处理器也可以不集成到处理器 170 中。在一些实施例中，处理器、存储器、可以在单一芯片上实现，在一些实施例中，他们也可以在独立的芯片上分

别实现。处理器 170 还可以用于产生相应的操作控制信号,发给计算处理设备相应的部件,读取以及处理软件中的数据,尤其是读取和处理存储器 120 中的数据和程序,以使其中的各个功能模块执行相应的功能,从而控制相应的部件按指令的要求进行动作。

- 5 终端 100 还包括外部接口 180,所述外部接口可以是标准的 Micro USB 接口,也可以使用多针连接器,可以用于连接终端 100 与其他装置进行通信,也可以用于连接充电器为终端 100 充电。

终端 100 还包括给各个部件供电的电源 190 (比如电池),优选的,电源可以通过电源管理系统与处理器 170 逻辑相连,从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。

- 10 尽管未示出,终端 100 还可以包括闪光灯、无线保真 (wireless fidelity, WiFi)模块、蓝牙模块、不同功能的传感器等,在此不再赘述。下文中描述的全部方法均可以应用在图 1 所示的终端中。此外,本领域技术人员可以理解,图 1 仅仅是便携式多功能装置的举例,并不构成对便携式多功能装置的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件。

- 15 参阅图 2 所示,本发明实施例提供一种物体建模运动方法,该方法应用于一种移动终端,该移动终端包括彩色相机和深度相机;且所述彩色相机和所述深度相机位于所述移动终端的同一侧;该方法包括如下步骤:

步骤 21:利用彩色相机和深度相机对目标物体(即被扫描物体,一些段落中简称物体)进行全景扫描,得到目标物体的 3D 模型;

- 20 步骤 22:获取目标骨骼模型;

步骤 23:将目标骨骼模型与目标物体的 3D 模型融合;

步骤 24:获取目标运动方式;

步骤 25:根据目标运动方式控制骨骼模型,使目标物体的 3D 模型根据目标运动方式进行运动。

- 25 其中,上述彩色相机和深度相机可以位于终端设备的前面,也可以位于终端设备的背面,它们具体的排布方式与数量可以根据设计者的需求灵活确定,本申请不做限定。

- 请参阅图 3,图 3 是对物体扫描到实现动画的主要过程。首选对物体进行扫描,通过深度相机扫描得到深度图,通过彩色相机扫描得到彩色图;根据深度图和彩色图进行融合得到带有纹理的网格模型,即物体的 3D 模型;将 3D 模型嵌入骨骼模型,根据骨骼动画使骨骼模型动起来(应理解,骨骼的运动通常不可见,当然一些特殊场景需求下,也可以对用户可见),视觉上呈现出物体的动画效果。下面将结合实例具体展开描述。
- 30

步骤 21

- 35 步骤 21 涉及到深度相机扫描、彩色相机扫描和 3D 重建;具体实例如下。

深度相机扫描

深度相机既可以包含 3D/深度传感器或 3D/深度传感器的模组,用于获取静态物体的深度信息,应理解,扫描的对象理论上应该是静止的物体,实际操作过程中,微小的动态

在一定程度上也是可以接受的。获取深度信息可以采用结构光技术和 TOF。随着深度信息获取方法的新兴，深度模组还可以包含更多的实现方式，本发明中不予以限定。

一种结构光原理图见图 4 所示，其中，301 为不可见的红外光源，302 为生成一定光图案的光栅，303 为被扫描物体，304 为红外相机，获取到 303 反射回来的光图案，跟期望的光图案对比，通过计算获取目标物体被扫描部分的深度信息。

一种 TOF 深度相机见图 5 所示，其中，311 为目标物体，312 为 TOF 相机中的红外发射端，313 为红外接收端，当 312 发射红外光线（例如但不限于：850nm-1200nm）于目标物体时，目标物体会反射红外线，反射的红外线由 313 接收，313 的传感器（例如但不限于：CMOS 阵列或 CCD 阵列，分辨率在 240*180 以上）会因反射的红外光线而产生一系列的电压差信号；深度计算单元 314 根据这一系列的电压差信号进行计算，最后得到目标物体被扫描部分的深度信息 315。

彩色相机扫描

在扫描目标物体的过程中，会同步调用深度相机和彩色相机，采用一定的校正算法，使得两者对应扫描到的目标物体的图像是一致的。彩色相在扫描时获取图像的方式和普通相机拍照的方式基本一致。本部分不做赘述。

在具体实现过程中，在扫描目标物体时，需要在一定的角度范围内（通常受到深度相机或者彩色相机中视场角的较小者的限制）以及一定的距离内扫描物体，受限于深度信息（例如：深度图）的质量，通常物体距离深度相机（或移动终端）的距离为介于 20cm 到 80cm。

一种具体的扫描方式可以是终端不动，手握目标物体置于深度相机前 30cm 到 70cm 范围，并全方位缓慢转动物体，使所有的扫描图的并集能够构建出完整物体。应注意，握住物体的手要尽量不遮挡住物体表面。

一种具体的扫描方式可以是物体不动，手握终端置于深度相机前 30cm 到 70cm 范围，并对物体进行环绕全景扫描，使所有的扫描图的并集能够构建出完整物体。应注意，握住终端的手要尽量不遮挡住物体表面。

一种具体的扫描方式可以是物体不动，手握终端置于深度相机前 30cm 到 70cm 范围，并对物体相隔预设角度进行扫描，直到所有的扫描图的并集能够构建出完整物体。应注意，握住终端的手要尽量不遮挡住物体表面。

具体地，可以拍摄多次，确保画面拍全，场景信息包括了物体的全貌，不留死角，因此，全景扫描的过程中，会对应出现多帧深度图（深度图序列），每一个深度图都是对应一次扫描时扫描范围内的场景；也会对应出现多帧彩色图（彩色图序列），每一个彩色图都是对应一次扫描时扫描范围内的场景。扫描目标物体时也可能包含了其他物体，但是采用上述适中的距离，可以在后续的 3D 重建过程中将目标物体以外的噪声去除。

一种可能的实现方式中，扫描过程中深度相机的拍摄帧率可以大于等于 25fps，例如 30fps、60fps、120fps。

一种可能的实现方式中，扫描的过程中，终端可以呈现出目标物体的扫描进度，以使用户观察是否已经覆盖了目标物体的全景，并且可以由用户自主选择继续扫描或者终止扫描。

由于深度相机和彩色相机的位置可以为前置或后置，因此也相应地存在前置扫描和后置扫描两种方式。如深度相机位于手机正面的上方时，可以与前置彩色摄像头配合使用，前置扫描可以实现自拍扫描；如深度相机位于手机背面的上方时，可以与后置彩色摄像头配合使用，后置扫描可以有更多目标物体的去选择，进而实现对目标物体的精细、稳定的扫描。应理解，随着折叠屏终端的出现，深度相机和彩色相机的物理位置是可以发生变化的，因此，传统意义上的前置和后置也不应该构成任何物理位置的限定。扫描物体进行 3D 建模时，为了保持画面的一致性，深度相机和彩色相机可以位于同一侧，至于与终端的相对位置和相对方向均不做限定，只要能够实现 3D 重建的摄像头位置组合方式都是可以的。在一种可能的实现方式中，终端还可以调用第三方的拍摄器件，如利用外部连接的拍摄杆、扫描仪、外接摄像头等，可选的，可以利用外部的彩色摄像头，或者，也可以利用外部的深度摄像头，也可以两者都用。

上述彩色相机扫描和深度相机扫描，可以由用户触发扫描功能时开启，触发操作包括定时、触发快门、手势操作、隔空感应、器件操控等。另外，当用户开启相机进行预览时，预览图中可以由系统提示哪些物体适合进行扫描或进行 3D 建模；例如可以用方框等标识预览图中的物体用于提示用户。

此外，上述深度相机和彩色相机涉及到的具体器件参数，与制造工艺和用户需求以及终端的设计约束有关，本发明中不做任何具体限定。

3D 重建（网格化+纹理映射）

如图 6 所示，对物体进行多帧的、360 度全景的扫描后，会得到深度图序列 321 和彩色图序列 322，其中深度相机得到的每一帧都是一个扫描场景的深度图（例如：Depth 图），同时彩色相机得到的每一帧都是一个扫描场景的彩色图（例如：RGB 图）；深度图序列 321 经过网格化处理得到目标物体的网格（Mesh）模型，根据彩色序列图 322 对网格模型进行纹理映射，得到纹理映射后的网格模型 323，即物体的 3D 模型。在一种可能的实现中，也可以根据彩色序列图中的全部帧或某几帧图像进行纹理映射。

1) 网格化

在一种具体实现过程中，参见图 7，一种具体的网格化实现方案如下：

步骤 331：获取目标物体每个扫描场景下的彩色图（包括但不限于 RGB）和深度图（Depth）。Depth Map（深度图）是包含目标物体的表面上的多个点与深度相机的距离有关的信息。其中，Depth Map 类似于灰度图像，只是它的一个像素点值表示的是深度相机距离目标物体的表面上一个点的实际距离。通常彩色图像和 Depth 图像是配准的。

步骤 332：包括但不限于，对深度图双边滤波去噪，对深度图下采样生成不同解析度的图像金字塔，将深度图转换成点云，估算每个顶点的法向量，裁切被扫描物体范围外的点。

步骤 333：在 332 中，采集到目标物体在不同扫描位置的深度图和彩色图序列，要想生成物体模型，需要将采集到的图像序列得到的单帧 3D 点云转换到统一的坐标系下，即

得到在物体在不同扫描位置之间的位姿变换关系，即位姿估计。位姿估计是基于图像序列估计物体的 3D 姿态。姿态估计有基于特征点的配准、基于点云的配准这两条思路。图像序列之间物体切换角度较小时，采用基于点云的精配准，例如迭代最近邻点算法 ICP

- (iterative closest point) 可以对物体姿态进行估计。当图像序列之间物体切换角度较大时，
5 基于物体 3D 特征，还可以对两个位姿之间进行粗配准，作为精配准的初值，这种方式可以支持更快速率的扫描。

如果不存在测量误差，则当前帧的 3D 点全部都在目标物体的 3D 体素模型(volumetric model)表面上，所以为了求解相机姿态(位姿变换关系)，转为求当前帧的 3D 点云与目标物体的 3D 体素模型表面点云之间距离的最小值，目标函数为：

$$10 \quad M_{opt} = \underset{M}{\operatorname{argmin}} \sum_i ((M \cdot s_i - d_i) \cdot n_i)^2$$

其中，M 为相机的位姿变换矩阵， s_i 为当前要计算位姿的那一帧的 3D 点云，并将该点云旋转到体素模型的观测坐标系下， d_i 为模型在观测坐标下的点云， n_i 为模型点云对应的法线，则目标函数表示当前帧点云到体素模型点云所在平面的最小距离平方和。

- 步骤 334: 实现 2D 深度图转换为 3D 信息并融合到统一的 3D 体素模型中。采用 TSDF
15 (Truncated Signed Distance Function, 截断的有符号的距离函数) 算法，融合后体素的值为 SDF (Signed Distance Function, 有符号的距离函数) 值和 Weight (权重) 值，以及可选的色彩值。TSDF 算法目前已经是 3D 点云融合的主流处理算法。权重计算，采用求均值的方法。每融合一次，旧权重值即加一。新值权重为 1，新旧 SDF 值，相应乘以其权重，相加，再除以融合次数(新的权重值)，即得到新的归一化 SDF 值。

- 20 步骤 335: 判断 Roll/Yaw/Pitch 三个方向每隔一定角度(例如但不限于 30、45、60、90 等预设角度)是否有保存右预设数量的关键帧，如已保存的关键帧数小于预设数量(以是否覆盖了目标物体的全景为准)时，需要去继续捕捉场景(彩色图和深度图)，终端会指示用户需要进行更多的扫描。当关键帧的数量足够覆盖目标物体的全景时，提示用户扫描完成，可以结束扫描进入后续步骤。

- 25 步骤 336: 实现实时融合过程中，选择并缓存纹理映射所需要的输入关键帧信息，包括彩色图像、位姿(不同图像之间的位置姿态差异)等信息。根据物体建模需要 360 度扫描的特性，Roll/Yaw/Pitch 方向都各选出预设数量(N)个关键帧，即可完全恢复物体 360 度的纹理。如，通过 ICP 结果，确定输入图像流里每一帧的角度(YAW/Pitch/Roll)，再计算每一帧的清晰度，根据角度和清晰度构建选择策略，选择出关键帧。

- 30 角度策略是不同方向上将 360 度分割出 N 个 360/N 的区域，每个区域一定要有一帧清晰的彩色图像。

清晰度检测原理：一般有梯度法和 sobel 算子评价图像的清晰度。可以选择梯度方法计算清晰度。如下图，对于图像中的某个像素，与其右侧及下侧两个像素一起，如下计算：

$$\begin{aligned}
\det_x &= a(i+1, j) - a(i, j) \\
\det_y &= a(i, j+1) - a(i, j) \\
sum &= \sum_{i=0, j=0}^{i=width-1, j=height-1} (\sqrt{(\det_x^2 + \det_y^2)} + |\det_x| + |\det_y|) \\
blur &= sum / (width * height)
\end{aligned}$$

blur 值越大，则图像越清晰。

- 步骤 337：采用 Marching Cubes 算法实现 3D 点云网格化，生成三角面片。Marching Cubes 算法的主要思路是以体元为单位来寻找三维点云中内容部分与背景部分的边界，在体元抽取三角片来拟合这个边界，简单来讲，可以讲包含体数据内容的体素点称为实点，而其外的背景体素点都称作虚点。这样一个三维点云就是由各种实点和虚点组成的点阵。例如，从单个体元的角度出发，体元的 8 个体素点每个都可能是实点或虚点，那么一个体元一共有 2 的 8 次方即 256 种可能的情况。Marching Cubes 算法的核心思想就是利用这 256 种可以枚举的情况来进行体元内的等值三角面片抽取。体元是在三维图像中由相邻的八个体素点组成的正方体方格，MarchingCubes 算法的 Cube 的语义也可以指这个体元。注意区别体元和体素，体元是 8 个体素构成的方格，而每个体素（除了边界上的之外）都为 8 个体元所共享。

2) 纹理映射

- 在一种具体实现过程中，参见图 8，一种具体的纹理映射实现方案如下：

步骤 341：根据网格模型（三角面片信息）以及关键帧的位姿信息，判断每个关键帧的位姿下所有面片是否可见。输入网格模型的所有三角面片信息以及关键帧的空间坐标，输出所有三角面片在每个关键帧的位姿下是否可见的信息。

- 原理描述：判断某个三角面片在某个位姿下可见还是不可见，则需要判断该位姿与该三角面片的顶点连接成的射线与模型的其他三角面片是否相交，如果有相交的情况，则说明该三角面片被其他面片挡住了，则不可见，否则在该位姿下该三角面片是可见的。

碰撞检测过程涉及到空间中的三角形法向量计算、判断射线与三角形是否相交、判断射线是否与 AABB ((Axis-aligned bounding box)) 包围盒相交以及层级二叉树的构建这几大部分。

- 一种具体示例处理流程如下：

- 1) 取该面片的一个顶点，与当前关键帧图像的相机视点连接得到射线；
- 2) 该射线与二叉层级树由根开始计算遮挡；
- 3) 判断该 BV(Bunding Volume)节点是否为叶子节点，若是叶子节点则跳至 6)；
- 4) 判断射线是否与该 BV 的 AABB 包围盒相交，如果不相交则返回 1)；
- 5) 若射线与该 BV 相交，则取 BV 的两个子节点，返回 3)；
- 6) 如果是叶子节点，先判断射线是否与 AABB 包围盒相交，若相交再判断是否与三角形相交，如果相交该顶点被遮挡；
- 7) 如果该面片中有 1 个或 1 个以上顶点被遮挡，则该面片在当前关键帧图像下不可见。

步骤 342: 根据步骤 341 的结果及网格模型, 采用区域划分和图割 (Graph Cuts) 的方法, 对网格模型上每个面片 (face) 进行标记, 确定其选择哪个关键帧图像 (view) 来生成纹理。面片标记的结果可作为仿射贴图 (Warping) 模块的输入, 用于生成初步的纹理图。

- 5 步骤 343: 将关键帧图像中对应区域的纹理映射到纹理图上, 并对不同关键帧的 patch(接缝处片) 做边界平滑。

由于物体 3D 模型是由多个关键帧图像来生成一张纹理图, 因此在关键帧图像选择后纹理图上会在不同关键帧图像生成接缝间有很多颜色不连续。在纹理不连续处, 每个顶点 Vertex 都可以看成两个顶点: 属于左边面片的 Vleft 和属于右边面片的 Vright。每一个顶点 V 的调整前颜色记为 fv, 通过下面的最小化方程得到每个顶点 V 的颜色矫正值 gv。

$$\arg \min_g [\sum_v (f_{v_{left}} + g_{v_{left}} - (f_{v_{right}} + g_{v_{right}}))^2 + \frac{1}{\lambda} \sum_v (g_{v_i} - g_{v_j})^2]$$

argmin 代表求差异最小值, 上面公式中包含两部分

$$1. \sum_v (f_{v_{left}} + g_{v_{left}} - (f_{v_{right}} + g_{v_{right}}))^2 \quad 2. \frac{1}{\lambda} \sum_v (g_{v_i} - g_{v_j})^2$$

- 15 第一部分中, 这里 v 代表关键帧接缝处的顶点(Vertex), 即从属于左边接缝处片, 也从属于

右边接缝处片。 $f_{v_{left}} + g_{v_{left}}$ 中的 fv 是调整前颜色值, gv 是颜色校正值, 也就是增加量(Δ),

这个公式的意思就是, 为了保证接缝处的平滑, 不同帧的图像的公共点矫正完后的差异要尽可能的小。第二部分中, Vi 和 Vj 表示同一纹理接缝处片上任意两个相邻顶点的增量要尽可能的小, 确保不会一个增加太多一个减少的太少而导致不平滑。

- 20 在一种具体实现过程中, 将面片标记的结果相邻区域相同标签的存为 patch, 对所有 patch 的顶点做边界平滑, 调整每个顶点的像素值, 对最终的顶点围城的三角行做基于位置和像素的仿射变换, 形成最终纹理图。

将物体的纹理图集绘制于物体的网格模型的表面便得到了物体的 3D 模型，一般保存为.obj 格式，如图 9 所示为例，狮子的纹理图集映射到狮子的网格模型上，得到狮子纹理映射后的 3D 模型。

5 步骤 22

通过步骤 21 得到了目标物体经 3D 重建后的 3D 模型，即带有纹理的网格模型。接下来，需要给带有纹理的网格加入骨骼，这就涉及到如何获取骨骼模型，即目标骨骼模型。

10 在一种具体实现过程中，可以提供给用户骨骼模型制作库，如一些线段和点，其中，线段表示骨骼，点表示关节节点。接收用户的操作指令，如手势、滑动、快捷键等，将至少两个线段和至少一个点组合成骨骼模型，从而得到骨骼模型，进一步地，将该骨骼模型上传到云端或存储在本地。

15 在一种具体实现过程中，可以提供给用户更加开放的制作库，线段和点完全由用户自由设计，其中，线段表示骨骼，点表示关节节点。接收用户的操作指令，如手势、滑动、快捷键等，将至少两个线段和至少一个点组合成骨骼模型，从而得到骨骼模型，进一步地，将该骨骼模型上传到云端或存储在本地。

20 在一种具体实现过程中，可以从至少一个预设骨骼模型中选择与所述目标物体的外形匹配度最高的骨骼模型作为目标骨骼模型。预设骨骼模型可以存储在网络、云端或者本地。例如本地存储了鸡的骨骼模型、狗的骨骼模型、鱼的骨骼模型，当目标物体为鸭子的时候，系统通过外形识别，将鸡的骨骼模型作为目标骨骼模型。相似的判定标准包括但不限于骨骼形态、骨骼的长度、骨骼的粗细、骨骼的数量、骨骼的组成方式等。

在一种具体实现过程中，可以接收用户的选择指令，该选择指令用于从至少一个预设骨骼模型中选择出目标骨骼模型。

25 步骤 23

将目标骨骼模型与目标物体的 3D 模型进行融合，或者说将目标骨骼模型嵌入到目标物体的 3D 模型中。需要一定计算物体/角色中骨骼关节节点的位置，使得最终结果的骨架尽可能好的符合目标物体的内部构造，并且同预设的(给定的)骨架看起来尽可能相像。

30 在一种具体实现过程中，参见图 10，一种具体的骨骼装配方案如下：

步骤 351：为了近似中轴面和利用其它的计算，计算三线性差值的自适应采样距离场 (adaptively Sampled Distance Fields)。可以通过构建 kd 树 (K-dimension tree) 来评估从任意点到物体表面的有符号的距离大小，kd 树是一种分割 k 维数据空间的数据结构。主要应用于多维空间关键数据的搜索。

35 步骤 352：计算近似位于物体中轴面的一组样本点，找到骨骼关节可能位于的点，滤除那些接近物体表面的点。

步骤 353：为了从中轴面中挑选出骨架图的顶点，可以在物体中填充球体，对所有中轴线上的点，按照距离 3D 模型表面的远近进行排序，从最远的点开始画出 3D 模型内部最

大的内切球体(不超出 3D 模型表面范围),得到球体半径;之后遍历中轴线上的每个点,只有当该点不被包含在之前任意一个填充的球体中时,才会对该点构建内切球体。

步骤 354: 可以通过连接一些球心来构建骨架图,连接任意两个球体相交的球的球心作为一条边。

5 以上步骤 351-步骤 354 可以称为骨骼识别。

步骤 355: 提取预设的骨骼模型,需要以一种最优化的方式嵌入步骤 354 构建的几何体骨架图 $G=(V,E)$ (V 代表顶点, E 代表边缘) 中。通常节点需要减少,骨架需要优化。

步骤 356: 识别骨骼层级关系,减少简单的层级而来近似骨骼形状。

经过骨骼适配(步骤 355+356)后,可以得到装配好骨骼模型的物体 3D 模型。

10

步骤 24

通过步骤 23 得到了嵌入骨骼模型后的 3D 模型。接下来,需要给骨骼模型寻找到一些可实现运动方式,进而实现目标物体 3D 模型的运动,这就涉及到如何获取骨骼的运动方式(也可形象地成为动画),即目标运动方式。

15

在一种具体实现过程中,可以获取第一物体的运动方式,将第一物体的运动方式作为目标运动方式。其中,第一物体可以是当前实时运动的物体(例如,拍摄一个人正在跑步,通过神经网络提取出人物的骨骼的运动方式);也可以是过去拍过且保存下来物体的运动方式(例如,之前拍摄过一只狗的一组可爱的动作,并通过算法将该动作的运动方式存储在本地或者云端);也可以是预设的某一特定物体的运动方式。(例如,只选择跟人类的有关的运动方式)

20

在一种具体实现过程中,可以从预设的目标运动方式选出一种。(例如,本地存储有人的动作、狗的动作、猫的动作、马的动作等,用户可以根据自己的喜好或者物体类型的符合度选择自己喜欢的特定类别)。

25

在一种具体实现过程中,运动方式可以是用户自己使用动画制作软件制作的,当然这个软件可以是系统内嵌的工具集或者在扫描重建运动的 APP 中载入的工具集,或者来源于第三方动画设计工具;可以是当前制作也可以是历史制作的运动方式或动画。

在一种具体实现过程中,运动方式可以根据物理属性在多个预先存储的运动方式中选择出属性匹配度最高的运动方式作为目标运动方式。例如,本地预先存储了鱼的游水动画,青蛙起跳的动画,马奔跑的动画;如果用户扫描的目标物体为一只鹿,则将马奔跑的动画作为鹿的目标运动方式。(鹿和马在外形、生物种类、骨骼结构等属性上比其鱼和青蛙相似度更高)

30

在一种具体实现过程中,运动方式还可以是基于目标物体的骨骼模型(可以采用上一步骤中的任一方法去获取)由系统或者用户对该骨骼模型进行自主设计,得到目标运动方式。这一种方式是对后续物体的 3D 模型实现动画的最契合的动画操作。

35

在一种具体实现过程中,运动方式可以是预设的骨骼动画,一般由专业的动画设计师制作而成。

应理解,一个骨骼动画描述了骨架中每一个节点变换随着时间改变的动态,通常使用关键帧的形式来进行储存和表达。通常具有 FPS(Frame Per Second)概念,即一秒钟包含多

少帧。骨骼动画不能脱离骨架而存在，否则无法驱动 3D 模型，因此骨骼动画通常是依赖于特定的骨架而存在的，骨架通常称为 Rig，其描述了一套骨架有哪些骨骼，每根骨骼的衔接关系，每根骨骼默认的变换（也就位姿），以及其他的一些额外信息。位姿描述了一个骨架中每一个节点的变换后的一个静态，诸如站立、奔跑的一帧。每个骨架都会储存一个绑定位姿，就是制作这个骨架时候的默认姿态。姿态一般不储存骨架的层次关系，而是用一个数组来依次储存每一个节点的变换，而节点从属于特定的某跟骨头，因而无法脱离骨架单独使用。另外，姿态是骨骼动画采样结果的一部分。可见，骨架、位姿和骨骼动画是互相关联的，共同实现后续的动画操作的。

10 步骤 25

骨骼动画本质上是记录了以树形结构存储的一系列对象的位置、旋转、缩放随着时间变化的动态，其中的每一个对象就是一根骨骼。动画的实现过程为，一组动画中的骨骼动画变换映射到上一个步骤“自动骨骼装配”装配好了骨骼的 3D 模型中，这种动作的映射的实现方式包括但不限于游戏引擎、动画引擎等；3D 模型根据骨骼变换进行姿态变换，连贯起来就是一系列的动画动作，用户感官上的感受就是扫描的物体“复活”了，静态的东西“动”起来了。在动画的实现过程中，蒙皮（skinning）技术是保证物体 3D 模型跟随骨骼模型运动的基础。

物体的 3D 模型的动画表现为装配了骨骼的 3D 模型映射到一组变化的骨骼模型的动作上，每一帧都需要根据骨骼的变化，实现 3D 模型表面（即物体的 3D 模型表皮）的形变，这个过程称为蒙皮。从而实现了 3D 模型到动作的映射，实现了动画效果。

在具体实现过程中，可以采用线性融合蒙皮(LBS)方案，对于 3D 模型中表面的任意一点，可以采用如下公式根据前一个状态的位置得到下一个状态的位置。其中， v_i 为上一个位置， v'_i 为下一个位置， $w_{i,j}$ 为第 j 根骨头在 i 点上的权重， T_j 为变换矩阵。3D 模型表面上达到一定数量的点都能计算出下一个位置，则 3D 模型在下一个位置的姿态便能够确定出来，进而实现动画。蒙皮技术的核心为找到每根骨头对每一个顶点的权重 $w_{i,j}$ 。

$$v'_i = \sum_{j=1}^m w_{i,j} T_j v_i$$

一种具体实现过程中，可以使用类似热平衡的方式来计算权重，将一个 3D 模型看成是一个绝缘的导热体，并将第 i 跟骨头温度设为 1° ，同时将其余的骨头的温度设定为 0° 。根据热平衡原理，这样我们就可以将表面顶点平衡之后的温度设定为该点的权重，权重值区间为 $[0, 1]$ 。基于热平衡的权重计算方法使得权重的计算结果具有平滑特性，呈现的动作效果会更加真实自然。

应理解，通过改变嵌入骨骼的位置（即通过动画）实现对物体 3D 模型的动作变换，最终用户看到的是蒙皮的效果。如果特别设定，也可以让用户看到只有骨骼而没有物体 3D 模型的动画。

一种具体的动画流程实现可参见图 11 所示,通过预设的骨骼图和多帧动作,即根据运动模型或动画模型对装配了目标骨骼模型的 3D 模型进行动作映射,控制目标骨骼模型使目标骨骼模型按照预设的运动模型进行运动,在运动的过程中实时计算和更新蒙皮数据,使得 3D 模型能够跟随目标骨骼模型实现平滑的运动,进而实现 3D 模型的动画。应理解,目标物体的骨骼模型和动画的骨骼结构可能并不完全一样,可以将物体的骨骼模型和动画的骨骼结构进行位置映射,比如关键节点要一致,至于骨骼长短可以按照比例设置;还可以将物体的骨骼模型和动画的骨骼结构互相适配,比如进行成比例的裁剪和延长,至少使得动画的骨骼结构不超出物体的 3D 模型的外表面,进一步地,可以经过一些物理运算,使得修整动画骨骼,使其最大限度地支撑起物体的 3D 模型,使得动画的骨骼模型和物体的 3D 模型契合地更和谐。

上述步骤 21-25 既可以一气呵成地逐步完成,也可以具有一定的时间间隔。例如,用户扫描完物体的 3D 模型之后,可以将 3D 模型存储在本地或者云端,若干时间以后可以直接调用该 3D 模型,自由选择骨骼装配或自由选择动画方式,还可以选择动画背景,包括但不限于实时拍摄的图像、本地已经存储的其它图像、云端数据图像等。此外,在终端显示物体动画的同时,还可以显示出物体的影子,或者增加其它的音效、特效等等。动画可以由移动终端自动播放,还可以在通过用户输入操作指令进行播放的控制。

应理解,上述实施例仅是本发明中一些可选的实施方式;并且由于相机参数设计、算法实现方式、用户设置、终端操作系统、终端所处环境、以及用户使用习惯的不同,前文所提及的器件参数、用户使用方法、实施例中涉及的算法均存在多种变形。无法通过穷举而一一列出,本领域技术人员应当理解,基于上述理论进行适应性地调整,包括一些常规方式的替代而产生的技术方案都应属于本发明保护范围内。

通过本发明,可以在移动终端上就能一体实现目标物体从扫描、3D 重建、到骨骼装配,再到预设的动画展示等一系列的操作,对于用户来说进行 3D 扫描能够轻松玩转;并且随着移动终端的拍摄技术的广泛应用,呈现二维图像可向呈现 3D 动画过渡,允许用户将现实扫描建模的物体最终实现虚拟的动画动作;大大提升了用户使用移动终端的趣味性和粘性,引领拍摄应用进入一个新的潮流。

基于上述实施例提供的物体建模运动方法,本发明实施例提供一种物体建模运动装置 700,所述装置 700 可以应用于各类拍照设备,如图 12 所示,该装置 700 包括扫描模块 701、第一获取模块 702、融合模块 703、第二获取模块 704 以及运动模块 705;该装置应用于移动终端,移动终端包括彩色相机和深度相机,彩色相机和深度相机位于移动终端的同一侧;相关特性可以参照前述方法实施例中的描述。

扫描模块 701,用于当彩色相机和深度相机对目标物体进行全景扫描时,得到目标物体的 3D 模型。该扫描模块 701 可以由处理器调用存储器中的程序指令对上述彩色相机和深度相机使能控制,进一步地,扫描时的采集的图片还可以选择性地存储在存储器中。

第一获取模块 702,用于获取目标骨骼模型。该第一获取模块 702 可以由处理器调用相应的程序指令实现,进一步地,可以通过调用本地存储器或云端服务器中的数据以及算法,进行相应计算实现。

融合模块 703, 用于将目标骨骼模型与目标物体的 3D 模型融合; 该融合模块 703 可以由处理器调用相应的程序指令实现, 进一步地, 可以通过调用本地存储器或云端服务器中的数据以及算法, 进行相应计算实现。

- 5 第二获取模块 704, 用于获取目标运动方式。该第二获取模块 704 可以由处理器调用相应的程序指令实现, 进一步地, 可以通过调用本地存储器或云端服务器中的数据以及算法, 进行相应计算实现。

运动模块 705, 用于根据目标运动方式控制骨骼模型, 使目标物体的 3D 模型根据目标运动方式进行运动。该运动模块 705 可以由处理器调用相应的程序指令实现, 进一步地, 也可以通过调用本地存储器或云端服务器中的数据以及算法实现。

- 10 在具体实现过程中, 扫描模块 701 具体用于执行步骤 21 中所提到的方法以及可以等同替换的方法; 第一获取模块 702 具体用于执行步骤 22 中所提到的方法以及可以等同替换的方法; 融合模块 703 具体用于执行步骤 23 中所提到的方法以及可以等同替换的方法; 第二获取模块 704 具体用于执行步骤 24 中所提到的方法以及可以等同替换的方法; 运动模块 705 具体用于执行步骤 25 中所提到的方法以及可以等同替换的方法。

- 15 更为具体地;

扫描模块 701 可以执行上述步骤 331-337, 步骤 341-343 的方法; 融合模块 703 可以执行上述步骤 351-356 的方法。

其中, 上述具体的方法实施例以及实施例中技术特征的解释、表述、以及多种实现形式的扩展也适用于装置中的方法执行, 装置实施例中不予以赘述。

- 20 本发明实施例提供的装置 700, 可以实现了物体从扫描、3D 重建、到骨骼装配, 再到预设的动画展示的一体化设计。无需用户采用专业的、笨重的、复杂的设备进行专业扫描, 也无需再跑到 PC 端做复杂的建模处理和动画处理, 将这些功能集成在一起, 提供给用户, 使得用户在一个移动终端上能够轻松玩转这一系列的操作方法, 使得用户身边任何一个“静态的物体(或近似静态的物体)”能够更鲜活, 更具生命力。增加用户使用终端的趣味性,
- 25 提升用户的体验度。

- 应理解以上装置 700 中的各个模块的划分仅仅是一种逻辑功能的划分, 实际实现时可以全部或部分集成到一个物理实体上, 也可以物理上分开。例如, 以上各个模块可以为单独设立的处理元件, 也可以集成在终端的某一个芯片中实现, 此外, 也可以以程序代码的形式存储于控制器的存储元件中, 由处理器的某一个处理元件调用并执行以上各个模块的功能。此外各个模块可以集成在一起, 也可以独立实现。这里所述的处理元件可以是一种集成电路芯片, 具有信号的处理能力。在实现过程中, 上述方法的各步骤或以上各个模块可以通过处理器元件中的硬件的集成逻辑电路或者软件形式的指令完成。该处理元件可以是通用处理器, 例如中央处理器(英文: central processing unit, 简称: CPU), 还可以是被配置成实施以上方法的一个或多个集成电路, 例如: 一个或多个特定集成电路(英文: application-specific integrated circuit, 简称: ASIC), 或, 一个或多个微处理器(英文: digital signal processor, 简称: DSP), 或, 一个或者多个现场可编程门阵列(英文: field-programmable gate array, 简称: FPGA)等。
- 30
- 35

本领域内的技术人员应明白, 本发明的实施例可提供为方法、系统、或计算机程序产品。因此, 本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实

实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

5 本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的
10 装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

15 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

20 尽管已描述了本发明的部分实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例作出另外的变更和修改。所以，所附权利要求意欲解释为包括已列举实施例以及落入本发明范围的所有变更和修改。显然，本领域的技术人员可以对本发明实施例进行各种改动和变型而不脱离本发明实施例的精神和范围。倘若本发明实施例的这些修改和变型属于本发明权利要求及其等同技术的范围之内，则本发明也包含这些改动和变型在内。

25