

图像处理方法、装置、电子设备及计算机可读存储介质

- 5 本公开要求于 2019 年 12 月 18 日提交中国专利局、申请号为 201911306421.6、申请名称为“图像处理方法、装置、电子设备及计算机可读存储介质”的中国专利申请
- 的优先权，其全部内容通过引用结合在本公开中。

技术领域

- 10 本公开涉及图像处理领域，尤其涉及一种图像处理方法、装置、电子设备及计算机可读存储介质。

背景技术

- 15 随着计算机网络的发展，和智能手机的普及，普通的用户已经不能满足于仅仅利用单调的图片和文字来表达自己的情感。视频以呈现出更为丰富多样的内容和形式及带来的直观感受深受用户的喜爱，并且逐渐流行开来，普通用户制作原创视频逐渐成为一个趋势。但是另外一方面是原创自拍类视频表现形式平淡无味，与此同时，我们可以看到视频特效在影视作品中的运用越来越丰富，在内容表现形式方面也更加的多样化，可以说视频特效是一个成功的影视作品的支撑和保障。
- 20 然而现有技术中制作特效一般是直接将特效覆盖于目标对象（如人脸）上，被覆盖的地方被特效遮住，无法体现出目标对象原有的特征，导致特效的实际效果比较生硬，缺乏真实感。

发明内容

- 25 提供该发明内容部分以便以简要的形式介绍构思，这些构思将在后面的具体实施方式部分被详细描述。该发明内容部分并不旨在标识要求保护的技术方案的关键特征或必要特征，也不旨在用于限制所要求的保护的技术方案的范围。
- 第一方面，本公开实施例提供一种图像处理方法，包括：
- 识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象；
- 30 按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像；其中，在所述第二视频帧图像中所述第三对象覆盖所述第一对象；
- 根据所述第二对象在所述第一视频帧图像中的位置，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。
- 35 第二方面，本公开实施例提供一种图像处理装置，包括：
- 对象识别模块，用于识别第一视频帧图像中的第一对象和位于所述第一对象内的

第二对象；

第二视频帧图像生成模块，用于按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像；其中，在所述第二视频帧图像中所述第三对象覆盖所述第一对象；

5 第三视频帧生成模块，用于根据所述第二对象在所述第一视频帧图像中的位置，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

第三方面，本公开实施例提供一种电子设备，包括：至少一个处理器；以及，与所述至少一个处理器通信连接的存储器；其中，所述存储器存储有能被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行前述第一方面中的任一所述图像处理方法。

第四方面，本公开实施例提供一种非暂态计算机可读存储介质，其特征在于，该非暂态计算机可读存储介质存储计算机指令，该计算机指令用于使计算机执行前述第一方面中的任一所述图像处理方法。

15 第五方面，本公开实施例提供了一种计算机程序，当所述计算机程序被计算机运行时，所述计算机执行前述第一方面中的任一所述图像处理方法。

本公开实施例公开了一种图像处理方法、装置、电子设备和计算机可读存储介质。其中该图像处理方法包括：识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象；按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像；其中，在所述第二视频帧图像中所述第三对象覆盖所述第一对象；根据所述第二对象在所述第一视频帧图像中的位置，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。通过上述方法，解决了现有技术中覆盖在对象上的特效无法体现出对象原有的特征所导致的缺乏真实感的技术问题。

25 上述说明仅是本公开技术方案的概述，为了能更清楚了解本公开的技术手段，而可依照说明书的内容予以实施，并且为了让本公开的上述和其他目的、特征和优点能够更明显易懂，以下特举较佳实施例，并配合附图，详细说明如下。

附图说明

30 结合附图并参考以下具体实施方式，本公开各实施例的上述和其他特征、优点及方面将变得更加明显。贯穿附图中，相同或相似的附图标记表示相同或相似的元素。应当理解附图是示意性的，原件和元素不一定按照比例绘制。

- 图 1 为本公开提供的图像处理方法实施例的流程图；
- 图 2 为本公开提供的图像处理方法实施例中计算人脸的定位点的示意图；
- 35 图 3 为本公开提供的图像处理方法实施例中的步骤 S103 的具体实例流程图；
- 图 4 为本公开提供的图像处理方法实施例步骤 S302 的具体实例流程图；
- 图 5 为本公开实施例提供的图像处理装置的实施例的结构示意图；
- 图 6 为根据本公开实施例提供的电子设备的结构示意图。

具体实施方式

下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例，然而应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为限于这里阐述的实施例，相反提供这些实施例是为了更加透彻和完整地理解本公开。

应当理解的是，本公开的附图及实施例仅用于示例性作用，并非用于限制本公开的保护范围。

应当理解，本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行，和/或并行执行。此外，方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。

本文使用的术语“包括”及其变形是开放性包括，即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施例”。其他术语的相关定义将在下文描述中给出。

需要注意，本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

需要注意，本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的，本领域技术人员应当理解，除非在上下文另有明确指出，否则应该理解为“一个或多个”。

本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的，而并不是用于对这些消息或信息的范围进行限制。

图1为本公开实施例提供的图像处理方法的流程图，本实施例提供的该图像处理装置可以由一图像处理装置来执行，该图像处理装置可以实现为软件，或者实现为软件和硬件的组合，该图像处理装置可以集成设置在图像处理系统中的某设备中，比如图像处理服务器或者图像处理终端设备中。如图1所示，该方法包括如下步骤：

步骤S101，识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象；

在该步骤中，所述第一视频帧图像是通过图像传感器或者存储器获取到的。其中所述图像传感器指可以采集图像的各种设备，典型的图像传感器为摄像机、摄像头、相机等。在该实施例中，所述图像传感器可以是终端设备上的摄像头，比如智能手机上的前置或者后置摄像头，摄像头采集的图像可以直接显示在手机的显示屏上；其中所述存储器为本地存储器或者网络存储器，通过本地存储器的存储位置或者网络存储器的存储地址来获取视频帧图像，在存储器中的视频帧图像为预先拍摄好的视频帧图像文件，其可以通过终端设备中的播放器显示在终端设备的显示装置上。

所述第一视频帧图像中包括第一对象，所述第一对象可以是第一视频帧图像中的任意物体。示例性的，所述第一对象为人脸；所述第二对象为第一对象的一部分，位于第一对象内，示例性的，当所述第一对象为人脸时，所述第二对象为人脸上的眼睛、嘴巴等五官。

在该实施例中，使用对象的关键点识别对象；示例性的，可以根据预设的多个关键点来识别对象，如果所述多个关键点被识别出来，则所述对象被识别出来。可选的，

所述识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象，包括：在所述第一视频帧图像中识别所述第一对象的多个第一关键点和所述第二对象的多个第二关键点；根据所述第一对象的多个第一关键点识别所述第一对象；根据所述第二对象的多个第二关键点识别所述第二对象，其中所述多个第二关键点为所述第二对象的边缘关键点。其中所述边缘关键点确定出所述第二对象的轮廓。示例性的，所述第一对象为人脸，所述第二对象为人脸中的眼睛和嘴巴。所述多个第一关键可以为人脸上的两个眼睛的中心关键点和鼻尖关键点，当所述两个眼睛的中心关键点和鼻尖关键点被识别出来，则所述人脸被识别出来；所述多个第二关键点为所述眼睛的边缘关键点和嘴巴的边缘关键点，同样的，当所述眼睛的边缘关键点和嘴巴的边缘关键点，所述眼睛和嘴巴被识别出来。示例性的，所述第一关键点和第二关键点可以使用深度学习算法来识别，可以直接使用标记有上述第一关键点和第二关键点的帧图像集合训练深度学习网络，得到一个可以回归出第一关键点和第二关键点的网络模型，之后将第一视频帧图像输入所述网络模型中，如果所述第一视频帧图像中包括第一对象和第二对象，则所述网络模型会输出第一关键点和第二关键点的位置，由此直接识别出第一对象和第二对象，并且识别出第一关键点和第二关键点的位置。可以理解的，所述第一关键点和第二关键点也可以使用其他任何关键点识别算法来识别，在此不再赘述。

步骤 S102，按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像；其中，在所述第二视频帧图像中所述第三对象覆盖所述第一对象；

可选的，所述第三对象为三维虚拟对象，且所述三维虚拟对象在所述第一视频帧图像中的二维投影的区域大于所述第一对象的区域，即所述第三对象可以完全覆盖所述第一对象。示例性的，所述三维虚拟对象可以是三维的橘子、三维的球、三维的西瓜等圆球形的虚拟对象。所述第三对象包括第三对象的定位点，用于定位所述第三对象的位置。

可以理解的，所述第三对象可以是任何对象，并不局限于上述三维虚拟对象，如二维虚拟对象以及真实的对象，所述真实的对象是真实物体被拍摄到的图像中形成的对象，如在第一视频帧图像中包括人脸图像以及动物脸图像，此时动物脸图像可以作为第三对象覆盖到所述人脸图像上。

在该步骤中，在将第三对象覆盖到所述第一对象上时，可以使用单个关键点对齐的方法来确定第三对象的覆盖位置。示例性的，第三对象上包括第三对象的定位点，在所述第一对象上预选设置与所述第三对象的定位点对应的第一对象的定位关键点，所述第一对象的定位关键点可以是上述第一关键点中的一个。在将第三对象作为前景图像覆盖到所述第一视频帧图像上时，将所述第三对象的定位点与第一对象的定位关键点的位置进行对齐，这样就能够使得所述第三对象覆盖到所述第一对象上。

有时候，使用单个关键点对齐不太准确，因此可选的，所述按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像，包括：根据所述第一视频帧图像中所述第一对象的多个第一关键点计算所述第一视频帧图像中所述第一对象的定位点；通过对齐所述第三对象的定位点与所述第一视频帧图像中所述第一对象的定位点，将所述第三对象作为前景

图像覆盖到所述第一视频帧图像中，得到所述第二视频帧图像。在该可选实施例中，所述第一对象的定位点由多个第一关键点计算得到，这样由于所述第一对象的定位点中包括了多个第一关键点的位置信息，因此比单个第一关键点定位更加准确。示例性的，如图 2 所示，所述第一对象为人脸，所述多个第一关键点为左右两个眼睛的中心关键点 201 和 202、鼻尖关键点 203，所述人脸的定位点可以通过两个眼睛的中心关键点和鼻尖关键点各占 1/3 权重的方式计算出来，这样计算出来的是两个眼睛的中心关键点和鼻尖关键点所组成的三角形的中心点 204，这样单个关键点的位置的畸变对定位点的位置的影响要小很多，第一对象的定位点更加稳定准确。

在得到第一对象的定位点之后，将所述第三对象的定位点设置到所述第一对象的定位点的位置上，将第三对象作为前景来渲染第一视频帧图像以得到第二视频帧图像。

步骤 S103，根据所述第二对象在所述第一视频帧图像中的位置，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

在步骤 S101 中已经识别出第二对象，其中所述第二对象在所述第一视频帧对象中的位置可以由所述第二对象的中心点的位置确定或者由所述多个第二关键点确定，示例性的，在识别所述第二对象时，将所述第二对象的定位点作为一个关键点直接识别出来或者根据所述多个第二关键点计算出所述第二对象的定位点，在计算所述第二对象的定位点时，可以为每个参与计算的第二关键点设置权重值，所述第二对象的定位点的位置为所述多个第二关键点的位置的加权平均值。在得到所述第二对象在所述第一视频帧图像中的位置之后，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象的对应位置上。

可选的，当所述第三对象为三维对象时，由于三维对象具有一定的深度（Z 轴方向的信息），而第二对象为二维对象不具有深度，因此将第二对象叠加到第三对象上时，会有一定的偏差，所以第二对象在叠加到第三对象上时需要有一定的偏移量，才能使叠加的结果比较自然。此时，如图 3 所示，步骤 S103 可以包括：

步骤 S301，计算所述第一对象上预设的两个关键点之间的第一距离；

步骤 S302，根据所述第一距离计算所述第二对象的偏移量；

步骤 S303，根据所述第二对象在所述第一视频帧图像中的位置以及所述第二对象的偏移量，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

所述第一对象上预设的两个关键点可以是第一对象上的任意两个关键点，示例性的是，所述预设的两个关键点为第一对象上左右对称的两个关键点，如上述实施例中第一对象为人脸，则所述预设的两个关键点可以为左右两个眼睛的中心关键点。示例性的，所述根据所述第一距离计算所述第二对象的偏移量可以是第一距离乘以一个偏移系数，所述偏移系数与所述第一距离的值相关，如第一距离的值越大，偏移系数越大，这样第二对象的偏移量越大，带来的效果是第一对象距离镜头越近，所述第二对象就偏移的越多，以适应第三对象的三维表面。

上述偏移量计算是近似的计算，可选的，可以更加精确的计算所述偏移量，此时，如图 4 所示，所述步骤 S302 可以包括：

步骤 S401，获取所述第一对象的偏航角以及俯仰角，其中所述偏航角为所述第一

对象在所述第一视频帧图像中的朝向与镜头拍摄方向的水平夹角, 所述俯仰角为第一对象在所述第一视频帧图像中的朝向与镜头拍摄方向的竖直夹角;

步骤 S402, 根据所述第一距离、偏航角以及俯仰角计算所述第二对象的偏移量。

具体的, 上述步骤 S402 中的计算过程可以使用以下公式表示:

$$\begin{aligned} dx &= -\text{sign}(\text{yaw}) \times \sqrt{2} \times \text{dist}_{\text{eye}} \times \cos(\theta_1) \times \sqrt{1 - \cos(\text{yaw})} \\ dy &= -\text{sign}(\text{pitch}) \times \sqrt{2} \times \text{dist}_{\text{eye}} \times \cos(\theta_2) \times \sqrt{1 - \cos(\text{pitch})} \end{aligned}$$

- 5 其中, 其中, sign 为取符号操作, dx 为所述第二对象在 X 轴方向的偏移量, dy 为所述第二对象在 Y 轴方向的偏移量, dist_{eye} 为第一距离, yaw 为第一对象的偏航角, pitch 为第一对象俯仰角, θ_1 为第二对象正向前方时的初始相对俯仰角, θ_2 为第二对象正向前方时的初始相对偏航角。其中 θ_1 和 θ_2 为预设的角度值, 由于三维的第三对象的外表面具有一定的弧度, 因此在移动时, 需要有一些衰减。以第一对象为人脸, 第二对象为眼睛和嘴巴为例, 由于嘴巴位于第三对象中的上下区域的下半个区域, 因此嘴巴初始的相对俯仰角较大, 由于嘴巴位于第三对象的左右区域的中间, 因此嘴巴初始的相对偏航角为 0。其他第二对象的 θ_1 和 θ_2 可以根据第三对象的不同而预先设置不同的值, 在此不再赘述。
- 10

- 可选的, 所述步骤 S303, 根据所述第二对象在所述第一视频帧图像中的位置以及所述第二对象的偏移量, 将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上, 得到第三视频帧图像, 还包括: 按照所述第二对象的偏移量对所述第二对象在所述第一视频帧图像中的初始定位点进行位置偏移, 得到目标定位点; 在所述第二视频帧图像中, 将所述第二对象作为前景图像渲染到所述目标定位点的位置, 得到所述第三视频帧图像。在上述步骤中, 将所述第二对象在第一视频帧图像中位置成为初始定位点, 所述初始定位点包括 X 轴的坐标值和 Y 轴的坐标值, 将所述初始定位点的 X 轴坐标值以及 Y 轴坐标值分别与所述偏移量的 X 轴坐标值和 Y 轴坐标值相加, 得到目标定位点; 之后在第二视频帧图像中, 将所述第二对象作为前景渲染到所述第二目标定位点的位置上, 得到第三视频帧图像, 在所述第三视频帧图像上, 所述第三对象上叠加了第二对象, 可以真实的反应第一对象的特征, 由此可以使特效更具真实感。
- 20
- 25

- 可选的, 由于第三对象与第二对象的大小可能并不相同, 因此在将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上之前, 还包括: 根据第三对象将所述第二对象进行缩放。此时, 可以根据第三对象上预定两关键点之间的距离与第一对象上预定两关键点之间的距离的比值来确定第二对象的缩放比例, 这样当所述第三对象大于所述第一对象时, 对所述第二对象进行放大, 当所述第三对象小于所述第一对象时, 将所述第二对象进行缩小。在该步骤之后叠加到所述第二视频帧图像的第三对象上的第二对象即为进行缩放之后的第二对象。
- 30

- 进一步的, 当所述目标对象覆盖在所述三维虚拟对象上之后, 由于目标对象的颜色和三维虚拟对象的颜色可能存在差别, 可以通过进一步的颜色处理来消除这种差别, 以使得第二对象与第三对象的叠加效果更加自然。进一步, 在所述步骤 S103 之后, 还包括:
- 35

将所述第三视频帧图像中的第二对象以及第三对象的颜色空间从 RGB

(Red,Green,Blue; 红, 绿, 蓝)颜色空间转换为 HSV HSV(Hue, Saturation, Value;色调, 饱和度, 明度)颜色空间;

将所述第二对象在 HSV 颜色空间中的 H 通道的值替换为所述第三对象的 H 通道的值

- 5 将所述第二对象以及第三对象的颜色空间从 HSV 空间转换为 RGB 颜色空间得到第四视频帧图像。

10 通过上述步骤, 将 RGB 空间转换到 HSV 空间, 在保存第二对象的原始饱和度和亮度的情况下, 将第二对象的颜色转换为第三对象在所述目标位置上的颜色, 这样使得第二对象与第三对象融合度更好, 使得第二对象更像是第三对象的一部分, 更加逼真。

本公开实施例公开了一种图像处理方法、装置、电子设备和计算机可读存储介质。其中该图像处理方法包括: 识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象; 按照所述第一对象在所述第一视频帧图像中的位置, 将第三对象作为前景图像覆盖到所述第一视频帧图像上, 得到第二视频帧图像; 其中, 在所述第二视频帧图像中所述第三对象覆盖所述第一对象; 根据所述第二对象在所述第一视频帧图像中的位置, 将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上, 得到第三视频帧图像。通过上述方法, 解决了现有技术中覆盖在对象上的特效无法体现出对象原有的特征所导致的缺乏真实感的技术问题。

20 在上文中, 虽然按照上述的顺序描述了上述方法实施例中的各个步骤, 本领域技术人员应清楚, 本公开实施例中的步骤并不必然按照上述顺序执行, 其也可以倒序、并行、交叉等其他顺序执行, 而且, 在上述步骤的基础上, 本领域技术人员也可以再加入其他步骤, 这些明显变型或等同替换的方式也应包含在本公开的保护范围之内, 在此不再赘述。

25 图 5 为本公开实施例提供的图像处理装置实施例的结构示意图, 如图 5 所示, 该装置 500 包括: 对象识别模块 501、第二视频帧图像生成模块 502 和第三视频帧生成模块 503。其中,

对象识别模块 501, 用于识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象;

30 第二视频帧图像生成模块 502, 用于按照所述第一对象在所述第一视频帧图像中的位置, 将第三对象作为前景图像覆盖到所述第一视频帧图像上, 得到第二视频帧图像; 其中, 在所述第二视频帧图像中所述第三对象覆盖所述第一对象;

第三视频帧生成模块 503, 用于根据所述第二对象在所述第一视频帧图像中的位置, 将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上, 得到第三视频帧图像。

35 进一步的, 所述对象识别模块 501, 还用于:

在所述第一视频帧图像中识别所述第一对象的多个第一关键点和所述第二对象的多个第二关键点;

根据所述第一对象的多个第一关键点识别所述第一对象;

根据所述第二对象的多个第二关键点识别所述第二对象, 其中所述多个第二关键

点为所述第二对象的边缘关键点。

进一步的，所述第二视频帧图像生成模块 502，包括：

根据所述第一视频帧图像中所述第一对象的多个第一关键点计算所述第一视频帧图像中所述第一对象的定位点；

- 5 通过对齐所述第三对象的定位点与所述第一视频帧图像中所述第一对象的定位点，将所述第三对象作为前景图像覆盖到所述第一视频帧图像中，得到所述第二视频帧图像。

进一步的，所述第二视频帧图像生成模块 502，还用于：

计算所述第一对象上预设的两个关键点之间的第一距离；

- 10 根据所述第一距离计算所述第二对象的偏移量；

根据所述第二对象在所述第一视频帧图像中的位置以及所述第二对象的偏移量，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

进一步的，所述根据所述第一距离计算所述第二对象的偏移量，包括：

- 15 获取所述第一对象的偏航角以及俯仰角，其中所述偏航角为所述第一对象在所述第一视频帧图像中的朝向与镜头拍摄方向的水平夹角，所述俯仰角为第一对象在所述第一视频帧图像中的朝向与镜头拍摄方向的竖直夹角；

根据所述第一距离、偏航角以及俯仰角计算所述第二对象的偏移量。

- 20 进一步的，所述根据所述第二对象在所述第一视频帧图像中的位置以及所述第二对象的偏移量，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像，包括：

按照所述第二对象的偏移量对所述第二对象在所述第一视频帧图像中的初始定位点进行位置偏移，得到目标定位点；

- 25 在所述第二视频帧图像中，将所述第二对象作为前景图像渲染到所述目标定位点的位置，得到所述第三视频帧图像。

进一步的，所述图像处理装置 500，还包括：

- 30 第四视频图像生成模块，用于将所述第三视频帧图像中的第二对象以及第三对象的颜色空间从 RGB 颜色空间转换为 HSV 颜色空间；将所述第二对象在 HSV 颜色空间中的 H 通道的值替换为所述第三对象的 H 通道的值；将所述第二对象以及第三对象的颜色空间从 HSV 空间转换为 RGB 颜色空间得到第四视频帧图像。

图 5 所示装置可以执行图 1-图 4 所示实施例的方法，本实施例未详细描述的部分，可参考对图 1-图 4 所示实施例的相关说明。该技术方案的执行过程和技术效果参见图 1-图 4 所示实施例中的描述，在此不再赘述。

- 35 下面参考图 6，其示出了适于用来实现本公开实施例的电子设备（例如图 1 中的终端设备或服务器）600 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、PDA（个人数字助理）、PAD（平板电脑）、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、台式计算机等等的固定终端。图 6 示出的电子设备仅仅是一个示例，不对本公开实施例的功能和使用范围带来任何限制。

如图 6 所示, 电子设备 600 可以包括处理装置 (例如中央处理器、图形处理器等) 601, 其可以根据存储在只读存储器 (ROM) 602 中的程序或者从存储装置 608 加载到随机访问存储器 (RAM) 603 中的程序而执行各种适当的动作和处理。在 RAM 603 中, 还存储有电子设备 600 操作所需的各种程序和数据。处理装置 601、ROM 602 以及 RAM 603 通过总线 604 彼此相连。输入/输出 (I/O) 接口 605 也连接至总线 604。

通常, 以下装置可以连接至 I/O 接口 605: 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 606; 包括例如液晶显示器 (LCD)、扬声器、振动器等的输出装置 607; 包括例如磁带、硬盘等的存储装置 608; 以及通信装置 609。通信装置 609 可以允许电子设备 600 与其他设备进行无线或有线通信以交换数据。虽然图 6 示出了具有各种装置的电子设备 600, 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, 其包括承载在非暂态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 609 从网络上被下载和安装, 或者从存储装置 608 被安装, 或者从 ROM 602 被安装。在该计算机程序被处理装置 601 执行时, 执行本公开实施例的方法中限定的上述功能。

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一——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 (EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器 (CD-ROM)、光存储器件、磁存储器件、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、RF (射频) 等等, 或者上述的任意合适的组合。

在一些实施方式中, 客户端、服务器可以利用诸如 HTTP (HyperText Transfer Protocol, 超文本传输协议) 之类的任何当前已知或未来研发的网络协议进行通信, 并且可以与任意形式或介质的数字数据通信 (例如, 通信网络) 互连。通信网络的示例包括局域网 (“LAN”), 广域网 (“WAN”), 网际网 (例如, 互联网) 以及端对端网络 (例如, ad hoc 端对端网络), 以及任何当前已知或未来研发的网络。

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：识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象；按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像；其中，在所述第二视频帧图像中所述第三对象覆盖所述第一对象；根据所述第二对象在所述第一视频帧图像中的位置，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

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但不限于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。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(LAN)或广域网(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，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。

本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如，非限制性地，可以使用的示范类型的硬件逻辑部件包括：现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑设备(CPLD)等等。

在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM 或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

本公开实施例还提供了一种计算机程序，当所述计算机程序被计算机运行时，所述计算机执行本公开实施例提供的图像处理方法。

根据本公开的一个或多个实施例，提供了一种图像处理方法，包括：

5 识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象；

按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像；其中，在所述第二视频帧图像中所述第三对象覆盖所述第一对象；

10 根据所述第二对象在所述第一视频帧图像中的位置，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

进一步的，所述识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象，包括：

在所述第一视频帧图像中识别所述第一对象的多个第一关键点和所述第二对象的多个第二关键点；

15 根据所述第一对象的多个第一关键点识别所述第一对象；

根据所述第二对象的多个第二关键点识别所述第二对象，其中所述多个第二关键点为所述第二对象的边缘关键点。

进一步的，所述按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像，包括：

20 根据所述第一视频帧图像中所述第一对象的多个第一关键点计算所述第一视频帧图像中所述第一对象的定位点；

通过对齐所述第三对象的定位点与所述第一视频帧图像中所述第一对象的定位点，将所述第三对象作为前景图像覆盖到所述第一视频帧图像中，得到所述第二视频帧图像。

25 进一步的，所述根据所述第二对象在所述第一视频帧图像中的位置，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像，包括：

计算所述第一对象上预设的两个关键点之间的第一距离；

根据所述第一距离计算所述第二对象的偏移量；

30 根据所述第二对象在所述第一视频帧图像中的位置以及所述第二对象的偏移量，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

进一步的，所述根据所述第一距离计算所述第二对象的偏移量，包括：

35 获取所述第一对象的偏航角以及俯仰角，其中所述偏航角为所述第一对象在所述第一视频帧图像中的朝向与镜头拍摄方向的水平夹角，所述俯仰角为第一对象在所述第一视频帧图像中的朝向与镜头拍摄方向的竖直夹角；

根据所述第一距离、偏航角以及俯仰角计算所述第二对象的偏移量。

进一步的，所述根据所述第二对象在所述第一视频帧图像中的位置以及所述第二对象的偏移量，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象

上，得到第三视频帧图像，包括：

按照所述第二对象的偏移量对所述第二对象在所述第一视频帧图像中的初始定位点进行位置偏移，得到目标定位点；

5 在所述第二视频帧图像中，将所述第二对象作为前景图像渲染到所述目标定位点的位置，得到所述第三视频帧图像。

进一步的，在得到第三视频帧图像之后，还包括：

将所述第三视频帧图像中的第二对象以及第三对象的颜色空间从 RGB 颜色空间转换为 HSV 颜色空间；

10 将所述第二对象在 HSV 颜色空间中的 H 通道的值替换为所述第三对象的 H 通道的值；

将所述第二对象以及第三对象的颜色空间从 HSV 空间转换为 RGB 颜色空间得到第四视频帧图像。

根据本公开的一个或多个实施例，提供了一种图像处理装置，包括：

15 对象识别模块，用于识别第一视频帧图像中的第一对象和位于所述第一对象内的第二对象；

第二视频帧图像生成模块，用于按照所述第一对象在所述第一视频帧图像中的位置，将第三对象作为前景图像覆盖到所述第一视频帧图像上，得到第二视频帧图像；其中，在所述第二视频帧图像中所述第三对象覆盖所述第一对象；

20 第三视频帧生成模块，用于根据所述第二对象在所述第一视频帧图像中的位置，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

进一步的，所述对象识别模块，还用于：

在所述第一视频帧图像中识别所述第一对象的多个第一关键点和所述第二对象的多个第二关键点；

25 根据所述第一对象的多个第一关键点识别所述第一对象；

根据所述第二对象的多个第二关键点识别所述第二对象，其中所述多个第二关键点为所述第二对象的边缘关键点。

进一步的，所述第二视频帧图像生成模块，包括：

30 根据所述第一视频帧图像中所述第一对象的多个第一关键点计算所述第一视频帧图像中所述第一对象的定位点；

通过对齐所述第三对象的定位点与所述第一视频帧图像中所述第一对象的定位点，将所述第三对象作为前景图像覆盖到所述第一视频帧图像中，得到所述第二视频帧图像。

进一步的，所述第二视频帧图像生成模块，还用于：

35 计算所述第一对象上预设的两个关键点之间的第一距离；

根据所述第一距离计算所述第二对象的偏移量；

根据所述第二对象在所述第一视频帧图像中的位置以及所述第二对象的偏移量，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像。

进一步的，所述根据所述第一距离计算所述第二对象的偏移量，包括：

获取所述第一对象的偏航角以及俯仰角，其中所述偏航角为所述第一对象在所述第一视频帧图像中的朝向与镜头拍摄方向的水平夹角，所述俯仰角为所述第一对象在所述第一视频帧图像中的朝向与镜头拍摄方向的竖直夹角；

5 根据所述第一距离、偏航角以及俯仰角计算所述第二对象的偏移量。

进一步的，所述根据所述第二对象在所述第一视频帧图像中的位置以及所述第二对象的偏移量，将所述第二对象作为前景图像叠加到所述第二视频帧图像的第三对象上，得到第三视频帧图像，包括：

10 按照所述第二对象的偏移量对所述第二对象在所述第一视频帧图像中的初始定位点进行位置偏移，得到目标定位点；

在所述第二视频帧图像中，将所述第二对象作为前景图像渲染到所述目标定位点的位置，得到所述第三视频帧图像。

进一步的，所述图像处理装置，还包括：

15 第四视频图像生成模块，用于将所述第三视频帧图像中的第二对象以及第三对象的颜色空间从 RGB 颜色空间转换为 HSV 颜色空间；将所述第二对象在 HSV 颜色空间中的 H 通道的值替换为所述第三对象的 H 通道的值；将所述第二对象以及第三对象的颜色空间从 HSV 空间转换为 RGB 颜色空间得到第四视频帧图像。

根据本公开的一个或多个实施例，提供了一种计算机程序，当所述计算机程序被计算机运行时，所述计算机执行本公开实施例提供的图像处理方法。

20 以上描述仅为本公开的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。

25 此外，虽然采用特定次序描绘了各操作，但是这不应理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和并行处理可能是有利的。同样地，虽然在上面论述中包含了若干具体实现细节，但是这些不应被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实施例中。相反地，在单个实施例的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

30 尽管已经采用特定于结构特征和/或方法逻辑动作的语言描述了本主题，但是应当理解所附权利要求书中所限定的主题未必局限于上面描述的特定特征或动作。相反，上面所描述的特定特征和动作仅仅是实现权利要求书的示例形式。