



(12)发明专利申请

(10)申请公布号 CN 107578017 A

(43)申请公布日 2018.01.12

(21)申请号 201710806070.X

(22)申请日 2017.09.08

(71)申请人 百度在线网络技术(北京)有限公司
地址 100085 北京市海淀区上地十街10号
百度大厦三层

(72)发明人 何涛 张刚 刘经拓

(74)专利代理机构 北京英赛嘉华知识产权代理
有限责任公司 11204
代理人 王达佐 马晓亚

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

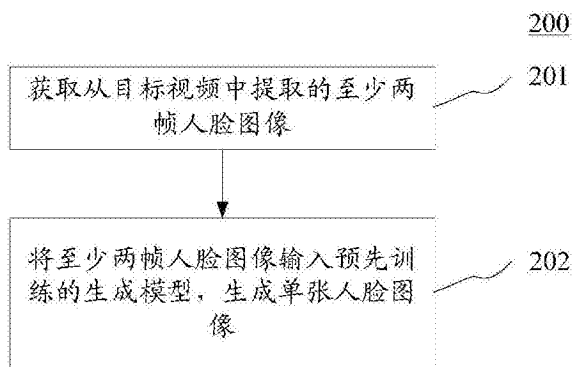
权利要求书2页 说明书9页 附图3页

(54)发明名称

用于生成图像的方法和装置

(57)摘要

本申请公开了用于生成图像的方法和装置。该方法的一具体实施方式包括:获取从目标视频中提取的至少两帧人脸图像;将至少两帧人脸图像输入预先训练的生成模型,生成单张人脸图像,生成模型在训练过程中利用损失函数更新模型参数,而该损失函数是基于单张人脸生成图像为真实人脸图像的概率和该单张人脸生成图像与标准人脸图像之间的相似度确定的。该实施方式可以提高生成模型所生成的单张人脸图像的真实性,进而提高了基于视频得到的人脸图像的质量。



1. 一种用于生成图像的方法,其特征在于,所述方法包括:

获取从目标视频中提取的至少两帧人脸图像;

将所述至少两帧人脸图像输入预先训练的生成模型,生成单张人脸图像,所述生成模型是通过以下训练步骤得到的:

将由初始生成模型输出的单张人脸生成图像输入预先训练的判别模型,生成所述单张人脸生成图像为真实人脸图像的概率;

基于所述概率和所述单张人脸生成图像与标准人脸图像之间的相似度确定所述初始生成模型的损失函数,其中,所述标准人脸图像和所述单张人脸生成图像包含同一人的的人脸信息;

利用所述损失函数更新所述初始生成模型的模型参数,得到生成模型。

2. 根据权利要求1所述的方法,其特征在于,所述确定所述初始生成模型的损失函数,包括:

使用预先训练的识别模型分别提取所述单张人脸生成图像的特征信息和所述标准人脸图像的特征信息,并计算所述单张人脸生成图像的特征信息与所述标准人脸图像的特征信息之间的欧式距离;

基于所述概率和所述欧式距离得到所述初始生成模型的损失函数。

3. 根据权利要求1所述的方法,其特征在于,所述初始生成模型是通过以下方式训练得到的:

利用机器学习方法,将从初始训练用视频中提取的至少两帧初始训练用人脸样本图像作为输入,将预先设定的初始训练用人脸图像作为输出,训练得到初始生成模型,其中,所述至少两帧初始训练用人脸样本图像和所述初始训练用人脸图像包含同一人的的人脸信息。

4. 根据权利要求1所述的方法,其特征在于,所述判别模型是通过以下方式训练得到的:

利用机器学习方法,将第一样本图像作为输入,将第一样本图像的标注信息作为输出,训练得到判别模型,其中,所述第一样本图像包括带有标注信息的正样本图像和带有标注信息的负样本图像,其中,所述负样本图像为所述生成模型输出的图像。

5. 根据权利要求2所述的方法,其特征在于,所述识别模型是通过以下方式训练得到的:

利用机器学习方法,将第二样本图像作为输入,所述第二样本图像的特征信息作为输出,训练得到识别模型。

6. 根据权利要求1所述的方法,其特征在于,所述生成模型为长短期记忆网络模型。

7. 一种用于生成图像的装置,其特征在于,所述装置包括:

获取单元,用于获取从目标视频中提取的至少两帧人脸图像;

生成单元,用于将所述至少两帧人脸图像输入预先训练的生成模型,生成单张人脸图像;

生成模型训练单元,用于训练所述生成模型;以及

所述生成模型训练单元包括:

概率生成单元,用于将由初始生成模型输出的单张人脸生成图像输入预先训练的判别模型,生成所述单张人脸生成图像为真实人脸图像的概率;

确定单元,用于基于所述概率和所述单张人脸生成图像与标准人脸图像之间的相似度确定所述初始生成模型的损失函数,其中,所述标准人脸图像和所述单张人脸生成图像包含同一人的脸信息;

更新单元,用于利用所述损失函数更新所述初始生成模型的模型参数,得到生成模型。

8. 根据权利要求7所述的装置,其特征在于,所述确定单元进一步用于:

使用预先训练的识别模型分别提取所述单张人脸生成图像的特征信息和所述标准人脸图像的特征信息,并计算所述单张人脸生成图像的特征信息与所述标准人脸图像的特征信息之间的欧式距离;

基于所述概率和所述欧式距离得到所述初始生成模型的损失函数。

9. 根据权利要求7所述的装置,其特征在于,所述装置还包括初始生成模型生成单元,所述初始生成模型生成单元用于:

利用机器学习方法,将从初始训练用视频中提取的至少两帧初始训练用人脸样本图像作为输入,将预先设定的初始训练用人脸图像作为输出,训练得到初始生成模型,其中,所述至少两帧初始训练用人脸样本图像和所述初始训练用人脸图像包含同一人的脸信息。

10. 根据权利要求7所述的装置,其特征在于,所述装置还包括判别模型训练单元,所述判别模型训练单元用于:

利用机器学习方法,将第一样本图像作为输入,将第一样本图像的标注信息作为输出,训练得到判别模型,其中,所述第一样本图像包括带有标注信息的正样本图像和带有标注信息的负样本图像,其中,所述负样本图像为所述生成模型输出的图像。

11. 根据权利要求8所述的装置,其特征在于,所述装置还包括识别模型训练单元,所述识别模型训练单元用于:

利用机器学习方法,将第二样本图像作为输入,所述第二样本图像的特征信息作为输出,训练得到识别模型。

12. 根据权利要求7所述的装置,其特征在于,所述生成模型为长短期记忆网络模型。

13. 一种终端,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,

当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如权利要求1-6中任一所述的方法。

14. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该计算机程序被处理器执行时实现如权利要求1-6中任一所述的方法。

用于生成图像的方法和装置

技术领域

[0001] 本申请涉及计算机技术领域,具体涉及图像处理领域,尤其涉及用于生成图像的方法和装置。

背景技术

[0002] 近年来,人脸识别技术发展迅速,应用的领域也越来越广。由于视频在拍摄中会受外界因素(例如,噪声、光线等)的影响,且视频中存在大量质量较差,姿态较大的人脸,这些都给人脸识别带来了很大的干扰,造成人脸识别效果不理想,例如不能从视频中识别出清楚、真实的人脸图像,又例如,将视频中的A错误的识别为B,这些问题会给后续的工作带来极大的不便,因此,如何从视频中获取高质量的人脸图像,并且保证获取的人脸图像是视频中的人的人脸图像是亟需解决的问题。

发明内容

[0003] 本申请的目的在于提出一种改进的用于生成图像的方法和装置,来解决以上背景技术部分提到的技术问题。

[0004] 第一方面,本申请实施例提供了一种用于生成图像的方法,该方法包括:获取从目标视频中提取的至少两帧人脸图像;将上述至少两帧人脸图像输入预先训练的生成模型,生成单张人脸图像,上述生成模型是通过以下训练步骤得到的:将由初始生成模型输出的单张人脸生成图像输入预先训练的判别模型,生成上述单张人脸生成图像为真实人脸图像的概率;基于上述概率和上述单张人脸生成图像与标准人脸图像之间的相似度确定上述初始生成模型的损失函数,其中,上述标准人脸图像和上述单张人脸生成图像包含同一人的的人脸信息;利用上述损失函数更新上述初始生成模型的模型参数,得到生成模型。

[0005] 在一些实施例中,上述确定上述初始生成模型的损失函数,包括:使用预先训练的识别模型分别提取上述单张人脸生成图像的特征信息和上述标准人脸图像的特征信息,并计算上述单张人脸生成图像的特征信息与上述标准人脸图像的特征信息之间的欧式距离;基于上述概率和上述欧式距离得到上述初始生成模型的损失函数。

[0006] 在一些实施例中,上述初始生成模型是通过以下方式训练得到的:利用机器学习方法,将从初始训练用视频中提取的至少两帧初始训练用人脸样本图像作为输入,将预先设定的初始训练用人脸图像作为输出,训练得到初始生成模型,其中,上述至少两帧初始训练用人脸样本图像和上述初始训练用人脸图像包含同一人的的人脸信息。

[0007] 在一些实施例中,上述判别模型是通过以下方式训练得到的:利用机器学习方法,将第一样本图像作为输入,将第一样本图像的标注信息作为输出,训练得到判别模型,其中,上述第一样本图像包括带有标注信息的正样本图像和带有标注信息的负样本图像,其中,上述负样本图像为上述生成模型输出的图像。

[0008] 在一些实施例中,上述识别模型是通过以下方式训练得到的:利用机器学习方法,将第二样本图像作为输入,上述第二样本图像的特征信息作为输出,训练得到识别模型。

[0009] 在一些实施例中,上述生成模型为长短期记忆网络模型。

[0010] 第二方面,本申请实施例提供了一种用于生成图像的装置,该装置包括:获取单元,用于获取从目标视频中提取的至少两帧人脸图像;生成单元,用于将上述至少两帧人脸图像输入预先训练的生成模型,生成单张人脸图像;生成模型训练单元,用于训练上述生成模型;以及上述生成模型训练单元包括:概率生成单元,用于将由初始生成模型输出的单张人脸生成图像输入预先训练的判别模型,生成上述单张人脸生成图像为真实人脸图像的概率;确定单元,用于基于上述概率和上述单张人脸生成图像与标准人脸图像之间的相似度确定上述初始生成模型的损失函数,其中,上述标准人脸图像和上述单张人脸生成图像包含同一人的人脸信息;更新单元,用于利用上述损失函数更新上述初始生成模型的模型参数,得到生成模型。

[0011] 在一些实施例中,上述确定单元进一步用于:使用预先训练的识别模型分别提取上述单张人脸生成图像的特征信息和上述标准人脸图像的特征信息,并计算上述单张人脸生成图像的特征信息与上述标准人脸图像的特征信息之间的欧式距离;基于上述概率和上述欧式距离得到上述初始生成模型的损失函数。

[0012] 在一些实施例中,上述装置还包括初始生成模型生成单元,上述初始生成模型生成单元用于:利用机器学习方法,将从初始训练用视频中提取的至少两帧初始训练用人脸样本图像作为输入,将预先设定的初始训练用人脸图像作为输出,训练得到初始生成模型,其中,上述至少两帧初始训练用人脸样本图像和上述初始训练用人脸图像包含同一人的人脸信息。

[0013] 在一些实施例中,上述装置还包括判别模型训练单元,上述判别模型训练单元用于:利用机器学习方法,将第一样本图像作为输入,将第一样本图像的标注信息作为输出,训练得到判别模型,其中,上述第一样本图像包括带有标注信息的正样本图像和带有标注信息的负样本图像,其中,上述负样本图像为上述生成模型输出的图像。

[0014] 在一些实施例中,上述装置还包括识别模型训练单元,上述识别模型训练单元用于:利用机器学习方法,将第二样本图像作为输入,上述第二样本图像的特征信息作为输出,训练得到识别模型。

[0015] 在一些实施例中,上述生成模型为长短期记忆网络模型。

[0016] 第三方面,本申请实施例提供了一种终端,该终端包括:一个或多个处理器;存储装置,用于存储一个或多个程序,当上述一个或多个程序被上述一个或多个处理器执行时,使得上述一个或多个处理器实现如第一方面中任一实现方式描述的方法。

[0017] 第四方面,本申请实施例提供了一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该计算机程序被处理器执行时实现如第一方面中任一实现方式描述的方法。

[0018] 本申请实施例提供的用于生成图像的方法和装置,基于预先训练的生成模型使用从目标视频中提取的至少两帧人脸图像生成单张人脸图像,上述生成模型在训练过程中利用损失函数更新模型参数,而该损失函数是基于单张人脸生成图像为真实人脸图像的概率和该单张人脸生成图像与标准人脸图像之间的相似度确定的,因此,可以提高生成模型所生成的单张人脸图像的真实性,进而提高了基于视频得到的人脸图像的质量。

附图说明

[0019] 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述,本申请的其它特征、目的和优点将会变得更明显:

[0020] 图1是本申请可以应用于其中的示例性系统架构图;

[0021] 图2是根据本申请的用于生成图像的方法的一个实施例的流程图;

[0022] 图3是根据本申请的用于生成图像的方法的一个应用场景的示意图;

[0023] 图4是根据本申请的用于生成图像的装置的一个实施例的结构示意图;

[0024] 图5是适于用来实现本申请实施例的终端设备的计算机系统的结构示意图。

具体实施方式

[0025] 下面结合附图和实施例对本申请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相关发明,而非对该发明的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与有关发明相关的部分。

[0026] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

[0027] 图1示出了可以应用本申请的用于生成图像的方法或用于生成图像的装置的实施例的示例性系统架构100。

[0028] 如图1所示,系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

[0029] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种客户端应用,例如游戏类应用、动画展示类应用、即时通信工具、社交平台软件等。

[0030] 终端设备101、102、103可以是具有显示屏并且支持图像显示的各种电子设备,包括但不限于智能手机、平板电脑、电子书阅读器、MP3播放器(Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3)、MP4(Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面4)播放器、膝上型便携计算机和台式计算机等等。

[0031] 服务器105可以是提供各种服务的服务器,例如对终端设备101、102、103上显示的图像或图形提供支持的后台服务器。后台服务器可以将数据(例如图像数据)反馈给终端设备,以供终端设备进行展示。

[0032] 需要说明的是,本申请实施例所提供的用于生成图像的方法可以通过终端设备101、102、103执行,也可以通过服务器105执行,还可以通过服务器105和终端设备101、102、103共同执行,相应地,用于生成图像的装置可以设置于服务器105中,也可以设置于终端设备101、102、103中,还可以部分单元设置于服务器105中并将其他单元设置于终端设备101、102、103中。本申请对此不做限定。

[0033] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器。

[0034] 继续参考图2,其示出了根据本申请的用于生成图像的方法的一个实施例的流程200。该用于生成图像的方法,包括以下步骤:

[0035] 步骤201,获取从目标视频中提取的至少两帧人脸图像。

[0036] 在本实施例中,用于生成图像的方法运行于其上的电子设备(例如图1所示的终端设备101、102、103或服务器105)可以从本地或者远程获取从目标视频中提取的至少两帧人脸图像,上述至少两帧人脸图像中的各帧人脸图像为同一人的人脸图像,上述至少两帧人脸图像可以是同一段目标视频中提取出的、同一人的人脸图像序列。从目标视频中提取的至少两帧人脸图像可以为低质量的人脸图像,例如,存在噪音、模糊、分辨率低、人脸姿态大等问题的人脸图像。

[0037] 步骤202,将至少两帧人脸图像输入预先训练的生成模型,生成单张人脸图像。

[0038] 在本实施例中,基于步骤201中获取的至少两帧人脸图像,上述电子设备可以将上述至少两帧人脸图像输入预先训练的生成模型生成单张人脸图像,该单张人脸图像可以为图像分辨率高于预先设定的第一分辨率阈值的、且不存在噪声、模糊、光线暗或过度曝光等问题的高质量人脸图像。在这里,上述第一分辨率阈值可以根据实际需要人工设定的。

[0039] 在这里,上述生成模型可以是上述电子设备或者其他用于训练上述生成模型的电子设备通过以下训练步骤得到的:首先,可以将由初始生成模型输出的单张人脸生成图像输入预先训练的判别模型,生成上述单张人脸生成图像为真实人脸图像的概率。上述初始生成模型可以通过各种方式得到的神经网络模型,例如,基于现有的神经网络(例如长短期记忆网络),随机生成该神经网络的网络参数,得到的神经网络模型。然后,可以基于上述概率和上述单张人脸生成图像与标准人脸图像之间的相似度(例如,余弦相似度、杰卡德相似系数、欧氏距离等等)确定上述初始生成模型的损失函数,其中,上述标准人脸图像和上述单张人脸生成图像包含同一人的人脸信息。在这里,上述标准人脸图像可以是预先设定的、不存在噪声、模糊、光线暗或过度曝光等问题的高质量人脸图像,例如,可以是预先拍摄的符合需求的高质量人脸图像,还可以是证件照等等。最后,可以利用上述损失函数更新上述初始生成模型的模型参数,得到生成模型。例如,将上述损失函数反向传播回上述初始生成模型,以更新上述初始生成模型的模型参数。需要说明的是,上述生成模型的训练过程仅仅用于说明生成模型参数的调整过程,可以认为初始生成模型为参数调整前的模型,生成模型为参数调整后的模型,模型参数的调整过程并不仅限于一次,可以根据模型的优化程度以及实际需要等重复多次。

[0040] 常见的生成模型可以包括但不限于深度神经网络模型、隐马尔可夫模型(Hidden Markov Model,HMM)、朴素贝叶斯模型、高斯混合模型等等。上述生成模型可以是生成式对抗网络(Generative Adversarial Networks,GAN)包含的生成模型,生成式对抗网络(Generative Adversarial Networks,GAN)启发自博弈论中的二人零和博弈(two-player game),GAN模型中的两位博弈方分别由生成模型(generative model)和判别模型(discriminative model)充当。生成模型捕捉样本数据的分布,生成一个类似真实训练数据的样本,追求效果是越像真实样本越好。判别模型是一个二分类器,判别一个样本来自于真实训练数据(而非生成模型的生成数据)的概率,常见的判别模型可以包括但不限于线性回归模型、线性判别分析、支持向量机(Support Vector Machine,SVM)、神经网络等等。此处,生成模型和判别模型可以同时训练:固定判别模型,调整生成模型参数;固定生成模型,

调整判别模型参数。本实施例中,生成模型通过不断学习,生成越来越逼真的人脸图像;而判别模型通过不断地学习,增强对生成的人脸图像和真实人脸图像的区分能力。通过生成模型与判别模型之间的对抗,最终,生成模型生成的人脸图像接近于真实人脸图像而成功“欺骗”了判别模型。这样的生成式对抗网络可以用于提高生成人脸图像的真实性。

[0041] 在本实施例的一些可选的实现方式中,上述确定上述初始生成模型的损失函数,可以具体包括:首先,可以使用预先训练的识别模型分别提取上述单张人脸生成图像的特征信息和上述标准人脸图像的特征信息,并计算上述单张人脸生成图像的特征信息与上述标准人脸图像的特征信息之间的欧式距离,其中,上述标准人脸图像和上述单张人脸生成图像包含同一人的人脸信息,然后,可以基于上述概率和上述欧式距离得到上述初始生成模型的损失函数。在实际的深度神经网络模型的训练过程中通常一个批次(batch)同时训练多个样本数据。在一些可选的实现方式中,假设上述生成模型的损失函数为 $J^{(G)}$, $J^{(G)}$ 的计算公式可以为:

$$[0042] \quad J^{(G)} = -\frac{1}{2} E_x \log(D(x)) + \|F(x) - F(x_0)\|_2$$

[0043] 其中, x 表示人脸生成图像的像素矩阵; $D(x)$ 表示 x 输入到判别模型后的输出; $E_x \log(D(x))$ 表示多个 x 同时训练时多个 $\log(D(x))$ 的期望; $F(x)$ 表示人脸生成图像输入到识别网络后输出的特征向量; x_0 表示标准人脸图像的像素矩阵; $F(x_0)$ 表示标准人脸图像输入到识别模型后的输出的特征向量; $\|F(x) - F(x_0)\|_2$ 表示 $F(x)$ 与 $F(x_0)$ 的2-范数,用于表示 $F(x)$ 和 $F(x_0)$ 的欧氏距离。生成模型的损失函数 $J^{(G)}$ 中加入 $F(x)$ 和 $F(x_0)$ 的欧氏距离用于保证生成模型输出的人脸图像的特征信息与标准人脸图像的特征信息的欧式距离尽可能小,从而保证了生成人脸图像与标准人脸图像之间的相似性,即保证了生成人脸图像为视频中的人的人脸图像。

[0044] 在一些可选的实现方式中,上述初始生成模型可以是通过以下方式训练得到的:上述电子设备或者其他用于训练上述初始生成模型的电子设备可以利用机器学习方法(例如长短期记忆网络),将从初始训练用视频中提取的至少两帧初始训练用人脸样本图像作为输入,将预先设定的初始训练用人脸图像作为输出,上述初始训练用人脸图像可以是图像分辨率超过上述第一分辨率阈值的人脸图像,训练得到初始生成模型,其中,上述至少两帧初始训练用人脸样本图像和上述初始训练用人脸图像包含同一人的人脸信息,初始训练用人脸图像为不存在噪声、模糊、光线暗或过度曝光等问题的高质量人脸图像。

[0045] 在一些可选的实现方式中,上述判别模型可以是通过以下方式训练得到的:上述电子设备或者其他用于训练上述判别模型的电子设备可以利用机器学习方法,将第一样本图像作为输入,将第一样本图像的标注信息作为输出,训练得到判别模型,其中,上述第一样本图像包括带有标注信息的正样本图像和带有标注信息负样本图像,其中,上述正样本图像可以是图像分辨率高于上述第一分辨率阈值的真实人脸图像,上述负样本图像为生成模型输出的图像。例如,在判别模型的训练过程中,可以将真实人脸图像作为人脸正样本图像,标注为1,将上述生成模型输出的图像作为人脸负样本图像,标注为0,计算判别模型的损失函数,将判别模型的损失函数反向传播回判别模型,以利用该损失函数更新判别模型的模型参数,从而实现判别模型参数的调整。在一些可选的实现方式中,假设判别网络损失函数为 $J^{(D)}$,其中, $J^{(D)}$ 的计算公式为:

$$[0046] \quad J^{(D)} = -\frac{1}{2} E_{x_1} \log(D(x_1)) - \frac{1}{2} E_x \log(1-D(x))$$

[0047] 其中： x_1 表示图像分辨率高于上述第一分辨率阈值的真实人脸图像的像素矩阵； $D(x_1)$ 表示 x_1 输入到判别模型后的输出； $E_{x_1} \log(D(x_1))$ 表示多个 x_1 同时训练时多个 $\log(D(x_1))$ 的期望； x 表示人脸生成图像的像素矩阵； $D(x)$ 表示 x 输入到判别模型后的输出； $E_x \log(1-D(x))$ 表示多个 x 同时训练时多个的 $\log(1-D(x))$ 的期望。上述判别模型可以为卷积神经网络模型。

[0048] 在一些可选的实现方式中，上述识别模型可以通过以下方式训练得到的：上述电子设备或者其他用于训练识别模型的电子设备可以利用机器学习方法，将第二样本图像作为输入，上述第二样本图像的特征信息作为输出，训练得到识别模型。其中，上述第二样本图像为人脸图像，上述第二样本图像的特征信息可以是指一个代表人脸的特征向量，例如一个代表人脸的512维的特征向量。

[0049] 在一些可选的实现方式中，上述生成模型可以为长短期记忆网络模型(Long-Short Term Memory, LSTM)。LSTM是一种时间递归神经网络，适合于处理和预测时间序列中间隔和延迟相对较长的重要事件。LSTM在基于从目标视频中提取的至少两帧人脸图像生成单张人脸图像时可以充分利用上述至少两帧人脸图像在视频中的时序关系，从而提高生成的人脸图像的质量。

[0050] 继续参见图3，图3是根据本实施例的用于生成图像的方法的应用场景的一个示意图。在图3的应用场景中，首先，终端设备（例如，手机、电脑等等）获取从目标视频中提取的10帧人脸图像 $C_1 \sim C_{10}$ ，这10帧图像存在噪声、模糊、分辨率低、人脸姿态大等问题，终端设备将获取的10帧人脸图像 $C_1 \sim C_{10}$ 输入预先训练的生成模型301，由生成模型301生成人脸图像 C ，人脸图像 C 为不存在噪声、模糊、分辨率低、人脸姿态大等问题的高质量人脸图像。需要说明的是，图3中的图像仅仅用于示意性的说明基于从目标视频中提取的至少两帧人脸图像生成单张人脸图像的过程，而非对输入图像张数、内容等的限定。

[0051] 本申请的上述实施例提供的方法使用预先训练的生成模型生成人脸图像，上述生成模型在训练过程中利用损失函数更新模型参数，而该损失函数是基于单张人脸生成图像为真实人脸图像的概率和该单张人脸生成图像与标准人脸图像之间的相似度确定的，因此，可以提高生成模型所生成的单张人脸图像的真实性，进而提高了基于视频得到的人脸图像的质量。

[0052] 进一步参考图4，作为对上述各图所示方法的实现，本申请提供了一种用于生成图像的装置的一个实施例，该装置实施例与图2所示的方法实施例相对应，该装置具体可以应用于各种电子设备中。

[0053] 如图4所示，本实施例的用于生成图像的装置400包括：获取单元401、生成单元402和生成模型训练单元403。其中，获取单元401用于获取从目标视频中提取的至少两帧人脸图像；生成单元402用于将上述至少两帧人脸图像输入预先训练的生成模型，生成单张人脸图像；生成模型训练单元403用于训练上述生成模型；以及上述生成模型训练单元403包括：概率生成单元4031，用于将由初始生成模型输出的单张人脸生成图像输入预先训练的判别模型，生成上述单张人脸生成图像为真实人脸图像的概率；确定单元4032，用于基于上述概率和上述单张人脸生成图像与标准人脸图像之间的相似度确定上述初始生成模型的损失

函数,其中,上述标准人脸图像和上述单张人脸生成图像包含同一人的脸信息;更新单元4033,用于利用上述损失函数更新上述初始生成模型的模型参数,得到生成模型。在本实施例中,用于生成图像的装置400的获取单元401、生成单元402和生成模型训练单元403的具体处理及其所带来的技术效果可分别参考图2对应实施例中步骤201和步骤202的相关说明,在此不再赘述。

[0054] 在本实施例的一些可选的实现方式中,上述确定单元4032进一步用于:使用预先训练的识别模型分别提取上述单张人脸生成图像的特征信息和上述标准人脸图像的特征信息,并计算上述单张人脸生成图像的特征信息与上述标准人脸图像的特征信息之间的欧式距离;基于上述概率和上述欧式距离得到上述初始生成模型的损失函数。

[0055] 在本实施例的一些可选的实现方式中,上述装置还包括初始生成模型生成单元(图中未示出),上述初始生成模型生成单元用于:利用机器学习方法,将从初始训练用视频中提取的至少两帧初始训练用人脸样本图像作为输入,将预先设定的初始训练用人脸图像作为输出,训练得到初始生成模型,其中,上述至少两帧初始训练用人脸样本图像和上述初始训练用人脸图像包含同一人的脸信息。

[0056] 在本实施例的一些可选的实现方式中,上述装置还包括判别模型训练单元(图中未示出),上述判别模型训练单元用于:利用机器学习方法,将第一样本图像作为输入,将第一样本图像的标注信息作为输出,训练得到判别模型,其中,上述第一样本图像包括带有标注信息的正样本图像和带有标注信息的负样本图像,其中,上述负样本图像为上述生成模型输出的图像。

[0057] 在本实施例的一些可选的实现方式中,上述装置还包括识别模型训练单元(图中未示出),上述识别模型训练单元用于:利用机器学习方法,将第二样本图像作为输入,上述第二样本图像的特征信息作为输出,训练得到识别模型。

[0058] 在本实施例的一些可选的实现方式中,上述生成模型为长短期记忆网络模型。

[0059] 下面参考图5,其示出了适于用来实现本申请实施例的终端设备的计算机系统500的结构示意图。图5示出的终端设备仅仅是一个示例,不应对本申请实施例的功能和使用范围带来任何限制。

[0060] 如图5所示,计算机系统500包括中央处理单元(CPU,Central Processing Unit)501,其可以根据存储在只读存储器(ROM,Read Only Memory)502中的程序或者从存储部分508加载到随机访问存储器(RAM,Random Access Memory)503中的程序而执行各种适当的动作和处理。在RAM 503中,还存储有系统500操作所需的各种程序和数据。CPU 501、ROM 502以及RAM 503通过总线504彼此相连。输入/输出(I/O,Input/Output)接口505也连接至总线504。

[0061] 以下部件连接至I/O接口505:包括键盘、鼠标等的输入部分506;包括诸如阴极射线管(CRT,Cathode Ray Tube)、液晶显示器(LCD,Liquid Crystal Display)等以及扬声器等的输出部分507;包括硬盘等的存储部分508;以及包括诸如LAN(局域网,Local Area Network)卡、调制解调器等网络接口卡的通信部分509。通信部分509经由诸如因特网的网络执行通信处理。驱动器510也根据需要连接至I/O接口505。可拆卸介质511,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器510上,以便于从其上读出的计算机程序根据需要被安装入存储部分508。

[0062] 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部分509从网络上被下载和安装,和/或从可拆卸介质511被安装。在该计算机程序被中央处理单元(CPU)501执行时,执行本申请的方法中限定的上述功能。需要说明的是,本申请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(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。在本申请中,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。而在本申请中,计算机可读的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。计算机可读的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:无线、电线、光缆、RF等等,或者上述的任意合适的组合。

[0063] 附图中的流程图和框图,图示了按照本申请各种实施例的系统、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,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。

[0064] 描述于本申请实施例中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。所描述的单元也可以设置在处理器中,例如,可以描述为:一种处理器包括获取单元、生成单元和生成模型训练单元。其中,这些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,获取单元还可以被描述为“获取从目标视频中提取的至少两帧人脸图像的单元”。

[0065] 作为另一方面,本申请还提供了一种计算机可读介质,该计算机可读介质可以是上述实施例中描述的装置中所包含的;也可以是单独存在,而未装配入该装置中。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装置执行时,使得该装置:获取从目标视频中提取的至少两帧人脸图像;将上述至少两帧人脸图像输入预先训练的生成模型,生成单张人脸图像,上述生成模型是通过以下训练步骤得到的:将由初始生

成模型输出的单张人脸生成图像输入预先训练的判别模型,生成上述单张人脸生成图像为真实人脸图像的概率;基于上述概率和上述单张人脸生成图像与标准人脸图像之间的相似度确定上述初始生成模型的损失函数,其中,上述标准人脸图像和上述单张人脸生成图像包含同一人的脸信息;利用上述损失函数更新上述初始生成模型的模型参数,得到生成模型。

[0066] 以上描述仅为本申请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,本申请中所涉及的发明范围,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发明构思的情况下,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。例如上述特征与本申请中公开的(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。

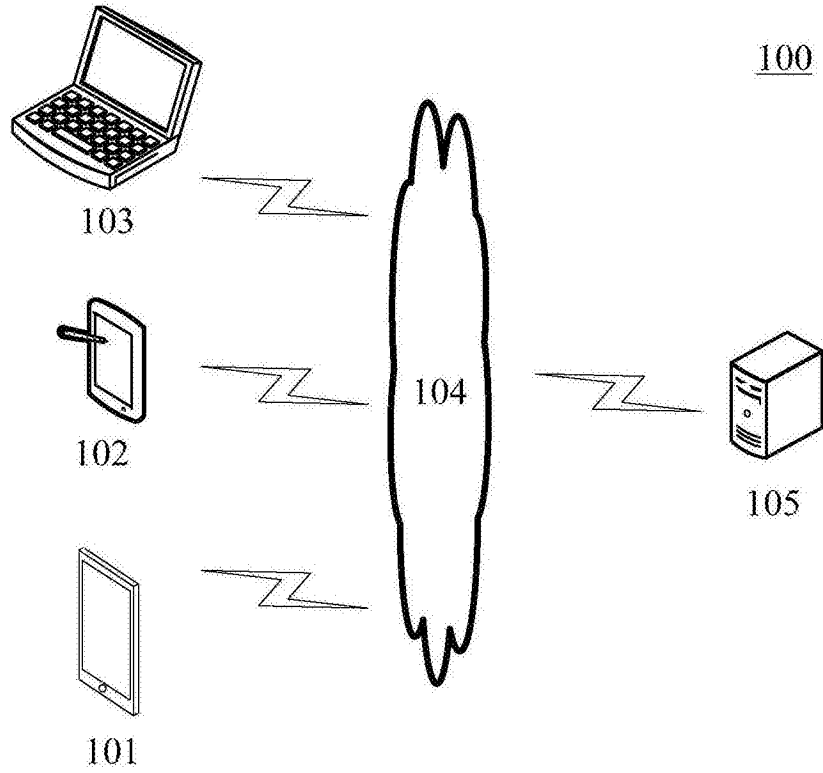


图1

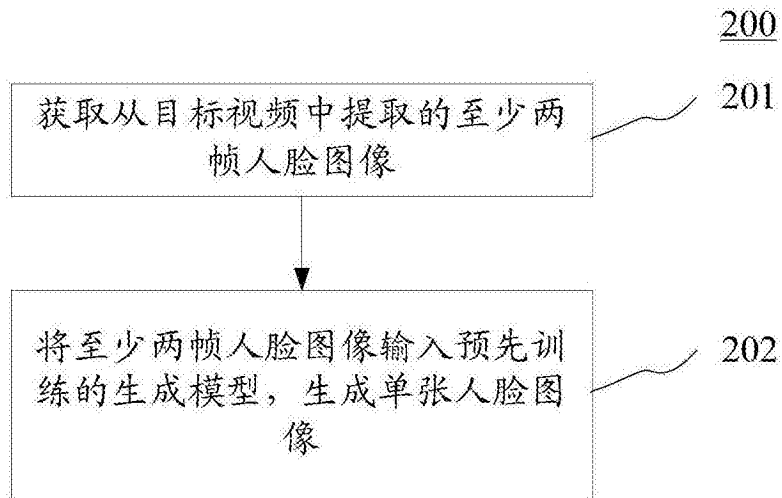


图2

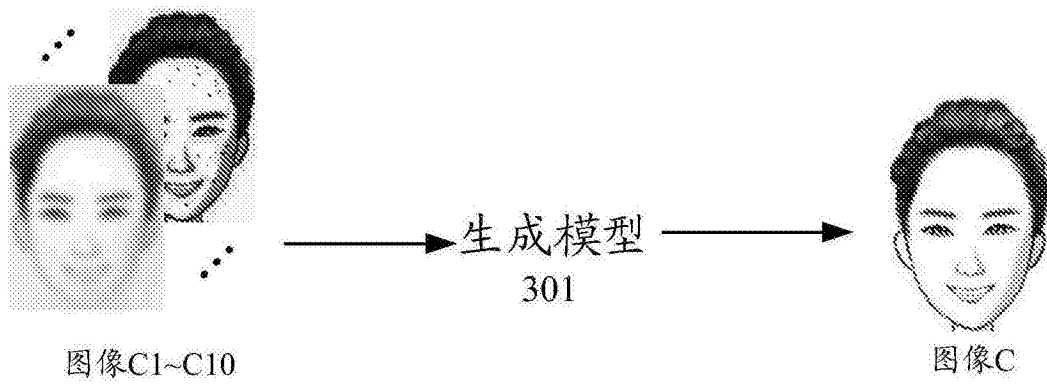


图3

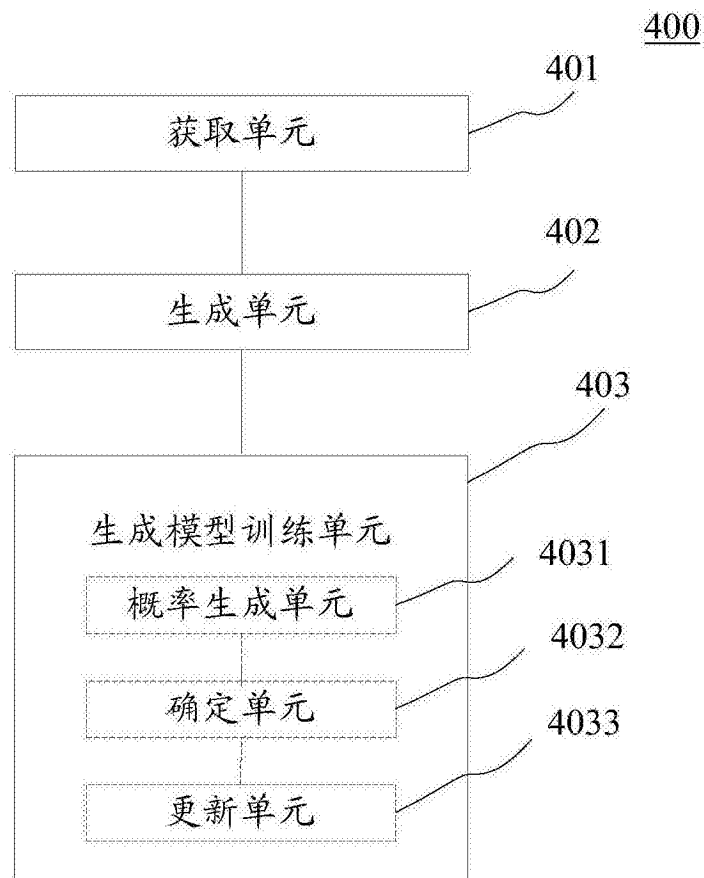


图4

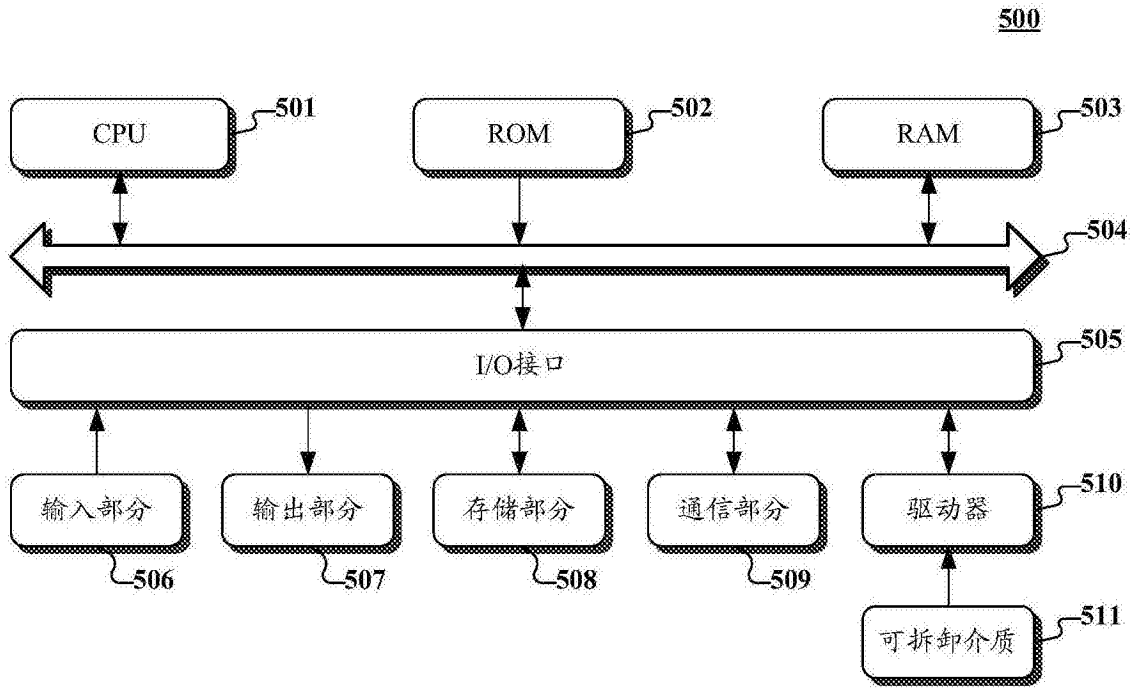


图5