



(12)发明专利申请

(10)申请公布号 CN 111669647 A

(43)申请公布日 2020.09.15

(21)申请号 202010537321.0

(22)申请日 2020.06.12

(71)申请人 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 李甫 林天威 何栋梁
张赫男 孙昊 文石磊 丁二锐

(74)专利代理机构 北京品源专利代理有限公司
11332
代理人 孟金喆

(51)Int.Cl.
H04N 21/4402(2011.01)
G06K 9/00(2006.01)
G06K 9/62(2006.01)

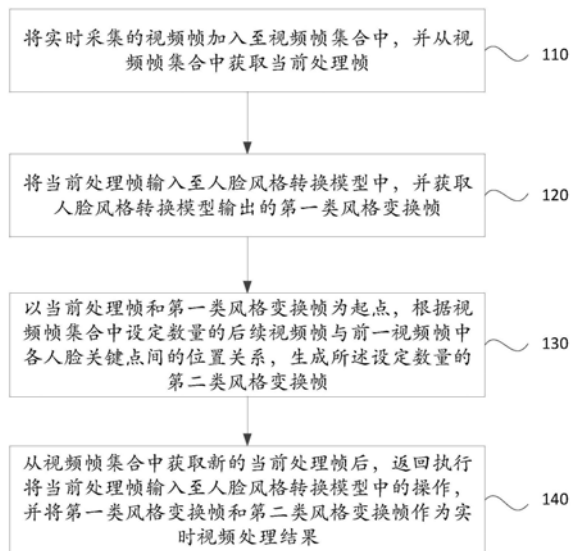
权利要求书3页 说明书12页 附图4页

(54)发明名称

实时视频的处理方法、装置、设备及存储介质

(57)摘要

本申请公开了一种实时视频的处理方法、装置、设备及存储介质,涉及数字图像处理、深度学习技术领域。具体实现方案为:将视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧;将当前处理帧输入至人脸风格转换模型中,获取模型输出的第一类风格变换帧;以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成对应的第二类风格变换帧;从视频帧集合中获取新的当前处理帧后,返回执行将当前处理帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果。本申请实施例的技术方案可以基于视频中的真实人脸实时生成匹配的风格人脸。



1. 一种实时视频的处理方法,包括:

将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧,视频帧中包括真实人脸;

将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧,风格变换帧中包括风格人脸;

以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧;

从视频帧集合中获取新的当前处理帧后,返回执行将当前处理帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果。

2. 根据权利要求1所述的方法,其中,以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧,包括:

将当前处理帧作为处理起点帧,在视频帧集合中,获取处理起点帧的下一视频帧;

根据下一视频帧与处理起点帧中各人脸关键点分别在对应视频帧中的图像位置,生成人脸关键点变换矩阵;

根据人脸关键点变换矩阵,和与处理起点帧匹配的第一类风格变换帧或第二类风格变换帧,生成下一视频帧的第二类风格变换帧;

将所述下一视频帧作为处理起点帧,返回执行在视频帧集合中,获取处理起点帧的下一视频帧的操作,直至下一视频帧的处理数量达到所述设定数量。

3. 根据权利要求1所述的方法,在将当前处理帧输入至人脸风格转换模型中之前,还包括:

获取训练样本集,训练样本集中的包括多个样本图像对,每个样本图像对中包括原始图像以及变换后图像;

使用训练样本集中各样本图像对训练设定机器学习模型,得到所述人脸风格转换模型;

其中,所述原始图像中包括真实人脸,所述变换后图像中包括与真实人脸匹配的风格人脸。

4. 根据权利要求3所述的方法,其中,所述机器学习模型包括:生成对抗网络。

5. 根据权利要求1-4任一项所述的方法,其中,将实时采集的视频帧加入至视频帧集合中,包括:

在视频录制过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中;

其中,将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果,包括:

对第一类风格变换帧和第二类风格变换帧进行实时录制播放,并生成录制视频。

6. 根据权利要求5所述的方法,其中,对第一类风格变换帧和第二类风格变换帧进行实时录制播放,包括:

将实时生成的第一类风格变换帧和第二类风格变换帧顺序存储于设定缓存队列中;

当所述缓存队列中满足预设的硬延时条件时,从所述缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,进行实时录制播放。

7. 根据权利要求1-4任一项所述的方法,其中,将实时采集的视频帧加入至视频帧集合中,包括:

在视频直播过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中;

其中,将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果,包括:

根据第一类风格变换帧和第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器,以进行视频直播。

8. 根据权利要求7所述的方法,其中,根据第一类风格变换帧和第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器,包括:

将实时生成的第一类风格变换帧和第二类风格变换帧顺序存储于设定缓存队列中;

当所述缓存队列中满足预设的硬延时条件时,从所述缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器。

9. 一种实时视频的处理装置,包括:

获取模块,用于将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧,视频帧中包括真实人脸;

第一变换模块,用于将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧,风格变换帧中包括风格人脸;

第二变换模块,用于以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧;

循环模块,用于从视频帧集合中获取新的当前处理帧后,返回执行将当前图像帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果。

10. 根据权利要求9所述的装置,其中,第二变换模块,包括:

后一视频帧获取单元,用于将当前处理帧作为处理起点帧,在视频帧集合中,获取处理起点帧的后一视频帧;

变换矩阵生成单元,用于根据后一视频帧与处理起点帧中各人脸关键点分别在对应视频帧中的图像位置,生成人脸关键点变换矩阵;

第二类风格变换帧生成单元,用于根据人脸关键点变换矩阵,和与处理起点帧匹配的第一类风格变换帧或第二类风格变换帧,生成后一视频帧的第二类风格变换帧;

循环操作单元,用于将所述后一视频帧作为处理起点帧,返回执行在视频帧集合中,获取处理起点帧的后一视频帧的操作,直至后一视频帧的处理数量达到所述设定数量。

11. 根据权利要求9所述的装置,其中,还包括:

模型训练模块,用于在将当前处理帧输入至人脸风格转换模型中之前,获取训练样本集,训练样本集中的包括多个样本图像对,每个样本图像对中包括原始图像以及变换后图像;

使用训练样本集中各样本图像对训练设定机器学习模型,得到所述人脸风格转换模型;

其中,所述原始图像中包括真实人脸,所述变换后图像中包括与真实人脸匹配的风格

人脸。

12. 根据权利要求11所述的装置,其中,所述机器学习模型包括:生成对抗网络。

13. 根据权利要求9-12任一项所述的装置,其中,获取模块,包括:

录制获取单元,用于在视频录制过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中;

其中,循环模块,包括:

录制播放单元,用于对第一类风格变换帧和第二类风格变换帧进行实时录制播放,并生成录制视频。

14. 根据权利要求13所述的装置,其中,录制播放单元,具体用于:

将实时生成的第一类风格变换帧和第二类风格变换帧顺序存储于设定缓存队列中;

当所述缓存队列中满足预设的硬延时条件时,从所述缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,进行实时录制播放。

15. 根据权利要求9-12任一项所述的装置,其中,获取模块,包括:

直播获取单元,用于在视频直播过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中;

其中,循环模块,包括:

直播播放单元,用于根据第一类风格变换帧和第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器,以进行视频直播。

16. 根据权利要求15所述的装置,其中,直播播放单元,具体用于:

将实时生成的第一类风格变换帧和第二类风格变换帧顺序存储于设定缓存队列中;

当所述缓存队列中满足预设的硬延时条件时,从所述缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器。

17. 一种电子设备,其特征在于,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,

所述存储器存储有可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权利要求1-8中任一项所述的方法。

18. 一种存储有计算机指令的非瞬时计算机可读存储介质,其特征在于,所述计算机指令用于使所述计算机执行权利要求1-8中任一项所述的方法。

实时视频的处理方法、装置、设备及存储介质

技术领域

[0001] 本申请实施例涉及图像处理、深度学习技术,尤其涉及数字图像处理技术,具体涉及一种实时视频的处理方法、装置、设备及存储介质。

背景技术

[0002] 随着生活水平的不断提高,用户对娱乐化的需求越来越多样化,将视频中的真实人脸变换成动漫风格人脸也受到越来越多用户的关注和喜爱。

[0003] 现有技术中,通常在离线状态下,基于视频中的真实人脸生成匹配的风格人脸,或者,使用视频中的真实人脸实时驱动一个已经预先生成的固定风格人脸,生成与真实人脸表情一致的风格人脸。但这两种方式都不能基于视频中的真实人脸,实时生成与该真实人脸匹配的风格人脸。

发明内容

[0004] 本申请实施例提供了一种实时视频的处理方法、装置、设备及存储介质,实现了基于视频中的真实人脸,实时生成与真实人脸匹配的风格人脸。

[0005] 第一方面,本申请实施例提供了一种实时视频的处理方法,包括:

[0006] 将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧,视频帧中包括真实人脸;

[0007] 将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧,风格变换帧中包括风格人脸;

[0008] 以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧;

[0009] 从视频帧集合中获取新的当前处理帧后,返回执行将当前处理帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果。

[0010] 第二方面,本申请实施例还提供了一种实时视频的处理装置,包括:

[0011] 获取模块,用于将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧,视频帧中包括真实人脸;

[0012] 第一变换模块,用于将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧,风格变换帧中包括风格人脸;

[0013] 第二变换模块,用于以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧;

[0014] 循环模块,用于从视频帧集合中获取新的当前处理帧后,返回执行将当前图像帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果。

[0015] 第三方面,本申请实施例还提供了一种电子设备,包括:

[0016] 至少一个处理器;以及

[0017] 与所述至少一个处理器通信连接的存储器;其中,

[0018] 所述存储器存储有可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本申请任意实施例提供的实时视频的处理方法。

[0019] 第四方面,本申请实施例还提供了一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使所述计算机执行本申请任意实施例提供的实时视频的处理方法。

[0020] 本申请实施例的技术方案,通过将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧,视频帧中包括真实人脸;将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧,风格变换帧中包括风格人脸;以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧;从视频帧集合中获取新的当前处理帧后,返回执行将当前处理帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果,实现了基于视频中的真实人脸,实时生成与真实人脸匹配的风格人脸,解决了现有技术中无法实时生成与视频中的真实人脸匹配的风格人脸的问题。

[0021] 应当理解,本部分所描述的内容并非旨在标识本申请的实施例的关键或重要特征,也不用于限制本申请的范围。本申请的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0022] 附图用于更好地理解本方案,不构成对本申请的限定。其中:

[0023] 图1是本申请第一实施例中的一种实时视频的处理方法的流程图;

[0024] 图2a是本申请第二实施例中的一种实时视频的处理方法的流程图;

[0025] 图2b是本申请第二实施例中的一种实时视频的处理流程图;

[0026] 图3是本申请第三实施例中的一种实时视频的处理装置的结构示意图;

[0027] 图4是用来实现本申请实施例的实时视频的处理方法的电子设备的框图。

具体实施方式

[0028] 以下结合附图对本申请的示范性实施例做出说明,其中包括本申请实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本申请的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0029] 第一实施例

[0030] 图1是本申请第一实施例中的一种实时视频的处理方法的流程图,本实施例的技术方案适用于实时生成与视频中真实人脸匹配的风格人脸的情况,该方法可以由实时视频的处理执行,该装置可以通过软件,和/或硬件的方式实现,并一般可以集成在电子设备中,例如终端设备中,本实施例的方法具体包括以下步骤:

[0031] 步骤110、将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧。

[0032] 其中,视频帧中包括真实人脸。

[0033] 本实施例中,视频帧可以是终端设备拍摄视频过程中采集的某一帧图像,示例性的,视频帧可以是视频录制过程中实时采集的一帧图像,也可以是视频直播过程中实时采集的一帧图像。视频帧集合是用于存储实时采集的视频帧的集合,通常一个视频的全部视频帧放入至一个视频帧集合中,以通过不同的视频帧集合区分不同视频的视频帧。

[0034] 可选的,将实时采集的视频帧加入至视频帧集合中,可以包括:在视频录制过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中。

[0035] 本可选的实施例中,视频录制过程是指只进行视频拍摄,不会将拍摄的内容实时播出的过程,在视频录制过程中,如果用户对当前拍摄页面中的人脸风格变换选项进行点击操作,则响应于人脸风格变换请求,将用户点击人脸风格变换选项之后,实时采集的视频帧加入至视频帧集合中,或者,将从视频录制开始,所有实时采集的视频帧都加入至视频帧集合中。

[0036] 可选的,将实时采集的视频帧加入至视频帧集合中,可以包括:在视频直播过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中。

[0037] 本可选的实施例中,视频直播过程是指在同时进行视频拍摄的同时,将拍摄的内容实时播出的过程,在视频直播过程中,如果用户对当前直播页面中的人脸风格变换选项进行点击操作,则响应于人脸风格变换请求,将用户点击人脸风格变换选项之后,实时采集的视频帧加入至视频帧集合中,或者,将从视频录制开始,所有实时采集的视频帧都加入至视频帧集合中。

[0038] 本实施例中,将实时采集的视频帧加入至视频帧集合中之后,按照视频帧的采集时间的先后顺序,从视频帧集合中获取当前采集时间最前,并且尚未进行人脸风格变换处理的一个视频帧,作为当前处理帧。

[0039] 步骤120、将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧。

[0040] 其中,风格变换帧中包括风格人脸。

[0041] 本实施例中,人脸风格转换模型用于根据输入的视频帧中的真实人脸,生成与真实人脸匹配的风格人脸,并输出包括风格人脸的第一类风格变换帧。其中,第一类风格变换帧是将当前处理帧中的真实人脸替换为风格人脸后的视频帧,每个风格人脸只与唯一的真实人脸匹配,即每个真实人脸有专属的风格人脸,两者的尺寸大小以及包括的人脸特征都一致。

[0042] 步骤130、以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧。

[0043] 本实施例中,考虑到单独采用人脸风格转换模型对实时采集的所有视频帧中的真实人脸进行风格转换,计算量比较大,耗时比较长,无法达到实时的效果,因此,将根据人脸风格转换模型对视频帧中的真实人脸进行风格转换的方式,与根据相邻两视频帧中各人脸关键点间的位置变化量,来对视频帧中的真实人脸进行风格转换的方式相结合,通过这两

种方式交替使用,实时生成与视频帧中的真实人脸匹配的风格人脸。

[0044] 本实施例中,可以在通过人脸风格转换模型得到与当前处理帧匹配的第一类风格变换帧后,根据相邻两视频帧中各人脸关键点间的位置变化量,生成与视频帧集合中在当前处理帧之后的设定数量的后续视频帧分别匹配的设定数量的第二类风格变换帧,然后重复上述过程,生成与视频帧集合中后续的视频帧分别对应的第一类风格变换帧或第二类风格变换帧。

[0045] 可选的,以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧,可以包括:将当前处理帧作为处理起点帧,在视频帧集合中,获取处理起点帧的后一视频帧;根据后一视频帧与处理起点帧中各人脸关键点分别在对应视频帧中的图像位置,生成人脸关键点变换矩阵;根据人脸关键点变换矩阵,和与处理起点帧匹配的第一类风格变换帧或第二类风格变换帧,生成后一视频帧的第二类风格变换帧;将所述后一视频帧作为处理起点帧,返回执行在视频帧集合中,获取处理起点帧的后一视频帧的操作,直至后一视频帧的处理数量达到所述设定数量。

[0046] 本可选的实施例中,提供了一种具体的以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧的方式,具体过程如下:

[0047] 首先,以当前处理帧为处理起点帧,并从视频帧集合中获取处理起点帧的后一视频帧,然后,根据后一视频帧中各人脸关键点在视频帧中的位置坐标,与处理起点帧中各人脸关键点在视频帧中的位置坐标,计算出相邻两视频帧中各人脸关键点之间的坐标变化量,进而得到人脸关键点变换矩阵,接着,根据人脸关键点变换矩阵,对与处理起点帧匹配的第一类风格变换帧中的风格人脸的各人脸关键点的坐标进行调整,生成后一视频帧的第二类风格变换帧,然后,判断连续生成的第二类风格变换帧的数量是否等于设定数量,若是,则停止对视频帧集合中后一视频帧之后的视频帧进行操作,若不是,则将后一视频帧作为处理起点帧,返回执行在视频帧集合中获取处理起点帧的后一视频帧的操作,继续根据与相邻两视频帧对应的人脸关键点变换矩阵,对与处理起点帧匹配的第二类风格变换帧中的风格人脸的各人脸关键点的坐标进行调整,生成后一视频帧的第二类风格变换帧。

[0048] 需要说明的是,本申请实施例中的设定数量可以是1、2或者其他根据需求设定的数值。由于根据人脸关键点变换矩阵生成第二类风格变换帧时,只需对与前一视频帧匹配的风格变换帧的风格人脸进行人脸关键点的位置调整,计算量较少,耗时较短,因此,设定数量越大,对视频帧集合中的视频帧实现人脸风格转换的速度越快。但为了维持较高的风格人脸与真实人脸的匹配度,设定数量不宜过大。

[0049] 步骤140、从视频帧集合中获取新的当前处理帧后,返回执行将当前处理帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果。

[0050] 本实施例中,当根据与相邻两视频帧对应的人脸关键点变换矩阵,连续生成第二类风格变换帧的数量等于设定数量之后,需要从视频帧集合中获取新的当前处理帧,并返回执行将当前处理帧输入至人脸风格转换模型中的操作,重新通过人脸风格转换模型得到与当前处理帧匹配的第一类风格变换帧,以及,根据与相邻两视频帧对应的人脸关键点变

换矩阵,生成与视频帧集合中在当前处理帧之后的设定数量的后续视频帧分别匹配的设定数量的第二类风格变换帧。

[0051] 可选的,将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果,可以包括:对第一类风格变换帧和第二类风格变换帧进行实时录制播放,并生成录制视频。

[0052] 本可选的实施例中,如果视频帧集合中包括的视频帧是在视频录制过程中实时采集的,则对根据人脸风格转换模型生成的第一类风格变换帧,以及,根据人脸关键点变换矩阵生成的第二类风格变换帧,进行实时录制播放,并生成录制视频。

[0053] 可选的,将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果,可以包括:根据第一类风格变换帧和第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器,以进行视频直播。

[0054] 本可选的实施例中,如果视频帧集合中包括的视频帧是在视频直播过程中实时采集的,则根据通过人脸风格转换模型生成的第一类风格变换帧,以及,根据人脸关键点变换矩阵生成的第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器,以进行视频直播。

[0055] 本申请实施例的技术方案,通过将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取包括真实人脸的当前处理帧;将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧,风格变换帧中包括风格人脸;以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧;从视频帧集合中获取新的当前处理帧后,返回执行将当前处理帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果,实现了基于视频中的真实人脸,实时生成与真实人脸匹配的风格人脸,解决了现有技术中无法实时生成与视频中的真实人脸匹配的风格人脸的问题。

[0056] 第二实施例

[0057] 图2a是本申请第二实施例中的一种实时视频的处理方法的流程图,本实施例在上述实施例的基础上进一步细化,提供在将当前处理帧输入至人脸风格转换模型中之前,确定人脸风格转换模型的具体步骤,对第一类风格变换帧和第二类风格变换帧进行实时录制播放的具体步骤,以及根据第一类风格变换帧和第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器的具体步骤。下面结合图2a对本申请第二实施例提供的一种实时视频的处理方法进行说明,包括以下步骤:

[0058] 步骤210、确定人脸风格转换模型。

[0059] 本实施例中,通过对预设的机器学习模型进行模型训练,生成可以根据输入视频帧中的真实人脸,生成与真实人脸匹配的风格人脸,并输出包括风格人脸的第一类风格变换帧的人脸风格转换模型。

[0060] 可选的,确定人脸风格转换模型,可以包括:获取训练样本集,训练样本集中的包括多个样本图像对,每个样本图像对中包括原始图像以及变换后图像;使用训练样本集中各样本图像对训练设定机器学习模型,得到所述人脸风格转换模型;其中,所述原始图像中包括真实人脸,所述变换后图像中包括与真实人脸匹配的风格人脸。

[0061] 本可选的实施例中,先获取多个优质的样本图像对组成训练样本集,每个样本图

像对中包括原始图像以及变换后图像,原始图像中包括真实人脸,如图2b中当前处理帧中的人脸,变换后图像中包括与真实人脸匹配的风格人脸,如图2b中第一类风格变换帧中的人脸。然后使用训练样本集中各样本图像对训练设定机器学习模型,使得设定机器学习模型根据输入的原始图像中的真实人脸,学习生成变换后图像中的与真实人脸匹配的风格人脸,训练好的机器学习模型即为人脸风格转换模型。

[0062] 可选的,所述机器学习模型包括:生成对抗网络。

[0063] 本可选的实施例中,生成对抗网络包括生成器和判别器,生成器用于生成与原始图像样本中的真实人脸匹配的风格人脸,判别器用于判断生成器生成的风格人脸是否是变换后图像样本中的风格人脸。通过对生成对抗网络进行训练,如果生成器生成的风格人脸与变换后图像样本中的风格人脸相同,并且,判别器找不出生成器生成的风格人脸与变换后图像样本中的风格人脸之间的区别,即,判定生成器生成的风格人脸与变换后图像样本中的风格人脸相同,则生成器和判别器博弈的思想达到两者收敛,此时,生成对抗网络即为人脸风格转换模型。

[0064] 本实施例中,通过人脸风格转换模型输出的第一类风格变换帧中的风格人脸,与真实人脸的匹配度很高,两者的人脸尺寸大小以及包括的人脸特征都一致,可以形象地将真实人脸转换为风格人脸,例如,动漫风格的二次元形象人脸,吸引用户的注意力,增加用户的兴趣。

[0065] 步骤220、将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧。

[0066] 其中,根据第一实施例中步骤110的解释可知,将实时采集的视频帧加入至视频帧集合中,可以指在视频录制过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中,也可以指在视频直播过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中。

[0067] 步骤230、将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧。

[0068] 示例性的,如图2b所示,将当前处理帧输入至人脸风格转换模型中,通过训练好的人脸风格转换模型,生成与当前处理帧中的真实人脸匹配的二次元形象人脸,并得到包括二次元形象人脸的第一类风格变换帧。

[0069] 步骤240、将当前处理帧作为处理起点帧,在视频帧集合中,获取处理起点帧的后一视频帧。

[0070] 步骤250、根据后一视频帧与处理起点帧中各人脸关键点分别在对应视频帧中的图像位置,生成人脸关键点变换矩阵。

[0071] 示例性的,如图2b所示,根据后一视频帧中各人脸关键点在视频帧中的位置坐标,与处理起点帧中各人脸关键点在视频帧中的位置坐标,计算出后一视频帧中各人脸关键点相对于处理起点帧中各人脸关键点的坐标变化量,例如,鼻子各关键点的坐标变化量,脸部轮廓的各关键点的坐标变化量等,进而生成人脸关键点变换矩阵。

[0072] 步骤260、根据人脸关键点变换矩阵,和与处理起点帧匹配的第一类风格变换帧或第二类风格变换帧,生成后一视频帧的第二类风格变换帧。

[0073] 本实施例中,根据人脸关键点变换矩阵中各人脸关键点的坐标变化量,对与处理

起点帧匹配的第一类风格变换帧中的风格人脸的各人脸关键点的坐标进行相应的调整,生成后一视频帧的第二类风格变换帧。

[0074] 步骤270、判断连续生成的第二类风格变换帧的数量是否等于设定数量,若是,则执行步骤290,否则,执行步骤280。

[0075] 步骤280、将后一视频帧作为处理起点帧,返回执行步骤240。

[0076] 步骤290、从视频帧集合中获取新的当前处理帧,并返回执行步骤230。

[0077] 步骤2110、将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果。

[0078] 本实施例中,在根据人脸关键点变换矩阵连续生成设定数量的第二类风格变换帧之后,从视频帧集合中获取新的当前处理帧,并返回执行步骤230,重复上述过程,生成与视频帧集合中未处理的视频帧匹配的第一类风格变换帧或第二类风格变换帧。

[0079] 本实施例中,步骤290与步骤2110可以分为两个线程并列执行,即,在一个线程中实现,从视频帧集合中获取新的当前处理帧,并重新根据人脸风格转换模型获取与当前处理帧匹配的第一类风格变换帧的同时,如果已经获取的第一类风格变换帧和第二类风格变换帧满足处理条件,可以在另一个线程中同时对第一类风格变换帧和第二类风格变换帧进行相应的处理。

[0080] 本实施例中,将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果,可以包括:对第一类风格变换帧和第二类风格变换帧进行实时录制播放,并生成录制视频,或者,根据第一类风格变换帧和第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器,以进行视频直播。

[0081] 可选的,对第一类风格变换帧和第二类风格变换帧进行实时录制播放,可以包括:将实时生成的第一类风格变换帧和第二类风格变换帧顺序存储于设定缓存队列中;当所述缓存队列中满足预设的硬延时条件时,从所述缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,进行实时录制播放。

[0082] 本可选的实施例中,在视频录制场景下,可以将实时生成的第一类风格变换帧和第二类风格变换帧,按照生成时间先后顺序存储于设定缓存队列中,当缓存队列中的风格变换帧的数量达到预设数量,例如20条,或者,距离上次从缓存队列中获取风格变换帧进行播放,间隔时间等于预设时间,例如3秒,则从缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,进行实时录制播放。

[0083] 可选的,根据第一类风格变换帧和第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器,可以包括:将实时生成的第一类风格变换帧和第二类风格变换帧顺序存储于设定缓存队列中;当所述缓存队列中满足预设的硬延时条件时,从所述缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器。

[0084] 本可选的实施例中,在视频直播场景下,可以将实时生成的第一类风格变换帧和第二类风格变换帧,按照生成时间先后顺序存储于设定缓存队列中,当缓存队列中的风格变换帧的数量达到预设数量,例如20条,或者,距离上次从缓存队列中获取风格变换帧进行播放,间隔时间等于预设时间,例如3秒,则从缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器。

[0085] 本实施例中,可以将风格变换帧的存储,以及从缓存队列中获取风格变换帧进行

播放,分为两个线程并行执行,即,当获取到第一类风格变换帧,或者第二类风格变换帧时,可以在一个线程中,将获取的风格变换帧顺序存储到缓存队列中,如果,发现缓存队列中的风格变换帧的数量达到预设数量,或者,距离上次从缓存队列中获取风格变换帧的间隔时间等于预设时间,则可以同时在另一个线程中实现从缓存队列中顺序获取风格变换帧的操作。

[0086] 本申请实施例的技术方案,通过对预设的机器学习模型进行模型训练,得到可以生成与真实人脸高度匹配的风格人脸的人脸风格转换模型,然后将根据人脸风格转换模型进行人脸风格转换的方式,与根据人脸关键点变换矩阵进行人脸风格转换的方式相结合,通过这两种方式交替使用,实时生成与视频帧中的真实人脸匹配的风格人脸,解决了现有技术中无法实时生成与视频中的真实人脸匹配的风格人脸的问题。

[0087] 第三实施例

[0088] 图3是本申请第三实施例中的一种实时视频的处理装置的结构示意图,该实时视频的处理装置,包括:获取模块310、第一变换模块320、第二变换模块330和循环模块340。

[0089] 获取模块310,用于将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取当前处理帧,视频帧中包括真实人脸;

[0090] 第一变换模块320,用于将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧,风格变换帧中包括风格人脸;

[0091] 第二变换模块330,用于以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧;

[0092] 循环模块340,用于从视频帧集合中获取新的当前处理帧后,返回执行将当前图像帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果。

[0093] 本申请实施例的技术方案,通过将实时采集的视频帧加入至视频帧集合中,并从视频帧集合中获取包括真实人脸的当前处理帧;将当前处理帧输入至人脸风格转换模型中,并获取人脸风格转换模型输出的第一类风格变换帧,风格变换帧中包括风格人脸;以当前处理帧和第一类风格变换帧为起点,根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系,生成所述设定数量的第二类风格变换帧;从视频帧集合中获取新的当前处理帧后,返回执行将当前处理帧输入至人脸风格转换模型中的操作,并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果,实现了基于视频中的真实人脸,实时生成与真实人脸匹配的风格人脸,解决了现有技术中无法实时生成与视频中的真实人脸匹配的风格人脸的问题。

[0094] 可选的,第二变换模块330,包括:

[0095] 后一视频帧获取单元,用于将当前处理帧作为处理起点帧,在视频帧集合中,获取处理起点帧的后一视频帧;

[0096] 变换矩阵生成单元,用于根据后一视频帧与处理起点帧中各人脸关键点分别在对应视频帧中的图像位置,生成人脸关键点变换矩阵;

[0097] 第二类风格变换帧生成单元,用于根据人脸关键点变换矩阵,和与处理起点帧匹配的第一类风格变换帧或第二类风格变换帧,生成后一视频帧的第二类风格变换帧;

- [0098] 循环操作单元,用于将所述后一视频帧作为处理起点帧,返回执行在视频帧集合中,获取处理起点帧的后一视频帧的操作,直至后一视频帧的处理数量达到所述设定数量。
- [0099] 可选的,还包括:模型训练模块,用于在将当前处理帧输入至人脸风格转换模型中之前,获取训练样本集,训练样本集中的包括多个样本图像对,每个样本图像对中包括原始图像以及变换后图像;
- [0100] 使用训练样本集中各样本图像对训练设定机器学习模型,得到所述人脸风格转换模型;
- [0101] 其中,所述原始图像中包括真实人脸,所述变换后图像中包括与真实人脸匹配的风格人脸。
- [0102] 可选的,所述机器学习模型包括:生成对抗网络。
- [0103] 可选的,获取模块310,包括:
- [0104] 录制获取单元,用于在视频录制过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中;
- [0105] 其中,循环模块340,包括:
- [0106] 录制播放单元,用于对第一类风格变换帧和第二类风格变换帧进行实时录制播放,并生成录制视频。
- [0107] 可选的,录制播放单元,具体用于:
- [0108] 将实时生成的第一类风格变换帧和第二类风格变换帧顺序存储于设定缓存队列中;
- [0109] 当所述缓存队列中满足预设的硬延时条件时,从所述缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,进行实时录制播放。
- [0110] 可选的,获取模块310,包括:
- [0111] 直播获取单元,用于在视频直播过程中,响应于人脸风格变换请求,将实时采集的视频帧加入至视频帧集合中;
- [0112] 其中,循环模块340,包括:
- [0113] 直播播放单元,用于根据第一类风格变换帧和第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器,以进行视频直播。
- [0114] 可选的,直播播放单元,具体用于:
- [0115] 将实时生成的第一类风格变换帧和第二类风格变换帧顺序存储于设定缓存队列中;
- [0116] 当所述缓存队列中满足预设的硬延时条件时,从所述缓存队列中顺序获取第一类风格变换帧,或者第二类风格变换帧,生成直播视频流,并将直播视频流发送至直播服务器。
- [0117] 本申请实施例所提供的实时视频的处理装置可执行本申请任意实施例所提供的实时视频的处理方法,具备执行方法相应的功能模块和有益效果。
- [0118] 第四实施例
- [0119] 根据本申请的实施例,本申请还提供了一种电子设备和一种可读存储介质。
- [0120] 如图4所示,是根据本申请实施例的实时视频的处理方法的电子设备的框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字

助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本申请的实现。

[0121] 如图4所示,该电子设备包括:一个或多个处理器401、存储器402,以及用于连接各部件的接口,包括高速接口和低速接口。各个部件利用不同的总线互相连接,并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在电子设备内执行的指令进行处理,包括存储在存储器中或者存储器上以在外部输入/输出装置(诸如,耦合至接口的显示设备)上显示GUI的图形信息的指令。在其它实施方式中,若需要,可以将多个处理器和/或多条总线与多个存储器和多个存储器一起使用。同样,可以连接多个电子设备,各个设备提供部分必要的操作(例如,作为服务器阵列、一组刀片式服务器、或者多处理器系统)。图4中以一个处理器401为例。

[0122] 存储器402即为本申请所提供的非瞬时计算机可读存储介质。其中,所述存储器存储有可由至少一个处理器执行的指令,以使所述至少一个处理器执行本申请所提供的实时视频的处理方法。本申请的非瞬时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本申请所提供的实时视频的处理方法。

[0123] 存储器402作为一种非瞬时计算机可读存储介质,可用于存储非瞬时软件程序、非瞬时计算机可执行程序以及模块,如本申请实施例中的实时视频的处理方法对应的程序指令/模块(例如,附图3所示的获取模块310、第一变换模块320、第二变换模块330和循环模块340)。处理器401通过运行存储在存储器402中的非瞬时软件程序、指令以及模块,从而执行服务器的各种功能应用以及数据处理,即实现上述方法实施例中的实时视频的处理方法。

[0124] 存储器402可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储根据实时视频的处理电子设备的使用所创建的数据等。此外,存储器402可以包括高速随机存取存储器,还可以包括非瞬时存储器,例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些实施例中,存储器402可选包括相对于处理器401远程设置的存储器,这些远程存储器可以通过网络连接至实时视频的处理电子设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0125] 实时视频的处理方法的电子设备还可以包括:输入装置403和输出装置404。处理器401、存储器402、输入装置403和输出装置404可以通过总线或者其他方式连接,图4中以通过总线连接为例。

[0126] 输入装置403可接收输入的数字或字符信息,以及产生与实时视频的处理电子设备的用户设置以及功能控制有关的键信号输入,例如触摸屏、小键盘、鼠标、轨迹板、触摸板、指示杆、一个或者多个鼠标按钮、轨迹球、操纵杆等输入装置。输出装置404可以包括显示设备、辅助照明装置(例如,LED)和触觉反馈装置(例如,振动电机)等。该显示设备可以包括但不限于,液晶显示器(LCD)、发光二极管(LED)显示器和等离子体显示器。在一些实施方式中,显示设备可以是触摸屏。

[0127] 此处描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、专用ASIC(专用集成电路)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种

实施方式可以包括：实施在一个或者多个计算机程序中，该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释，该可编程处理器可以是专用或者通用可编程处理器，可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令，并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0128] 这些计算程序(也称作程序、软件、软件应用、或者代码)包括可编程处理器的机器指令，并且可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。如本文使用的，术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供给可编程处理器的任何计算机程序产品、设备、和/或装置(例如，磁盘、光盘、存储器、可编程逻辑装置(PLD))，包括，接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”指的是用于将机器指令和/或数据提供给可编程处理器的任何信号。

[0129] 为了提供与用户的交互，可以在计算机上实施此处描述的系统和技术，该计算机具有：用于向用户显示信息的显示装置(例如，CRT(阴极射线管)或者LCD(液晶显示器)监视器)；以及键盘和指向装置(例如，鼠标或者轨迹球)，用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互；例如，提供给用户的反馈可以是任何形式的传感反馈(例如，视觉反馈、听觉反馈、或者触觉反馈)；并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0130] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如，作为数据服务器)、或者包括中间件部件的计算系统(例如，应用服务器)、或者包括前端部件的计算系统(例如，具有图形用户界面或者网络浏览器的用户计算机，用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如，通信网络)来将系统的部件相互连接。通信网络的示例包括：局域网(LAN)、广域网(WAN)和互联网。

[0131] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。

[0132] 根据本申请实施例的技术方案，通过将实时采集的视频帧加入至视频帧集合中，并从视频帧集合中获取包括真实人脸的当前处理帧；将当前处理帧输入至人脸风格转换模型中，并获取人脸风格转换模型输出的第一类风格变换帧，风格变换帧中包括风格人脸；以当前处理帧和第一类风格变换帧为起点，根据视频帧集合中设定数量的后续视频帧与前一视频帧中各人脸关键点间的位置关系，生成所述设定数量的第二类风格变换帧；从视频帧集合中获取新的当前处理帧后，返回执行将当前处理帧输入至人脸风格转换模型中的操作，并将第一类风格变换帧和第二类风格变换帧作为实时视频处理结果，实现了基于视频中的真实人脸，实时生成与真实人脸匹配的风格人脸，解决了现有技术中无法实时生成与视频中的真实人脸匹配的风格人脸的问题。

[0133] 应该理解，可以使用上面所示的各种形式的流程，重新排序、增加或删除步骤。例如，本发申请中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行，

只要能够实现本申请公开的技术方案所期望的结果,本文在此不进行限制。

[0134] 上述具体实施方式,并不构成对本申请保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本申请的精神和原则之内所作的修改、等同替换和改进等,均应包含在本申请保护范围之内。

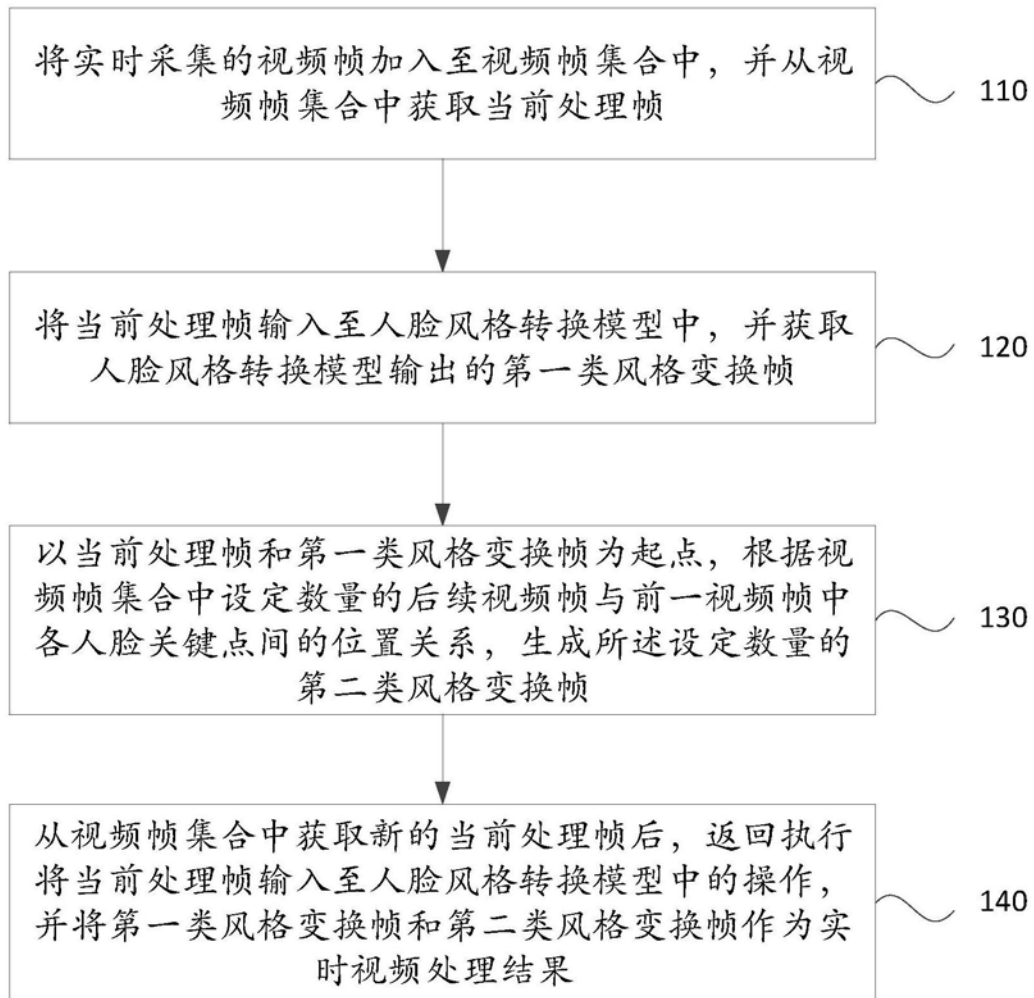


图1

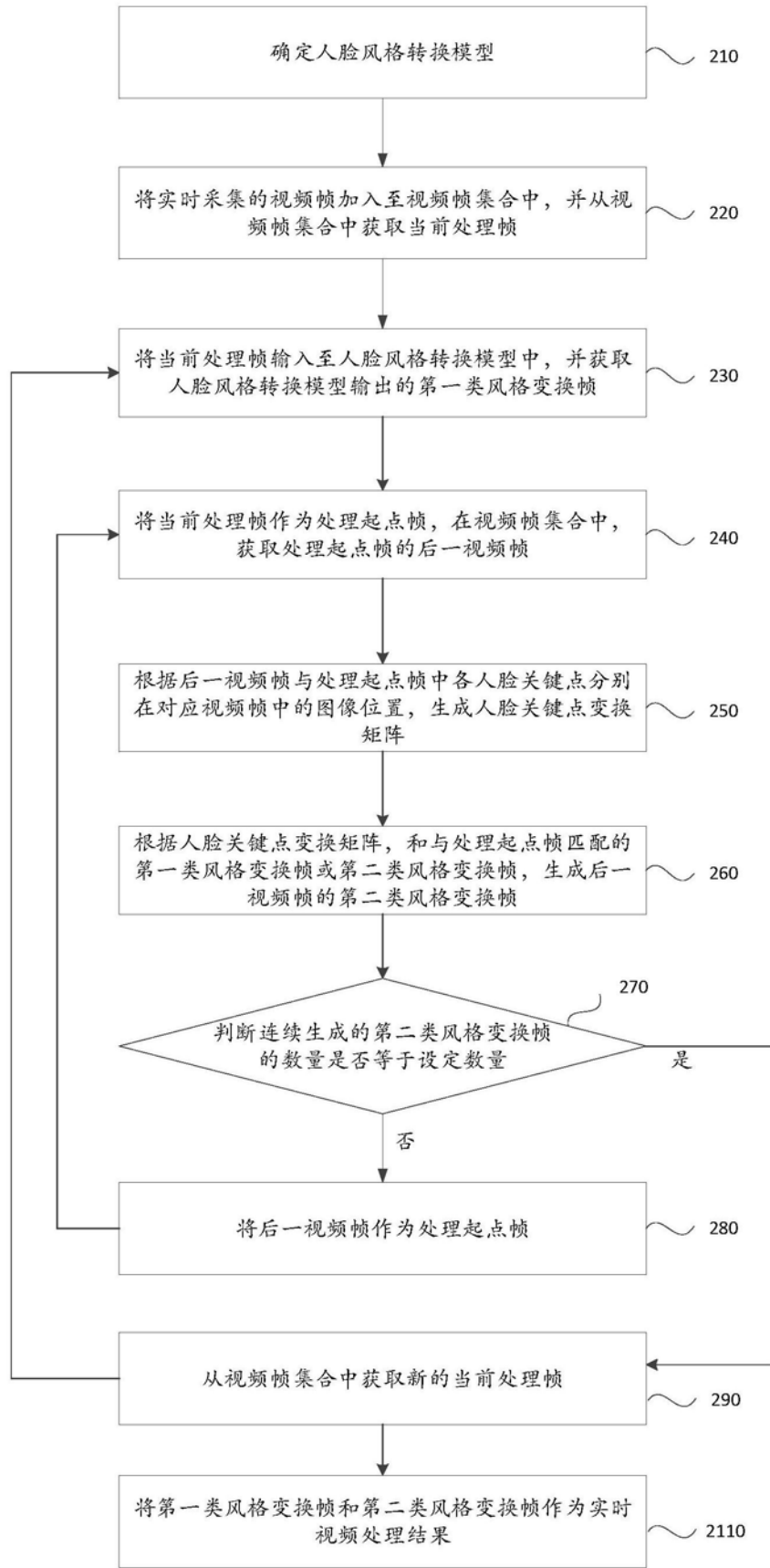


图2a

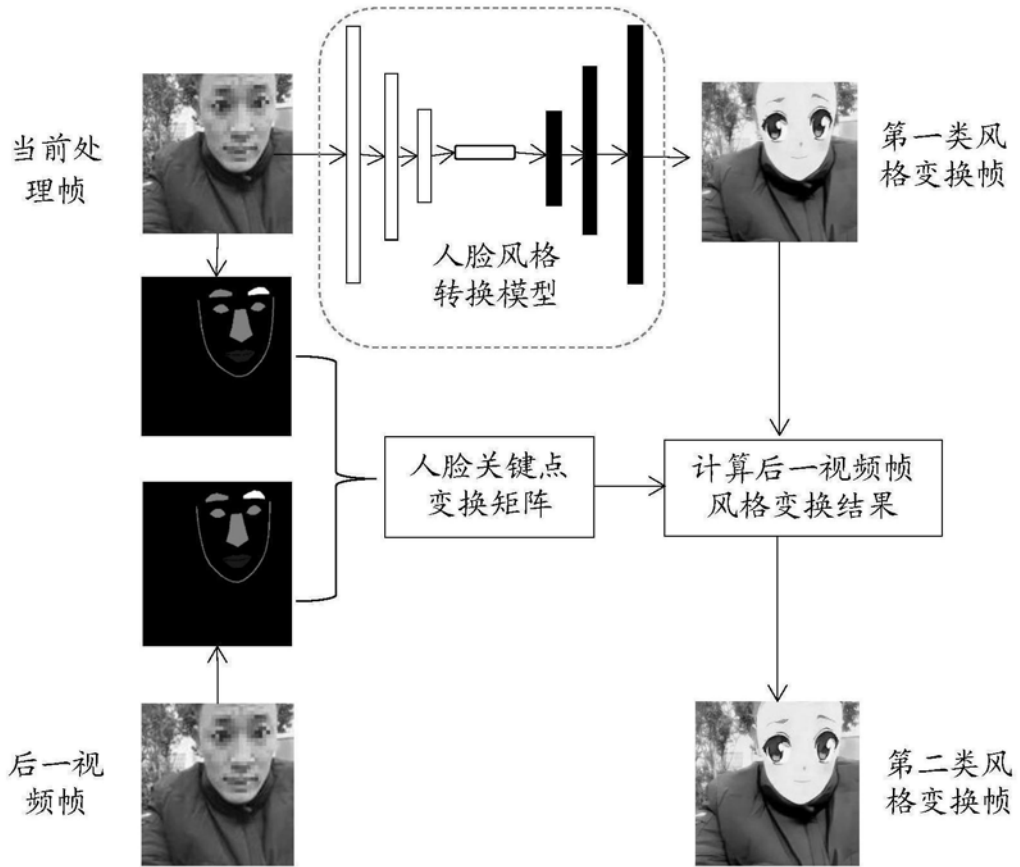


图2b

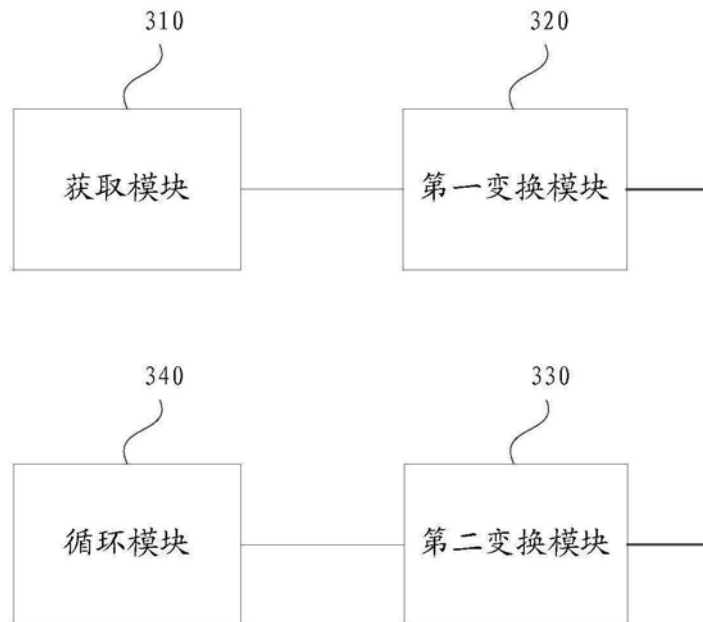


图3

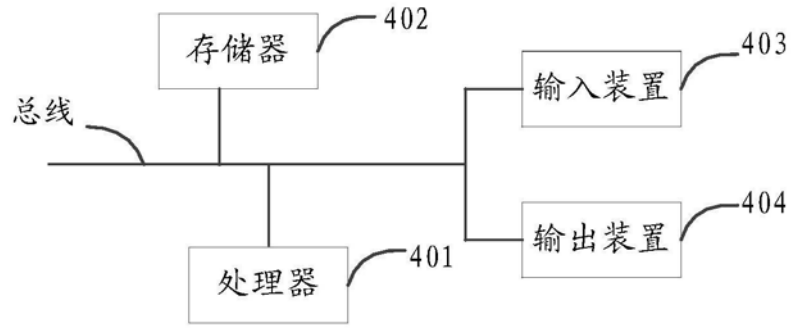


图4