



(12) 发明专利

(10) 授权公告号 CN 111275784 B

(45) 授权公告日 2023. 06. 13

(21) 申请号 202010065575.7

(22) 申请日 2020.01.20

(65) 同一申请的已公布的文献号
申请公布号 CN 111275784 A

(43) 申请公布日 2020.06.12

(73) 专利权人 北京百度网讯科技有限公司
地址 100085 北京市海淀区上地十街10号
百度大厦2层

(72) 发明人 李鑫 刘霄 张赫男 孙昊 李甫
何栋梁 周志超 龙翔 王平
文石磊 丁二锐

(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
专利代理师 王达佐 马晓亚

(51) Int. Cl.

G06T 11/00 (2006.01)

G06V 40/16 (2022.01)

G06N 3/0464 (2023.01)

G06N 3/0475 (2023.01)

G06N 3/094 (2023.01)

(56) 对比文件

CN 109858445 A, 2019.06.07

US 2019102640 A1, 2019.04.04

审查员 毛洁

权利要求书2页 说明书11页 附图7页

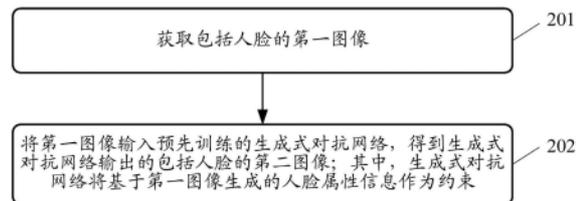
(54) 发明名称

生成图像的方法和装置

(57) 摘要

本公开实施例公开了生成图像的方法和装置。方法包括：获取包括第一人脸的第一图像；将第一图像输入预先训练的生成式对抗网络，得到生成式对抗网络输出的包括第二人脸的第二图像；其中，生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束。该方法可以提高生成式对抗网络基于输入的第一图像生成第二图像的准确率和效率，减少误生成图像的概率，并且减少了对输入的图像只能为预定区域内的图像的约束。

200



1. 一种生成图像的方法,包括:
 - 获取包括第一人臉的第一图像;
 - 将所述第一图像输入预先训练的生成式对抗网络,得到生成式对抗网络输出的包括第二人脸的第二图像;其中,所述生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束;
 - 其中,所述将所述第一图像输入预先训练的生成式对抗网络包括:
 - 对所述第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络;或
 - 检测所述第一图像的纹理特征参数值是否大于纹理阈值,若是,对所述第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络。
2. 根据权利要求1所述的方法,其中,所述生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束包括:生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束。
3. 根据权利要求2所述的方法,其中,所述生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束包括:
 - 所述生成式对抗网络将基于输入的图像生成的多通道人脸图像作为输入;其中,所述多通道人脸图像包括输入的图像的RGB三通道和输入的图像的以下至少一项:
 - 人脸关键点的二值图一通道或RGB三通道;
 - 人脸语义分割结果的二值图一通道或RGB三通道;以及
 - 头发的二值图一通道。
4. 根据权利要求1所述的方法,其中,所述预先训练的生成式对抗网络用于训练的输入的图像样本包括:
 - 对初始图像样本进行数据增强后得到的图像样本。
5. 根据权利要求4所述的方法,其中,所述对初始图像样本进行数据增强包括:
 - 对所述初始图像样本进行以下至少一项操作:旋转、翻转、缩放和不同程度的高斯模糊。
6. 根据权利要求1所述的方法,其中,所述生成式对抗网络至少包括以下任意一项:生成式对抗网络GAN、循环一致生成式对抗网络CycleGan、人脸高精度属性编辑模型AttGAN、星型生成式对抗网络StarGAN、空间变换器生成式对抗网络STGAN。
7. 根据权利要求1-6任意一项所述的方法,其中,所述第一图像为真实的人脸图像;所述第二图像为卡通图像。
8. 一种生成图像的装置,包括:
 - 获取单元,被配置成获取包括第一人臉的第一图像;
 - 生成单元,被配置成将所述第一图像输入预先训练的生成式对抗网络,得到生成式对抗网络输出的包括第二人脸的第二图像;其中,所述生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束;
 - 所述生成单元进一步被配置成:
 - 对所述第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入

预先训练的生成式对抗网络;或

检测所述第一图像的纹理特征参数值是否大于纹理阈值,若是,对所述第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络。

9. 根据权利要求8所述的装置,其中,所述生成单元中所采用的所述生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束包括:所述生成单元中所采用的生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束。

10. 根据权利要求9所述的装置,其中,所述生成单元中所采用的生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束包括:

所述生成单元中所采用的生成式对抗网络将基于输入的图像生成的多通道人脸图像作为输入;其中,所述多通道人脸图像包括输入的图像的RGB三通道和输入的图像的以下至少一项:

人脸关键点的二值图一通道或RGB三通道;

人脸语义分割结果的二值图一通道或RGB三通道;以及

头发的二值图一通道。

11. 根据权利要求8所述的装置,其中,所述生成单元中所采用的所述预先训练的生成式对抗网络用于训练的输入的图像样本包括:

对初始图像样本进行数据增强后得到的图像样本。

12. 根据权利要求11所述的装置,其中,所述生成单元中所采用的所述预先训练的生成式对抗网络用于训练的输入的图像样本包括:

对初始图像样本进行以下至少一项操作:旋转、翻转、缩放和不同程度的高斯模糊。

13. 根据权利要求8所述的装置,其中,所述生成式对抗网络至少包括以下任意一项:生成式对抗网络GAN、循环一致生成式对抗网络CycleGan、人脸高精度属性编辑模型AttGAN、星型生成式对抗网络StarGAN、空间变换器生成式对抗网络STGAN。

14. 根据权利要求8-13任意一项所述的装置,其中,所述第一图像为真实的人脸图像;所述第二图像为卡通图像。

15. 一种电子设备,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;

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中任一所述的方法。

16. 一种计算机可读介质,其上存储有计算机程序,该程序被处理器执行时实现如权利要求1-7中任一所述的方法。

生成图像的方法和装置

技术领域

[0001] 本公开涉及计算机技术领域,具体涉及图像转换技术领域,尤其涉及生成图像的方法和装置。

背景技术

[0002] 目前市场上有很多自拍特效玩法,比如变小孩,换脸等等,有很大的趣味性,很多用户都喜欢玩,甚至出现一夜爆火的现象级应用(APP)。

[0003] 现有的技术方案有两种:一种是基于生成式对抗网络的照片漫画化方法(cartoon gan),类似于在输入的图像上加个卡通滤镜,基本不能生成可用的卡通形象。另一种是最近发表的论文-无监督生成注意网络(UGATIT)的算法,但是在实际使用中会发现卡通形象的生成率比较低,很多坏例子(bad case)(比如衣服上出现眼睛,面部很多纹理)或者根本没法生成,并且对输入的图像有严格限制(方形,只能输入正面脸部图像),而且现有的方法只能生成女生。

发明内容

[0004] 本公开实施例提供了生成图像的方法和装置。

[0005] 第一方面,本公开实施例提供了一种生成图像的方法,包括:获取包括第一人脸的第一图像;将第一图像输入预先训练的生成式对抗网络,得到生成式对抗网络输出的包括第二人脸的第二图像;其中,生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束。

[0006] 在一些实施例中,将第一图像输入预先训练的生成式对抗网络包括:对第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络;或检测第一图像的纹理特征参数值是否大于纹理阈值,若是,对第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络。

[0007] 在一些实施例中,生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束包括:生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束。

[0008] 在一些实施例中,生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束包括:生成式对抗网络将基于输入的图像生成的多通道人脸图像作为输入;其中,多通道人脸图像包括输入的图像的RGB三通道和输入的图像的以下至少一项:人脸关键点的二值图一通道或RGB三通道;人脸语义分割结果的二值图一通道或RGB三通道;以及头发的二值图一通道。

[0009] 在一些实施例中,预先训练的生成式对抗网络用于训练的输入的图像样本包括:对初始图像样本进行数据增强后得到的图像样本。

[0010] 在一些实施例中,对初始图像样本进行数据增强包括:对初始图像样本进行以下至少一项操作:旋转、翻转、缩放和不同程度的高斯模糊。

[0011] 在一些实施例中,生成式对抗网络至少包括以下任意一项:生成式对抗网络GAN、循环一致生成式对抗网络CycleGan、人脸高精度属性编辑模型AttGAN、星型生成式对抗网络StarGAN、空间变换器生成式对抗网络STGAN。

[0012] 在一些实施例中,第一图像为真实的人脸图像;第二图像为卡通图像。

[0013] 第二方面,本公开实施例提供了一种生成图像的装置,包括:获取单元,被配置成获取包括第一人臉的第一图像;生成单元,被配置成将第一图像输入预先训练的生成式对抗网络,得到生成式对抗网络输出的包括第二人脸的第二图像;其中,生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束。

[0014] 在一些实施例中,生成单元进一步被配置成:对第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络;或检测第一图像的纹理特征参数值是否大于纹理阈值,若是,对第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络。

[0015] 在一些实施例中,生成单元中所采用的述生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束包括:生成单元中所采用的生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束。

[0016] 在一些实施例中,生成单元中所采用的生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束包括:生成单元中所采用的生成式对抗网络将基于输入的图像生成的多通道人脸图像作为输入;其中,多通道人脸图像包括输入的图像的RGB三通道和输入的图像的以下至少一项:人脸关键点的二值图一通道或RGB三通道;人脸语义分割结果的二值图一通道或RGB三通道;以及头发的二值图一通道。

[0017] 在一些实施例中,生成单元中所采用的预先训练的生成式对抗网络用于训练的输入的图像样本包括:对初始图像样本进行数据增强后得到的图像样本。

[0018] 在一些实施例中,生成单元中所采用的预先训练的生成式对抗网络用于训练的输入的图像样本包括:对初始图像样本进行以下至少一项操作:旋转、翻转、缩放和不同程度的高斯模糊。

[0019] 在一些实施例中,生成式对抗网络至少包括以下任意一项:生成式对抗网络GAN、循环一致生成式对抗网络CycleGan、人脸高精度属性编辑模型AttGAN、星型生成式对抗网络StarGAN、空间变换器生成式对抗网络STGAN。

[0020] 在一些实施例中,第一图像为真实的人脸图像;第二图像为卡通图像。

[0021] 第三方面,本公开实施例提供了一种电子设备/终端/服务器,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上任一所述的生成图像的方法。

[0022] 第四方面,本公开实施例提供了一种计算机可读介质,其上存储有计算机程序,该程序被处理器执行时实现如上任一所述的生成图像的方法。

[0023] 本公开实施例提供的生成图像的方法和装置,首先获取包括第一人臉的第一图像;将第一图像输入预先训练的生成式对抗网络,得到生成式对抗网络输出的包括第二人脸的第二图像;其中,生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束。在这一过程中,可以提高生成式对抗网络基于输入的第一图像生成第二图像的准确率和效率,减少误生成图像的概率,并且减少了对输入的图像只能为预定区域内的图像的约束。

附图说明

[0024] 通过阅读参照以下附图所作的对非限制性实施例详细描述,本公开的其它特征、目的和优点将会变得更明显:

[0025] 图1是本公开可以应用于其中的示例性系统架构图;

[0026] 图2是根据本公开实施例的生成图像的方法的一个实施例的流程示意图;

[0027] 图3是根据本公开实施例的生成图像的方法的一个示例性应用场景;

[0028] 图4是根据本公开实施例的生成图像的方法的又一个实施例的流程示意图;

[0029] 图5a是根据本公开实施例的生成图像的方法的又一个实施例中输入预先训练的循环一致生成式对抗网络的第一图像的示例图;

[0030] 图5b是根据本公开实施例的基于图5a中的第一图像提取的采用RGB三通道表示的人脸关键点图像;

[0031] 图5c是根据本公开实施例的基于图5a中的第一图像提取的采用二值图一通道表示的头发的分割结果图像;

[0032] 图5d是根据本公开实施例的基于图5a中的第一图像提取的采用RGB三通道表示的人脸语义分割结果图像;

[0033] 图6是根据本公开实施例的生成图像的方法的预先训练的循环一致生成式对抗网络的结构示意图;

[0034] 图7是本公开的生成图像的装置的一个实施例的示例性结构图;

[0035] 图8是适于用来实现本公开实施例的服务器的计算机系统的结构示意图。

具体实施方式

[0036] 下面结合附图和实施例对本公开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。

[0037] 需要说明的是,在不冲突的情况下,本公开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公开。本领域技术人员还将理解的是,虽然本文中可使用用语“第一”、“第二”等来描述各种人脸、图像等,但是这些人脸、图像不应被这些用语限制。这些用语仅用于将一个人脸、图像与其它人脸、图像区分开。

[0038] 图1示出了可以应用本公开的生成图像的方法或生成图像的装置的实施例的示例性系统架构100。

[0039] 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。

[0040] 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上可以安装有各种通讯客户端应用,例如浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

[0041] 终端设备101、102、103可以是硬件,也可以是软件。当终端设备101、102、103为硬件时,可以是支持浏览器应用的各种电子设备,包括但不限于平板电脑、膝上型便携计算机和台式计算机等等。当终端设备101、102、103为软件时,可以安装在上述所列举的电子设备

中。其可以实现成例如用来提供分布式服务的多个软件或软件模块,也可以实现成单个软件或软件模块。在此不做具体限定。

[0042] 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上进行的浏览器应用提供支持的后台服务器。后台服务器可以对接收到的请求等数据进行分析等处理,并将处理结果反馈给终端设备。

[0043] 需要说明的是,服务器可以是硬件,也可以是软件。当服务器为硬件时,可以实现成多个服务器组成的分布式服务器集群,也可以实现成单个服务器。当服务器为软件时,可以实现成例如用来提供分布式服务的多个软件或软件模块,也可以实现成单个软件或软件模块。在此不做具体限定。

[0044] 在实践中,本公开实施例所提供的生成图像的方法可以由终端设备101、102、103和/或服务器105、106执行,生成图像的装置也可以设置于终端设备101、102、103和/或服务器105、106中。

[0045] 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。

[0046] 继续参考图2,图2示出了根据本公开的生成图像的方法的一个实施例的流程200。该生成图像的方法包括以下步骤:

[0047] 步骤201,获取包括第一人臉的第一图像。

[0048] 在本实施例中,生成图像的方法的执行主体(例如图1所示的终端或服务器)可以从本地或远程相册、数据库获取包括第一人臉的第一图像,也可以经由本地或远程的拍照服务获取包括第一人臉的第一图像。

[0049] 步骤202,将第一图像输入预先训练的生成式对抗网络,得到生成式对抗网络输出的包括第二人脸的第二图像;其中,生成式对抗网络将基于第一图像生成的人脸属性信息作为约束。

[0050] 在本实施例中,预先训练的生成式对抗网络,是指以生成式对抗网络(GAN, Generative Adversarial Networks)为基础发展而成的深度学习模型。例如,生成式对抗网络GAN、循环一致生成式对抗网络CycleGan、人脸高精度属性编辑模型AttGAN、星型生成式对抗网络StarGAN、空间变换器生成式对抗网络STGAN、对偶学习的生成对抗网络DualGAN、DiscoGAN等。

[0051] 该预先训练的生成式对抗网络,通常包括生成器G(generator)和鉴别器(Discriminator)。有两个数据域分别为X,Y。G负责将X域中的数据模仿成真实数据并把它们藏在真实数据中,而D负责将伪造数据和真实数据分开。经过二者的博弈以后,G的伪造技术越来越厉害,D的鉴别技术也越来越厉害。直到D再也分不出数据是真实的还是G生成的数据的时候,对抗的过程达到一个动态的平衡。

[0052] 训练生成式对抗网络需要两个损失函数:生成器的重建损失函数和判别器的判别损失函数。其中,重建损失函数用于确定生成的图片与原图是否尽可能的相似;判别损失函数用于将生成的假图片和原始真图片输入至判别器中,得到公式为0,1的二分类的损失。

[0053] 生成器由编码器、转换器和解码器构成。编码器采用卷积神经网络从输入的图象中提取特征。例如,将图像压缩成256个64*64的特征向量。转换器通过组合图像的不相近特征,将图像在DA域中的特征向量转换为DB域中的特征向量。例如,可以使用6层Reset模块,

每个Reset模块是一个由两个卷积层构成的神经网络层,从而达到在转换时同时保留原始图像特征的目标。解码器利用反卷积层(decovolution)完成从特征向量中还原出低级特征的工作,最后得到生成图像。

[0054] 鉴别器将一张图像作为输入,并尝试预测其为原始图像或是生成器的输出图像。鉴别器本身属于卷积网络,需要从图像中提取特征,再通过添加产生一维输出的卷积层来确定提取的特征是否属于特定类别。

[0055] 上述预先训练的生成式对抗网络,基于输入的图像生成的人脸属性信息作为约束。这里的人脸属性信息,是表征人脸特征的一系列生物特性信息,具有很强的自身稳定性和个体差异性,标识了人的身份。包括性别、肤色、年龄、表情等。

[0056] 生成式对抗网络在将基于输入的图像(例如应用生成式对抗网络时输入的第一图像或训练生成式对抗网络时输入的第一图像样本)生成的人脸属性信息作为约束时,可以采用现有技术或未来发展的技术中的用于机器学习网络的约束的方法来实现,本申请对此不做限定。

[0057] 在本申请的一个具体的示例中,上述基于输入的图像生成的人脸属性信息,可以作为生成式对抗网络中的生成器网络的任意一层或多层的输入,与该层网络原有的输入联合作为输入,从而在机器学习中改善引入约束的生成器网络的输出图像与人脸属性信息的关联性。

[0058] 在本申请的另一个具体的示例中,上述基于输入的图像生成的人脸属性信息,采用面部标志特征来定义一致性损失,指导在生成式对抗网络中训练鉴别器。首先,生成器基于输入的第一图像生成一张粗糙的第二图像。随后,生成式对抗网络根据生成的第二图像,生成一个预训练回归量(regressor),来预测面部界标landmark,对面部的关键点进行标记。最后,通过局部和全局两种鉴别器,细化在第二图像中对应第一图像的人脸特征。在这个阶段,强调了landmark的一致性,因此最后的生成结果逼真且有辨识度。

[0059] 本领域技术人员可以理解的是,生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束。其中,在训练该生成式对抗网络时,输入的图像可以为输入的第一图像样本。生成式对抗网络提取输入的第一图像样本的人脸属性信息,并将人脸属性信息作为生成式对抗网络的约束,得到输出的生成的图像。在应用该生成式对抗网络时,输入的图像可以为输入的包括第一人臉的第一图像。生成式对抗网络提取输入的包括第一人臉的第一图像的人脸属性信息,并将第一图像的人脸属性信息作为生成式对抗网络的约束,得到输出的包括第二人脸的第二图像。

[0060] 输入上述生成式对抗网络的第一图像,可以为包括人脸的图像。生成式对抗网络输出的第二图像,可以为风格或性别不同于第一图像且包括人脸的图像。

[0061] 在本实施例的一些可选的实现方式中,生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束包括:生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束。

[0062] 在本实现方式中,上述执行主体可以采用人脸关键点提取技术,提取输入的图像的人脸关键点,并将人脸关键点作为生成式对抗网络生成输出的图像时的约束。备选地或附加地,上述执行主体可以采用人脸语义分割技术,将人脸语义分割结果作为生成式对抗网络生成输出的图像时的约束。

[0063] 本实现方式中的生成式对抗网络,采用了基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束,可以使得生成器可以将输入的图像的五官和输出的图像的五官联系起来,从而不会在其它部位误生成人脸特征,使得输入的图像可以为包括人脸的更大的图像,而不局限于只输入人脸图像,并且提高了输出的图像的准确度和质量。

[0064] 在本实施例的一些可选实现方式中,将第一图像输入预先训练的生成式对抗网络包括:对第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络。

[0065] 在本实现方式中,高斯模糊(Gaussian Blur),也叫高斯平滑,可以减少图像噪声以及降低细节层次,增强图像在不同比例大小下的图像效果(参见尺度空间表示以及尺度空间实现)。从数学的角度来看,图像的高斯模糊过程就是图像与正态分布做卷积。由于正态分布又称为高斯分布,所以这项技术又称为高斯模糊。

[0066] 通过采用不同程度的高斯模糊后的第一图像作为生成式对抗网络的输入,可以得到不同清晰度的第二图像,从而可以在不同清晰度的第二图像的基础上,确定所需的清晰的第二图像。

[0067] 备选地或附加地,将第一图像输入预先训练的生成式对抗网络包括:检测第一图像的纹理特征参数值是否大于纹理阈值,若是,对第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络。

[0068] 在这里,图像的纹理特征参数值是指用于表征图像的纹理特征的值。例如,纹理的粗细度、密度及方向性等。当检测到第一图像的纹理特征参数值大于纹理阈值时,表明第一图像的纹理复杂。通常,基于纹理复杂的图像生成的图像内容较乱。因此,可以对第一图像进行不同程度的高斯模糊,从而对于不同程度的高斯模糊后的第一图像,分别生成第二图像,从而得到不同清晰度的第二图像。之后,可以在不同清晰度的第二图像的基础上,确定所需的清晰的第二图像,从而提高生成的图像的质量。

[0069] 上述的预先训练的生成式对抗网络,可以采用爬虫从网络爬取图像样本或直接获取图像样本数据集,这些图像中的每张图像都包括人脸;之后,可以直接将爬取或获取的图像样本直接作为训练生成式对抗网络的图像样本;或者,将爬取的包括人脸的图像作为初始图像样本,对初始图像样本进行进一步的数据处理,得到符合图像样本要求的筛选后的图像,并将筛选后的图像作为训练生成式对抗网络的图像样本。

[0070] 在本实施例的一些可选实现方式中,预先训练的生成式对抗网络用于训练的输入的图像样本包括:对初始图像样本进行数据增强后得到的图像样本。

[0071] 在本实现方式中,上述数据增强可以包括如旋转、平移、翻折等操作。预先训练的生成式对抗网络,由于训练数据需要风格一致且包含不同性别、不同角度、不同面部大小的图像样本,因此可以对初始图像样本进行数据增强,从而增加训练的数据量,提高生成式对抗网络的泛化能力,并且增加噪声数据,提升生成式对抗网络的鲁棒性。

[0072] 在本实施例的一些可选实现方式中,对初始图像样本进行数据增强包括:对初始图像样本进行以下至少一项操作:旋转、翻转、缩放和不同程度的高斯模糊。

[0073] 在本实现方式中,通过对初始图像样本进行旋转、翻转、缩放和不同程度的高斯模糊中的一项或多项操作,可以提升训练完成的生成式对抗网络的泛化能力和鲁棒性。

[0074] 本公开上述实施例的生成图像的方法,由于在生成式对抗网络基于第一图像生成

第二图像的过程中,采用了第一图像的人脸属性信息作为生成式对抗网络的约束,可以提高生成式对抗网络基于输入的第一图像生成第二图像的质量、准确率和效率,减少误生成图像的概率,并且减少了对输入的图像只能为预定区域内的图像的约束。

[0075] 以下结合图3,描述本公开的生成图像的方法的示例性应用场景。

[0076] 如图3所示,图3示出了根据本公开的生成图像的方法的一个示例性应用场景。

[0077] 如图3所示,生成图像的方法300运行于电子设备310中,可以包括:

[0078] 首先,获取包括第一人脸301的第一图像302;

[0079] 之后,基于第一图像302中的人脸301,生成人脸属性信息303;

[0080] 最后,将第一图像302输入预先训练的生成式对抗网络304,并将人脸属性信息303作为生成式对抗网络304的约束,得到生成式对抗网络304输出的包括第二人脸305的第二图像306。

[0081] 应当理解,上述图3中所示出的生成图像的方法的应用场景,仅为对于生成图像的方法的示例性描述,并不代表对该方法的限定。例如,上述图3中示出的各个步骤,可以进一步采用更为细节的实现方法。也可以在上述图3的基础上,进一步增加其它生成图像的步骤。

[0082] 进一步参考图4,图4示出了根据本公开的生成图像的方法中上传文件的方法的一个实施例的示意性流程图。

[0083] 如图4所示,本实施例的生成图像的方法中上传文件的方法400,可以包括以下步骤:

[0084] 步骤401,获取包括第一人脸的第一图像。

[0085] 在本实施例中,生成图像的方法的执行主体(例如图1所示的终端或服务器)可以从本地或远程相册、数据库获取包括第一人脸的第一图像,也可以经由本地或远程的拍照服务获取包括第一人脸的第一图像。

[0086] 步骤402,基于第一图像,获取第一图像的人脸关键点的RGB三通道图像。

[0087] 在本实施例中,可以采用现有技术或未来发展的技术中的提取人脸关键技术来基于第一图像获取第一图像的人脸关键点图像,本申请对此不做限定。例如,可以采用主动形状模型(ASM,Active Shape Model)、主动外观模型(AMM,Active Appearance Model,)、级联形状回归模型Cascaded Shape Regression(CSR)、或者人脸对齐算法(DAN,Deep Alignment Network)等进行人脸关键点提取。

[0088] 步骤403,基于第一图像,获取第一图像的头发的二值图。

[0089] 在本实施例中,可以采用现有技术或未来发展的技术中的头发分割技术的技术来基于第一图像获取第一图像的头发的二值图,本申请对此不做限定。例如,可以采用分割头发的技术先获取头发的分割结果,再将头发分割结果转换为头发的二值图。

[0090] 步骤404,将第一图像的RGB三通道图像、人脸关键点的RGB三通道图像和头发的二值图拼接为多通道人脸图像。

[0091] 在本实施例中,多通道人脸图像为包括第一图像的RGB三通道和第一图像的以下各项的人脸图像:人脸关键点RGB三通道和头发的二值图一通道。

[0092] 其中,图像的RGB三通道,是指每个像素点都由RGB三个值表示,三通道分别为:红色通道R(red),绿色通道G(green),蓝色通道B(blue)。

[0093] 第一图像的人脸关键点RGB三通道,是指从第一图像中提取人脸关键点,并且人脸关键点特征图的每个像素点由RGB三通道表示。在一个具体的示例中,输入的第一图像为图5a,可以提取如图5b所示的采用RGB三通道表示的人脸关键点。

[0094] 输入的图像的头发的二值图一通道,是指采用头发分割技术得到头发的二值图,并且头发分割结果的每个像素点由二值图一通道表示。二值图是指图像中,灰度等级只有两种,也就是说,图像中的任何像素点的灰度值均为0或者255,分别代表黑色和白色。在一个具体的示例中,输入的图像为图5a,可以得到如图5c所示的采用二值图一通道表示的头发的分割结果。

[0095] 步骤405,将多通道人脸图像输入预先训练的循环一致生成式对抗网络,得到预先训练的循环一致生成式对抗网络输出的包括第二人脸的第二图像。

[0096] 在本实施例中,可以将多通道人脸图像输入预先训练的循环一致生成式对抗网络(CycleGan),从而得到循环一致生成式对抗网络输出的第二图像。

[0097] 在一个具体的示例中,第一图像为真实的人脸图像;第二图像为卡通图像。这里的卡通图像可以为壁画、油画、地毯等的草图、底图,也可以为漫画、讽刺画、幽默画等。

[0098] 本领域技术人员可以理解的是,预先训练循环一致生成式对抗网络时所采用的多通道人脸图像样本,与应用该循环一致生成式对抗网络时输入的多通道人脸图像相适应,以便得到训练完成的、适合应用时输入的多通道人脸图像的循环一致生成式对抗网络。

[0099] 在一个具体的示例中,预先训练的循环一致生成式对抗网络的结构如图6所示,包括两个生成器G_A2B、G_B2A和两个鉴别器D_A和D_B。其中,生成器G_A2B用于基于输入的真实人图像生成卡通,生成器G_B2A用于基于卡通图像生成真人。鉴别器D_A用于根据输出结果为(0,1)判断是否是真正的人脸图像,鉴别器D_B用于根据输出结果为(0,1)判断是否是真正的卡通图像。

[0100] 训练的时候,两组生成器与鉴别器的组合网络交替训练,训练完成后采用真人到卡通的G_A2B生成器上线。

[0101] 本领域技术人员可以理解的是,上述图4中所示的实施例,仅为本申请的示例性实施例,并不代表对本申请的限定。

[0102] 例如,上述的预先训练的循环一致生成式对抗网络,也可以为基于GAN改进的其它生成式对抗网络。

[0103] 又例如,上述图4中的多通道人脸图像,可以为在包括第一图像的RGB三通道的的基础上,还包括第一图像的人脸语义分割结果RGB三通道。

[0104] 其中,第一图像的人脸语义分割结果RGB三通道,是指从输入的图像中提取人脸语义分割结果,并且人脸语义分割结果的每个像素点由RGB三通道表示。在一个具体的示例中,输入的图像为图5a,可以提取如图5d所示的采用RGB三通道表示的人脸语义分割结果。

[0105] 在这里,第一图像的人脸语义分割结果图像,语义分割是目标检测的扩展,其输出是目标的按类别区分的彩色掩码,可以对目标实现更精准的定位,并且不受目标的复杂形状所影响。

[0106] 在进行人脸语义分割时,可以采用现有技术或未来发展的技术中的分割人脸语义的技术来基于第一图像获取第一图像的人脸语义分割结果,本申请对此不做限定。在一些具体的示例中,分割人脸语义的技术可以采用全卷积神经网络FCN(Fully Convolutional

Networks for Semantic Segmentation)、语义分割网络(SegNet, Semantic Segmentation)、空洞卷积(Dilated Convolutions)、语义分割(DeepLab(v1、v2、V3等))、图像分割模型(RefineNet)、金字塔场景解析网络(PSPNet)等来基于第一图像获取第一图像的人脸语义分割结果。

[0107] 又例如,上述图4中的多通道人脸图像,可以在包括第一图像的RGB三通道的基礎上,可以进一步包括人脸语义分割结果的二值图一通道。备选地,上述图4中的多通道人脸图像,可以在包括第一图像的RGB三通道的基礎上,进一步包括第一图像的人脸关键点的二值图一通道和头发的二值图一通道。备选地,上述图4中的多通道人脸图像,也可以在包括第一图像的RGB三通道的基礎上,进一步包括第一图像的人脸关键点RGB三通道和人脸语义分割结果的二值图一通道。在此不再一一赘述。

[0108] 本公开图4所示的实施例中的生成图像的方法,在图2中所示的生成图像的方法的基础上,进一步公开了将多通道人脸图像输入预先训练的循环一致生成式对抗网络,得到预先训练的循环一致生成式对抗网络输出的包括第二人脸的第二图像。这一过程中,由于采用了多通道人脸图像作为生成第二图像的基准,提高了生成的第二图像的准确性和质量。

[0109] 作为对上述各图所示方法的实现,本公开实施例提供了一种生成图像的装置的一个实施例,该装置实施例与图2-图6中所示的方法实施例相对应,该装置具体可以应用于上述图1所示的终端或服务器中。

[0110] 如图7所示,本实施例的生成图像的装置700可以包括:获取单元710,被配置成获取包括第一人脸的第一图像;生成单元720,被配置成将第一图像输入预先训练的生成式对抗网络,得到生成式对抗网络输出的包括第二人脸的第二图像;其中,生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束。

[0111] 在一些实施例中,生成单元720进一步被配置成:对第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络;或检测第一图像的纹理特征参数值是否大于纹理阈值,若是,对第一图像进行不同程度的高斯模糊,将不同程度的高斯模糊后的第一图像输入预先训练的生成式对抗网络。

[0112] 在一些实施例中,生成单元720中所采用的述生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束包括:生成单元中所采用的生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束。

[0113] 在一些实施例中,生成单元720中所采用的生成式对抗网络将基于输入的图像生成的人脸关键点和/或人脸语义分割结果作为约束包括:生成单元中所采用的生成式对抗网络将基于输入的图像生成的多通道人脸图像作为输入;其中,多通道人脸图像包括输入的图像的RGB三通道和输入的图像的以下至少一项:人脸关键点的二值图一通道或RGB三通道;人脸语义分割结果的二值图一通道或RGB三通道;以及头发的二值图一通道。

[0114] 在一些实施例中,生成单元720中所采用的预先训练的生成式对抗网络用于训练的输入的图像样本包括:对初始图像样本进行数据增强后得到的图像样本。

[0115] 在一些实施例中,生成单元720中所采用的预先训练的生成式对抗网络用于训练的输入的图像样本包括:对初始图像样本进行以下至少一项操作:旋转、翻转、缩放和不同程度的高斯模糊。

[0116] 在一些实施例中,生成式对抗网络至少包括以下任意一项:生成式对抗网络GAN、循环一致生成式对抗网络CycleGan、人脸高精度属性编辑模型AttGAN、星型生成式对抗网络StarGAN、空间变换器生成式对抗网络STGAN。

[0117] 在一些实施例中,第一图像为真实的人脸图像;第二图像为卡通图像。

[0118] 应当理解,装置700中记载的各个单元与参考图2-图5描述的方法中记载的各个步骤相对应。由此,上文针对方法描述的操作和特征同样适用于装置700及其中包含的各个单元,在此不再赘述。

[0119] 下面参考图8,其示出了适于用来实现本公开的实施例的电子设备(例如图1中的服务器或终端设备)800的结构示意图。本公开的实施例中的终端设备可以包括但不限于诸如笔记本电脑、台式计算机等。图8示出的终端设备/服务器仅仅是一个示例,不应对本公开的实施例的功能和使用范围带来任何限制。

[0120] 如图8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(ROM)802中的程序或者从存储装置808加载到随机访问存储器(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RAM803通过总线804彼此相连。输入/输出(I/O)接口805也连接至总线804。

[0121] 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(LCD)、扬声器、振动器等的输出装置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8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。图8中示出的每个方框可以代表一个装置,也可以根据需要代表多个装置。

[0122] 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,执行本公开的实施例的方法中限定的上述功能。需要说明的是,本公开的实施例所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(EPR0M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。在本公开的实施例中,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。而在本公开的实施例中,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。

[0123] 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：获取包括第一人臉的第一图像；将第一图像输入预先训练的生成式对抗网络，得到生成式对抗网络输出的包括第二人脸的第二图像；其中，生成式对抗网络将基于输入的图像生成的人脸属性信息作为约束。

[0124] 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实施例的操作的计算机程序代码，所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)。

[0125] 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，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。

[0126] 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。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和生成单元。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“获取包括第一人臉的第一图像的单元”。

[0127] 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。

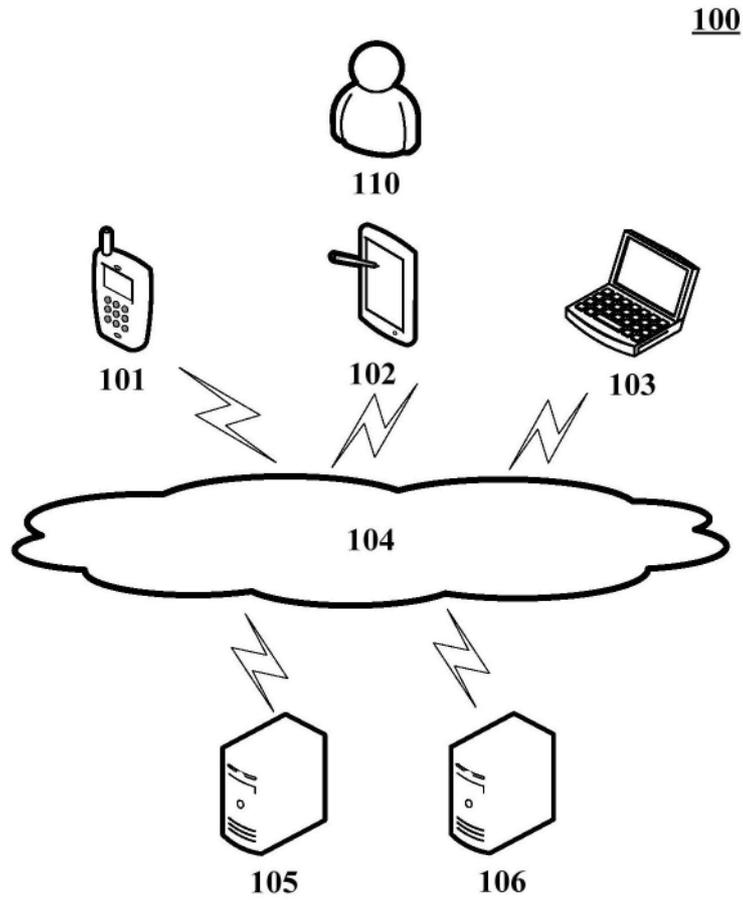


图1

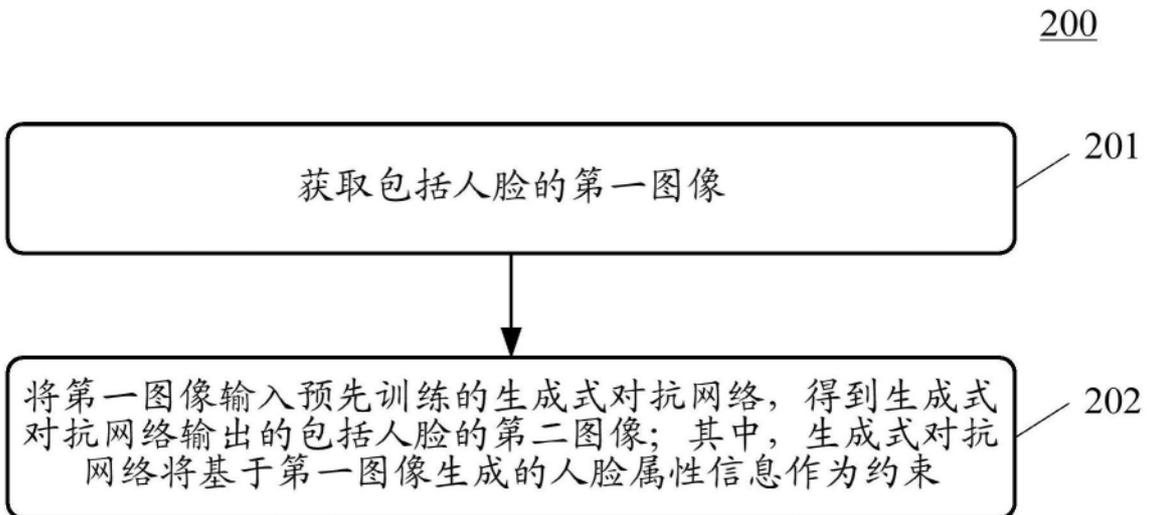


图2

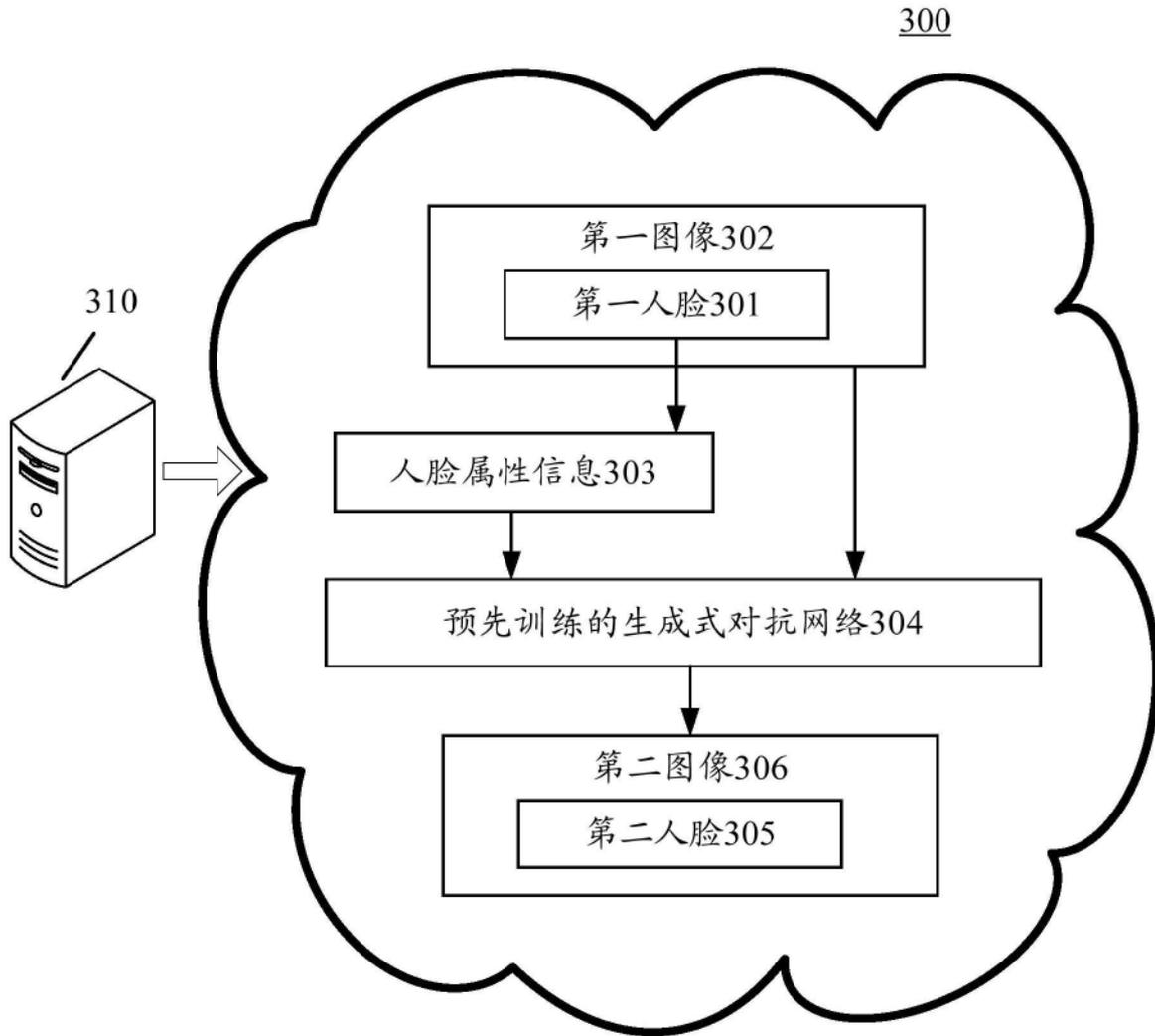


图3

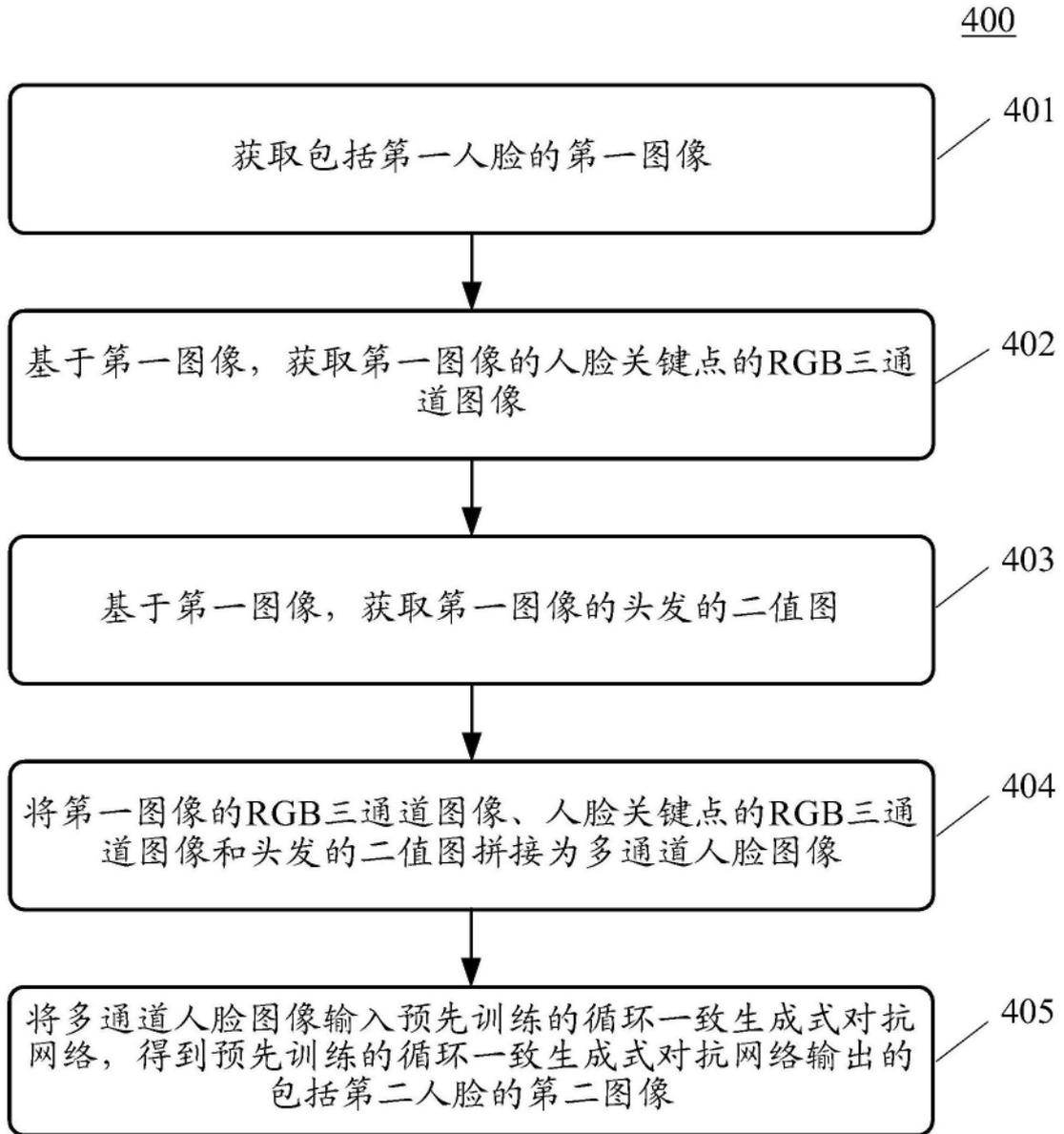


图4



图5a



图5b



图5c



图5d

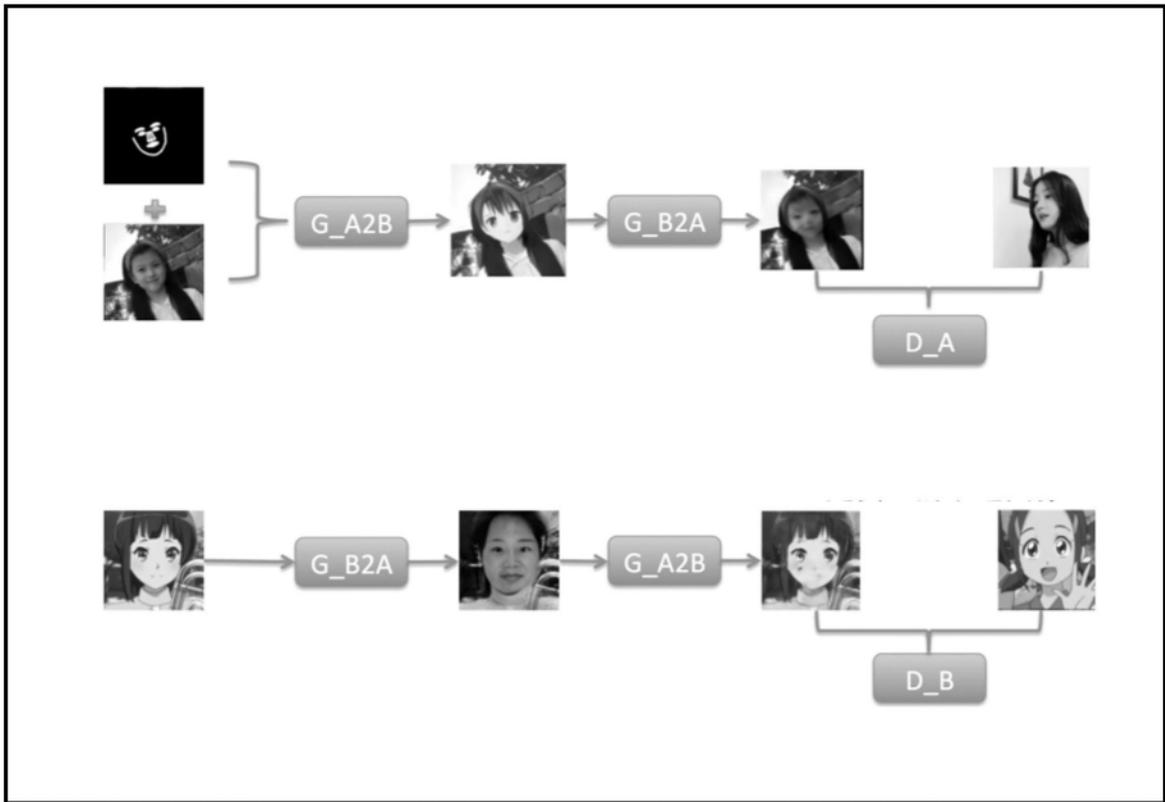


图6



图7

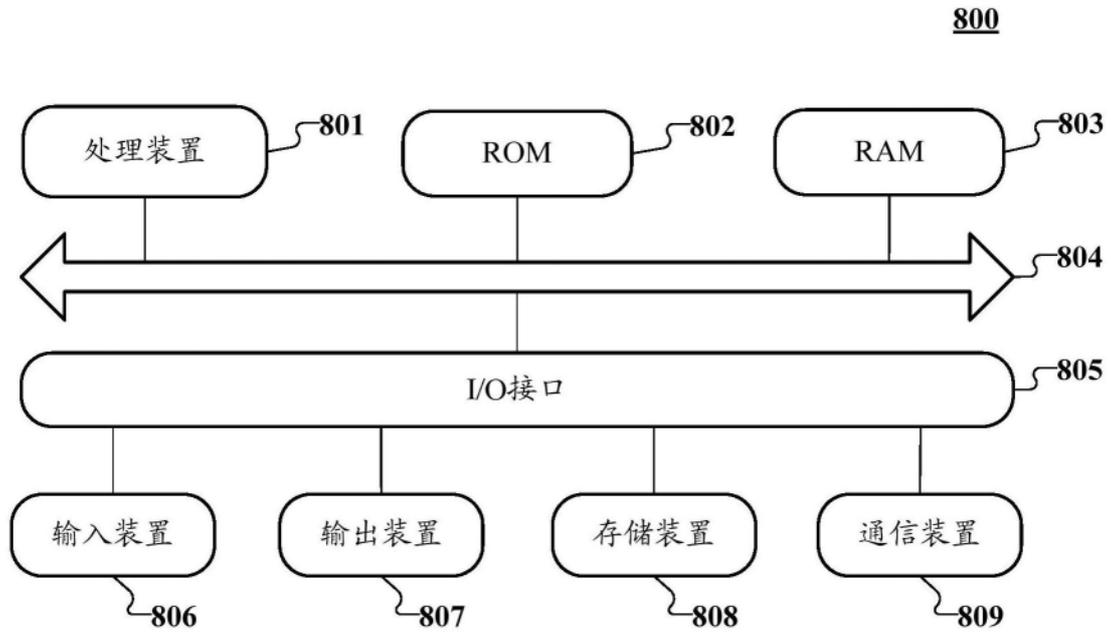


图8