

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2025 年 1 月 9 日 (09.01.2025)



(10) 国际公布号
WO 2025/007933 A1

(51) 国际专利分类号:
G06V 10/774 (2022.01)

(21) 国际申请号: PCT/CN2024/103675

(22) 国际申请日: 2024 年 7 月 4 日 (04.07.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310814024.X 2023年7月4日 (04.07.2023) CN

(71) 申请人: 支付宝(杭州)信息技术有限公司 (ALIPAY (HANGZHOU) INFORMATION TECHNOLOGY CO., LTD.) [CN/CN]; 中国浙江省杭州市西湖区西溪路 556 号 8 层 B 段 801-11, Zhejiang 310000 (CN)。

(72) 发明人: 童仁玲 (TONG, Renling); 中国浙江省杭州市西湖区西溪路 556 号 8 层 B 段 801-11, Zhejiang 310000 (CN)。

(74) 代理人: 北京博思佳知识产权代理有限公司 (BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街 9 号嘉华大厦 B 座 409, Beijing 100085 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN,

(54) Title: FACIAL ANIMATION GENERATION MODEL TRAINING AND FACIAL ANIMATION

(54) 发明名称: 人脸活化生成模型训练和人脸活化

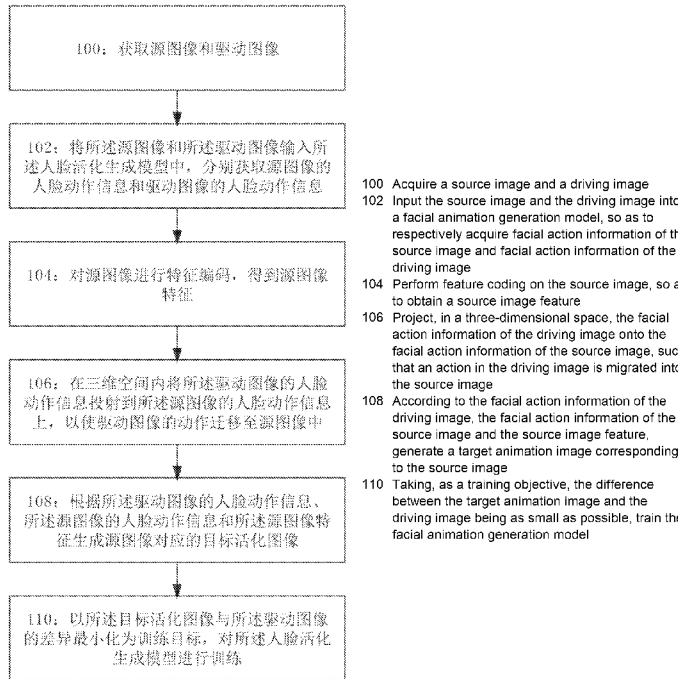


图 1

(57) Abstract: Disclosed in the embodiments of the present specification is a facial animation generation model training method. The method comprises: acquiring a source image and a driving image, and inputting the source image and the driving image into a facial animation generation model, so as to obtain facial action information of the source image and facial action information of the driving image; performing feature coding on the source image, so as to obtain a source image feature; projecting, in a three-dimensional space, the facial action information of the driving image onto the facial action information of the source image, such that an action in the

[见续页]

MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

driving image is migrated into the source image; generating a target animation image according to the facial action information of the driving image, the facial action information of the source image and the source image feature; and taking, as a training objective, the difference between the target animation image and the driving image being as small as possible, training the facial animation generation model. Further disclosed in the embodiments of the present specification is a facial animation method. Correspondingly, disclosed in the embodiments of the present specification are a facial animation generation model training apparatus and a facial animation apparatus.

(57) 摘要: 本说明书实施例公开了一种人脸活化生成模型训练方法, 所述方法包括: 获取源图像和驱动图像, 并输入所述人脸活化生成模型中, 得到源图像的人脸动作信息和驱动图像的人脸动作信息; 对源图像进行特征编码得到源图像特征; 在三维空间内将驱动图像的人脸动作信息投射到源图像的人脸动作信息上, 以使驱动图像的动作迁移至源图像中; 根据驱动图像的人脸动作信息、源图像的人脸动作信息和源图像特征生成目标活化图像; 以目标活化图像与驱动图像的差异最小化为训练目标, 对所述人脸活化生成模型进行训练。本说明书实施例还公开了人脸活化方法。相应地, 本说明书实施例公开了人脸活化生成模型训练装置和人脸活化装置。

人脸活化生成模型训练和人脸活化

技术领域

[01]本说明书实施例涉及一种图像处理方法，尤其涉及一种人脸活化生成模型训练方法。

背景技术

[02]手机已经成为如今人人必备的通讯工具，随着科技的发展，使用手机进行身份验证的方式经历了从密码到指纹再到人脸的发展，而手机等设备的安全性也正受到各方面的挑战。

[03]在 IIFAA 生物安全性攻击检测中，需要智能生成足够逼真的人脸攻击物料，来对手机的身份验证功能进行安全评测，以保证人脸安全评测的完备有效。该攻击手段首先要有通用性，在不同的手机黑盒测试中都具备较稳定的攻击力；其次要有高攻击率，具备强攻击性，从而能够区分人脸识别算法的抗攻击能力强弱，有效拉开不同待测手机厂商在身份验证功能的抗攻击性上的差距；另外该攻击手段需要通过算法智能生成，区别于人造面具等物理生成或普通的 ps 手段，使从人脸识别算法使用的模型中发现的潜在错误有迹可循，便于对模型进行改进，进而提高模型的可靠性与模型预测结果的可解释性。

[04]相比于 FOMM 等目前主流的人脸活化算法，仍需要更有效的通过动作迁移获得攻击物料的方法，以减少在转头等动作上发生的动作畸变现象，保留完好的人脸身份信息，从而使在评价人脸识别算法的抗攻击性时获得的结果更加可靠。

[05]鉴于此，希望获得一种新的人脸活化生成方法，能够生成更逼真的人脸活化图像，从而更有效地评测人脸识别算法。

发明内容

[06]本说明书实施例的目的之一在于提供一种人脸活化生成模型训练方法，该方法能够将指定的动作迁移到指定的人脸图像中，较大程度地保留指定人脸的身份信息，有效测试人脸识别算法的抗攻击性。

[07]根据上述目的，本说明书实施例提出了一种人脸活化生成模型训练方法，所述方法包括：获取源图像和驱动图像；将所述源图像和所述驱动图像输入所述人脸活化生成模型中，分别获取源图像的人脸动作信息和驱动图像的人脸动作信息；对源图像进行特征

编码，得到源图像特征；在三维空间内将所述驱动图像的人脸动作信息投射到所述源图像的人脸动作信息上，以使驱动图像的动作迁移至源图像中；根据所述驱动图像的人脸动作信息、所述源图像的人脸动作信息和所述源图像特征生成源图像对应的目标活化图像；以所述目标活化图像与所述驱动图像的差异最小化为训练目标，对所述人脸活化生成模型进行训练。

[08]在本说明书实施例中，利用三维空间内的人脸动作信息投射实现了从驱动图像到源图像的动作迁移，将三维的位置信息与人脸纹理信息相关联，很好地改善了动作畸变的问题，使源图像中的人脸身份信息可以较准确且完整地保留；通过使目标活化图像与驱动图像的差异最小化，较大程度地提高了目标活化图像的逼真度。综上，通过该方法训练出的人脸活化生成模型所生成的目标活化图像能够更有效地测试人脸识别算法的抗攻击性。

[09]进一步地，在一些实施方式中，所述人脸动作信息包括人脸三维姿态信息以及人脸关键点信息。

[10]更进一步地，在一些实施方式中，在三维空间内将所述驱动图像的人脸动作信息投射到所述源图像的人脸动作信息上，以使驱动图像的动作迁移至源图像中具体包括：通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点；将所述驱动关键点投射到所述源关键点上，以将驱动图像的动作迁移至源图像中。

[11]更进一步地，在一些实施方式中，所述人脸三维姿态信息包括头部转动角度。

[12]进一步地，在一些实施方式中，所述人脸活化生成模型通过生成对抗网络实现，所述生成对抗网络包括判别器和生成器；以所述目标活化图像与所述驱动图像的差异最小化为训练目标，对所述人脸活化生成模型进行训练具体包括：将所述驱动图像的人脸动作信息、所述源图像的人脸动作信息和所述源图像特征输入所述生成器，输出源图像对应的目标活化图像；利用判别器判断所述目标活化图像是否为所述驱动图像；根据所述目标活化图像与所述驱动图像计算生成损失；根据所述判别器的判别结果计算判别损失；分别以所述判别损失和所述生成损失最小化为训练目标，对所述生成对抗网络进行训练。

[13]更进一步地，在一些实施方式中，根据所述目标活化图像与所述驱动图像计算生成损失具体包括：计算所述目标活化图像与所述驱动图像之间基于像素的第一损失；计算

所述目标活化图像与所述驱动图像之间基于特征图的第二损失；将所述第一损失和所述第二损失加权融合，作为所述生成损失。

[14]更进一步地，在一些实施方式中，通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点具体包括：通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点。

[15]更进一步地，在一些实施方式中，通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点具体包括：通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，利用矩阵变化将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点。

[16]本说明书实施例的另一目的在于提供一种人脸活化方法，该方法可以将指定的动作迁移到指定的人脸图像中，生成逼真的人脸活化图像，且较大程度地保留指定人脸的身份信息，从而有效测试人脸识别算法的抗攻击性。

[17]根据上述目的，本说明书实施例提供了一种人脸活化方法，所述方法包括：获取待活化的源图像和包含目标动作的驱动视频，从所述驱动视频中抽取若干帧驱动图像；针对获得的每帧驱动图像，将所述源图像和所述驱动图像输入所述人脸活化生成模型中，生成源图像对应的目标活化图像，其中，所述人脸活化生成模型是采用如上任一方法所述的步骤训练得到的；将获得的所有目标活化图像进行连接，得到目标活化视频。

[18]本说明书实施例的又一目的在于提供一种人脸活化生成模型训练装置，该装置能够将指定的动作迁移到指定的人脸图像中，较大程度地保留指定人脸的身份信息，可以有效测试人脸识别算法的抗攻击性。

[19]根据上述目的，本说明书实施例提供了一种人脸活化生成模型训练装置，所述装置包括：样本获取模块，用于获取源图像和驱动图像；信息提取模块，用于将所述源图像和所述驱动图像输入所述人脸活化生成模型中，分别获取源图像的人脸动作信息和驱动图像的人脸动作信息；对源图像进行特征编码，得到源图像特征；生成模块，用于在三

维空间内将所述驱动图像的人脸动作信息投射到所述源图像的人脸动作信息上,以使驱动图像的动作迁移至源图像中;根据所述驱动图像的人脸动作信息、所述源图像的人脸动作信息和所述源图像特征生成源图像对应的目标活化图像;训练模块,用于以所述目标活化图像与所述驱动图像的差异最小化为训练目标,对所述人脸活化生成模型进行训练。

[20]进一步地,在一些实施方式中,所述人脸动作信息包括人脸三维姿态信息以及人脸关键点信息。

[21]更进一步地,在一些实施方式中,所述生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息,将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下,得到三维空间下的源关键点和驱动关键点;将所述驱动关键点投射到所述源关键点上,以将驱动图像的动作迁移至源图像中。

[22]更进一步地,在一些实施方式中,所述人脸三维姿态信息包括头部转动角度。

[23]进一步地,在一些实施方式中,所述人脸活化生成模型通过生成对抗网络实现,所述生成对抗网络包括判别器和生成器;所述训练模块具体包括:将所述驱动图像的人脸动作信息、所述源图像的人脸动作信息和所述源图像特征输入所述生成器,输出源图像对应的目标活化图像;利用判别器判断所述目标活化图像是否为所述驱动图像;根据所述目标活化图像与所述驱动图像计算生成损失;根据所述判别器的判别结果计算判别损失;分别以所述判别损失和所述生成损失最小化为训练目标,对所述生成对抗网络进行训练。

[24]更进一步地,在一些实施方式中,所述训练模块计算所述目标活化图像与所述驱动图像之间基于像素的第一损失;计算所述目标活化图像与所述驱动图像之间基于特征图的第二损失;将所述第一损失和所述第二损失加权融合,作为所述生成损失。

[25]更进一步地,在一些实施方式中,所述生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息,将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下,得到三维空间下的源关键点和驱动关键点。

[26]更进一步地,在一些实施方式中,所述生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息,利用矩阵变化将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下,得到三维空间下的源关

键点和驱动关键点。

[27]本说明书实施例的又一目的在于提供一种人脸活化装置，该装置可以将指定的动作迁移到指定的人脸图像中，生成逼真的人脸活化图像，且较大程度地保留指定人脸的身份信息，从而有效测试人脸识别算法的抗攻击性。

[28]根据上述目的，本说明书实施例提供了一种人脸活化装置，所述装置包括：样本采集模块，用于获取待活化的源图像和包含目标动作的驱动视频，从所述驱动视频中抽取若干帧驱动图像；活化生成模块，用于针对获得的每帧驱动图像，将所述源图像和所述驱动图像输入所述人脸活化生成模型中，生成源图像对应的目标活化图像，其中，所述人脸活化生成模型是采用如上任一方法所述的步骤训练得到的；连接模块，用于将获得的所有目标活化图像进行连接，得到目标活化视频。

[29]本说明书实施例的又一目的在于提供一种电子设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述程序时实现如上任一方法所述的步骤。

[30]本说明书实施例所述的人脸活化生成模型训练和人脸活化方法具有以下有益效果。

[31]利用三维空间内从驱动图像到源图像的人脸动作信息投射实现动作迁移，将三维的位置信息与人脸纹理信息相关联，能够很好地改善头部转动导致的动作畸变问题，使源图像中的人脸身份信息可以较准确且完整地保留；通过使目标活化图像与驱动图像的差异最小化，较大程度地提高了目标活化图像的逼真度，所生成的目标活化图像能够更有效地测试人脸识别算法的抗攻击性。

[32]基于人脸三维姿态信息进行关键点对齐矫正能够将原来 2D 的关键点信息投射到 3D 的人脸三维姿态信息中，人脸活化转变为伪 3D 的活化，使得人脸关键点信息与人脸纹理特征更好地融合，生成更逼真的活化图像。

[33]本说明书实施例所述的人脸活化生成模型训练和人脸活化装置同样具有上述有益效果。

附图说明

[34]图 1 示例性地显示了本说明书实施例所述的人脸活化生成模型训练方法在一种实施方式下的流程示意图。

[35]图 2 示例性地示出了本说明书实施例所述的人脸活化生成模型训练方法在一种具体

实施方式下执行的步骤。

[36]图 3 示例性地显示了本说明书实施例所述的人脸活化方法在一种实施方式下的流程示意图。

[37]图 4 示例性地显示了本说明书实施例所述的人脸活化生成模型训练装置在一种实施方式下的结构示意图。

[38]图 5 示例性地显示了本说明书实施例所述的人脸活化装置在一种实施方式下的结构示意图。

具体实施方式

[39]首先需要说明的是，在本发明实施例中使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本发明。在本发明实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。另外，在不冲突的情况下，本说明书中的实施例及实施例中的特征可以相互组合。

[40]需要说明的是，人脸活化的过程表现为将图像 B 中人脸的动作或者表情迁移至图像 A 中的人脸上，由于表情变化和表情迁移可以表现为脸部五官的动作变化和五官的动作迁移，因此本说明书实施例中的动作不仅包括摇头、点头等运动，也包括微笑、皱眉、闭眼等表情。

[41]下面将结合说明书附图和具体的实施例来对本发明所述的人脸活化生成模型训练和人脸活化方法以及装置进行进一步地详细说明，但是该详细说明不构成对本发明的限制。

[42]在本说明书的一个实施例中，提出了一种人脸活化生成模型训练方法。图 1 示例性地显示了本说明书实施例所述的人脸活化生成模型训练方法在一种实施方式下的流程示意图。

[43]如图 1 所示，包括如下步骤。

[44]100：获取源图像和驱动图像。

[45]使用到的源图像和驱动图像需要包含明显不同的人脸动作，以更直观地观察动作迁移的效果。源图像和驱动图像可以从现有的包含不同人脸动作的图像集中提取，也可以来自通过摄像头等采集设备拍摄的不同人的图像或者若干段视频，其中每段视频都记录了同一个人的人脸动作变化过程。

[46]在一些实施例中，每次获取一段视频作为训练样本，从中抽取人脸动作不同的两帧图像分别作为源图像 A 和驱动图像 B，人脸活化的目的是将驱动图像 B 中的动作迁移到源图像 A 中的人脸上。针对每一轮训练，利用同一个人的人脸作为样本数据对人脸活化生成模型进行训练，这样模型可以得到充分的训练，从而更容易收敛。

[47]102: 将源图像和驱动图像输入人脸活化生成模型中，分别获取源图像的人脸动作信息和驱动图像的人脸动作信息。

[48]在人脸活化生成模型中，可以采取通用的人脸动作检测方法分别获取源图像的人脸动作信息和驱动图像的人脸动作信息。人脸动作信息能够在三维空间内表现头部的运动情况以及五官的位置变化情况，相对于二维空间的动作信息来说具有更强的逼真度和表征性。

[49]在一些实施例中，人脸动作信息包括人脸三维姿态信息以及人脸关键点信息。

[50]相应地，人脸动作检测方法包括人脸三维姿态检测方法和人脸关键点检测方法。人脸三维姿态信息可以体现三维空间内头部的姿态，以确定人脸的朝向；而人脸关键点信息通过在脸部眼睛、鼻子、嘴巴等关键五官或者其它脸部关键区域添加关键点来记录位置信息，从而根据这些关键区域的位置变化表现人脸动作。

[51]在一些更具体的实施例中，人脸三维姿态信息包括头部转动角度。

[52]人脸三维姿态信息可以用人脸 3D 位姿信息来表示，用于描述头部在三维空间中互相垂直的三个坐标轴上转动的角度，换句话说，人脸位姿可以体现头部的朝向，表达头部位置信息。这三个互相垂直的坐标轴基于点头、摇头以及摆头动作所沿的方向构建以头部为中心的 XYZ 轴，能够体现头部分别在三个坐标轴上的转动角度，并汇总为人脸 3D 位姿信息。

[53]104: 对源图像进行特征编码，得到源图像特征。

[54]将源图像输入通用的图像编码器进行特征编码，输出对应的源图像特征，以便基于源图像特征对源图像作动作迁移等进一步处理，配合图像解码器生成目标活化图像。

[55]106: 在三维空间内将驱动图像的人脸动作信息投射到源图像的人脸动作信息上，以使驱动图像的动作迁移至源图像中。

[56]在一些实施例中，在三维空间内将驱动图像的人脸动作信息投射到源图像的人脸动作信息上，以使驱动图像的动作迁移至源图像中具体包括：通过比对源图像的人脸三维

姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点；将驱动关键点投射到源关键点上，以将驱动图像的动作迁移至源图像中。

[57]由于人脸三维姿态信息表现了三维空间内头部的姿态，可以基于此得到源图像的头部姿态与驱动图像的头部姿态。图像中的人脸关键点信息与头部姿态是对应的，为了使驱动图像中的动作与源图像中的人脸更好地融合，基于二者的头部姿态差异进行关键点矫正，使源图像中的关键点与驱动图像中的关键点背后所展现的人脸位于同一姿态下，得到源关键点和驱动关键点，以便进一步比对出位置相同与位置不同的关键点，位置不同的关键点即代表了人脸动作的变化，是需要进行重点迁移的对象。通过将驱动关键点投射到源关键点上，根据位置不同的关键点对源关键点进行调整，从而在源图像中获得需要迁移的人脸动作信息，完成驱动图像的动作迁移。将源图像的头部姿态与驱动图像的头部姿态进行统一，有利于将人脸调整至更方便进行动作迁移的姿态，获得更好的人脸活化效果，同时避免由于头部姿态的不同对活化结果产生影响。

[58]需要说明的是，在将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下时，可以固定源图像的头部姿态对应的关键点，以将源图像的头部姿态作为目标姿态，通过矫正驱动图像的关键点使得驱动图像的头部姿态与目标姿态一致；也可以固定驱动图像的头部姿态对应的关键点，以将驱动图像的头部姿态作为目标姿态，通过矫正源图像的关键点使得源图像的头部姿态与目标姿态一致；也可以预先设置可以更好地展现人脸活化效果的目标姿态，再将源图像的头部姿态与驱动图像的头部姿态矫正至该目标姿态下。

[59]在一些更具体的实施例，通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点具体包括：通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点。

[60]在对人脸关键点信息进行矫正时，可以将人脸关键点旋转至目标姿态对应的关键点位置，也可以对人脸关键点进行偏置，当然，还可以先将人脸关键点旋转至目标姿态对应的大致关键点位置，再经过偏置对人脸关键点位置进行进一步调整。

[61]在一些更具体的实施例，通过比对源图像的人脸三维姿态信息和驱动图像的人脸

三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点具体包括：通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，利用矩阵变化将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点。

[62]基于源图像的人脸关键点信息和驱动图像的人脸关键点信息，选择进行关键点矫正的目标姿态，并通过通用的关键点检测方法获取该目标姿态对应的目标关键点，进而分别确定将源图像的人脸关键点信息和驱动图像的人脸关键点信息调整至目标关键点的仿射变换矩阵，利用对应的仿射变换矩阵分别对源图像的人脸关键点信息和驱动图像的人脸关键点信息进行矫正以调整至目标姿态。

[63]108: 根据驱动图像的人脸动作信息、源图像的人脸动作信息和源图像特征生成源图像对应的目标活化图像。

[64]采用通用的图像解码器结合源图像特征与动作迁移获得的人脸动作信息生成目标活化图像。在一些实施例中，图像解码器可以基于 GAN 生成网络中的图像解码器构建。

[65]110: 以目标活化图像与驱动图像的差异最小化为训练目标，对人脸活化生成模型进行训练。

[66]图像之间的差异可以通过计算图像中逐像素值的差异获得，在一些实施例中还可以将需要比对的图像进一步处理为其他层面的图像，再计算逐像素值差异，例如提取待比对图像的特征图。优选地，通过多种途径计算目标活化图像与驱动图像之间的差异，再利用获得的多种损失训练模型有利于提高人脸活化生成模型的训练效率，加快模型性能的提升。

[67]在一些实施例中，人脸活化生成模型通过生成对抗网络实现，生成对抗网络包括判别器和生成器；以目标活化图像与驱动图像的差异最小化为训练目标，对人脸活化生成模型进行训练具体包括：将驱动图像的人脸动作信息、源图像的人脸动作信息和源图像特征输入生成器，输出源图像对应的目标活化图像；利用判别器判断目标活化图像是否为驱动图像；根据目标活化图像与驱动图像计算生成损失；根据判别器的判别结果计算判别损失；分别以判别损失和生成损失最小化为训练目标，对生成对抗网络进行训练。

[68]生成对抗网络由判别器和生成器组成，在训练生成对抗网络的过程中，首先训练判别器，使其具有良好的判别能力；再训练生成器，使得生成器生成的结果能过骗过判别

能力良好的判别器，拥有更强的逼真度。在上述实施例中，获得目标活化图像后将其输入判别器，以判断是真实的样本图像还是经过篡改的伪造图像，获得判别结果；再根据判别结果的准确性计算判别损失，通过使判别损失最小化训练判别器的判别能力。

[69]优选地，针对人脸活化生成模型的每一轮训练，采用的训练样本中源图像与驱动图像为同一个人。通过人脸活化将驱动图像中的人脸动作迁移至源图像中，与源图像中的人脸融合，所获得的目标活化图像在理想情况下应该与驱动图像一致，因此通过计算驱动图像与目标活化图像之间的差异确定生成损失，通过令生成损失最小化优化生成器的生成效果，使驱动图像与源图像融合得更加自然。

[70]随着生成器的不断优化，生成的目标活化图像越来越逼真，使得判别器越来越难以判断其真伪，当判别器已无法判断目标活化图像的真伪时，对人脸活化生成模型的生成对抗式训练完成。

[71]在不断地通过计算损失进行迭代训练的过程中，人脸活化生成模型还学习并迁移了源图像与驱动图像中的人脸纹理特征，即该人脸活化过程将表示人脸三维位置的人脸动作信息与人脸纹理特征进行关联与融合，实现三维空间下的人脸活化，改善了动作迁移后人脸动作发生畸变的问题。

[72]在一些更具体的实施例中，采用 GAN 网络模型构建生成对抗网络来实现人脸活化生成模型。

[73]在一些更具体的实施例中，根据目标活化图像与驱动图像计算生成损失具体包括：计算目标活化图像与驱动图像之间基于像素的第一损失；计算目标活化图像与驱动图像之间基于特征图的第二损失；将第一损失和第二损失加权融合，作为生成损失。

[74]优选地，第一损失可以通过计算目标活化图像与驱动图像之间逐像素的均方差损失获得。

[75]优选地，第二损失通过计算目标活化图像与驱动图像之间的人脸相似度获得，可以分别提取出目标活化图像的特征向量与驱动图像的特征向量后，采用 Arcface（又称加性角度间隔损失函数）与交叉熵结合的方式计算人脸相似度。通过 Arcface 对人脸活化生成模型进行训练可以在保证类间距离的情况下有效地增大类间距离，同时达到减小类内距离的效果，使得人脸活化生成模型能够更高效地生成更准确的活化结果。

[76]图 2 示例性地示出了本说明书实施例所述的人脸活化生成模型训练方法在一种具体实施方式下执行的步骤。

[77]如图 2 所示的, 在一些实施例中, 从同一人的一段视频中抽取两帧图像分别作为源图像和驱动图像, 分别提取源图像的人脸动作信息和驱动图像的人脸动作信息。其中, 通过通用的人脸关键点检测方法分别提取出源图像的人脸关键点以及驱动图像的人脸关键点, 由于源图像与驱动图像中的人脸相同, 可以仅对源图像中的人脸关键点进行提取, 并复制到驱动图像中; 再通过通用的人脸位姿检测方法分别提取出源图像的人脸 3D 位姿以及驱动图像的人脸 3D 位姿, 基于人脸 3D 位姿对比出源图像与驱动图像中人脸的头部姿态差异, 利用矩阵变化进行人脸关键点矫正, 使得源图像中的头部姿态与驱动图像中的头部姿态一致, 得到矫正后的源关键点与驱动关键点。

[78]接着, 对源图像中的图像特征进行提取, 将源图像特征与源关键点、驱动关键点一起输入基于 GAN 生成网络中的图像解码器构建的生成器中, 通过将驱动关键点投射到源关键点上进行动作迁移, 并结合源图像特征生成最终的目标活化图像。然后根据目标活化图像与驱动图像的差异最小化为训练目标, 对人脸活化生成模型进行迭代训练, 得到生成的活化图像中无动作畸变的人脸活化生成模型。

[79]本说明书实施例提供的人脸活化生成模型训练方法利用三维空间内从驱动图像到源图像的人脸动作信息投射实现动作迁移, 能够很好地改善头部转动导致的动作畸变问题; 通过使目标活化图像与驱动图像的差异最小化, 较大程度地提高了目标活化图像的逼真度, 使源图像中的人脸身份信息可以较准确且完整地保留, 所生成的目标活化图像能够更有效地测试人脸识别算法的抗攻击性。基于人脸三维姿态信息进行关键点对齐矫正能够将原来 2D 的关键点信息投射到 3D 的人脸三维姿态信息中, 人脸活化从而转变为伪 3D 的活化, 使得人脸关键点信息与人脸纹理特征更好地融合, 生成更逼真的活化图像。

[80]在本说明书的另一个实施例中, 提出了一种人脸活化方法, 图 3 示例性地显示了本说明书实施例所述的人脸活化方法在一种实施方式下的流程示意图。

[81]如图 3 所示, 包括如下步骤。

[82]200: 获取待活化的源图像和包含目标动作的驱动视频, 从驱动视频中抽取若干帧驱动图像。

[83]在采集样本时, 可以直接通过安装本地图像采集设备输入图像, 也可以从云端调取异地的图像采集设备进行采集, 或者从云端调取预先准备好的图像集直接作为模型输入。

[84]在进行人脸活化时, 源图像中的人脸与驱动视频中的人脸可以不是同一人, 换句话说, 使用该人脸活化方法可以将驱动视频中的人脸替换为源图像中的目标人脸。

[85]202: 针对获得的每帧驱动图像,将源图像和驱动图像输入人脸活化生成模型中,生成源图像对应的目标活化图像,其中,人脸活化生成模型是采用如上任一方法所述的步骤训练得到的。

[86]人脸活化生成模型通过提取源图像的人脸动作信息和驱动图像的人脸动作信息,在三维空间下将各自的人脸动作信息矫正为有利于进行动作迁移的姿态,从而生成逼真度极高的目标活化图像,能够更有效地用于对人脸识别算法进行抗攻击性测试,并解决动作畸变问题,较为完整地保留源图像中的人脸身份信息。

[87]204: 将获得的所有目标活化图像进行连接,得到目标活化视频。

[88]基于抽取的驱动图像帧的顺序对生成的对应目标活化图像帧进行连接,即可得到人脸已替换为源图像中人脸的活化视频。

[89]在本说明书的一个实施例中,提出了一种人脸活化生成模型训练装置,图4示例性地显示了本说明书实施例所述的人脸活化生成模型训练装置在一种实施方式下的结构示意图。

[90]如图4所示,包括:样本获取模块30,用于获取源图像和驱动图像;信息提取模块32,用于将源图像和驱动图像输入人脸活化生成模型中,分别获取源图像的人脸动作信息和驱动图像的人脸动作信息;对源图像进行特征编码,得到源图像特征;生成模块34,用于在三维空间内将驱动图像的人脸动作信息投射到源图像的人脸动作信息上,以使驱动图像的动作迁移至源图像中;根据驱动图像的人脸动作信息、源图像的人脸动作信息和源图像特征生成源图像对应的目标活化图像;训练模块36,用于以目标活化图像与驱动图像的差异最小化为训练目标,对人脸活化生成模型进行训练。

[91]样本获取模块获取到的源图像和驱动图像需要包含明显不同的人脸动作,以更直观地观察动作迁移的效果。样本获取模块可以从现有的包含不同人脸动作的图像集中提取源图像和驱动图像,也可以通过摄像头等采集设备拍摄不同人的图像或者若干段视频,其中每段视频都记录了同一个人的脸动作变化过程。

[92]在一些实施例中,样本获取模块每次获取一段视频作为训练样本,从中抽取人脸动作不同的两帧图像分别作为源图像A和驱动图像B,人脸活化的目的是将驱动图像B中的动作迁移到源图像A中的人脸上。针对每一轮训练,利用同一个人的脸作为样本数据对人脸活化生成模型进行训练,这样模型可以得到充分的训练,从而更容易收敛。

[93]信息提取模块可以采取通用的人脸动作检测方法分别获取源图像的人脸动作信息和

驱动图像的人脸动作信息。人脸动作信息能够在三维空间内表现头部的运动情况以及五官的位置变化情况，相对于二维空间的动作信息来说具有更强的逼真度和表征性。

[94]在一些实施例中，人脸动作信息包括人脸三维姿态信息以及人脸关键点信息。相应地，人脸动作检测方法包括人脸三维姿态检测方法和人脸关键点检测方法。人脸三维姿态信息可以体现三维空间内头部的姿态，以确定人脸的朝向；而人脸关键点信息通过在脸部眼睛、鼻子、嘴巴等关键五官或者其它脸部关键区域添加关键点来记录位置信息，从而根据这些关键区域的位置变化表现人脸动作。

[95]在一些更具体的实施例中，人脸三维姿态信息包括头部转动角度。人脸三维姿态信息可以用人脸 3D 位姿信息来表示，用于描述头部在三维空间中互相垂直的三个坐标轴上转动的角度，换句话说，人脸位姿可以体现头部的朝向，表达头部位置信息。这三个互相垂直的坐标轴基于点头、摇头以及摆头动作所沿的方向构建以头部为中心的 XYZ 轴，能够体现头部分别在三个坐标轴上的转动角度，并汇总为人脸 3D 位姿信息。

[96]信息提取模块还将源图像输入通用的图像编码器进行特征编码，输出对应的源图像特征，以便基于源图像特征对源图像作动作迁移等进一步处理，配合图像解码器生成目标活化图像。

[97]在一些更具体的实施例中，生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点；将驱动关键点投射到源关键点上，以将驱动图像的动作迁移至源图像中。

[98]由于人脸三维姿态信息表现了三维空间内头部的姿态，可以基于此得到源图像的头部姿态与驱动图像的头部姿态。图像中的人脸关键点信息与头部姿态是对应的，为了使驱动图像中的动作与源图像中的人脸更好地融合，生成模块基于二者的头部姿态差异进行关键点矫正，使源图像中的关键点与驱动图像中的关键点背后所展现的人脸位于同一姿态下，得到源关键点和驱动关键点，以便进一步比对应位置相同与位置不同的关键点，位置不同的关键点即代表了人脸动作的变化，是需要进行重点迁移的对象。生成模块通过将驱动关键点投射到源关键点上，根据位置不同的关键点对源关键点进行调整，从而在源图像中获得需要迁移的人脸动作信息，完成驱动图像的动作迁移。将源图像的头部姿态与驱动图像的头部姿态进行统一，有利于将人脸调整至更方便进行动作迁移的姿态，获得更好的人脸活化效果，同时避免由于头部姿态的不同对活化结果产生影响。

[99]在一些更具体的实施例中，生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点。

[100]生成模块在对人脸关键点信息进行矫正时，可以将人脸关键点旋转至目标姿态对应的关键点位置，也可以对人脸关键点进行偏置，当然，还可以先将人脸关键点旋转至目标姿态对应的大致关键点位置，再经过偏置对人脸关键点位置进行进一步调整。

[101]在一些更具体的实施例中，生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，利用矩阵变化将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点。

[102]生成模块基于源图像的人脸关键点信息和驱动图像的人脸关键点信息，选择进行关键点矫正的目标姿态，并通过通用的关键点检测方法获取该目标姿态对应的目标关键点，进而分别确定将源图像的人脸关键点信息和驱动图像的人脸关键点信息调整至目标关键点的仿射变换矩阵，利用对应的仿射变换矩阵分别对源图像的人脸关键点信息和驱动图像的人脸关键点信息进行矫正以调整至目标姿态。

[103]生成模块采用通用的图像解码器结合源图像特征与动作迁移获得的人脸动作信息生成目标活化图像。在一些实施例中，生成模块中的图像解码器可以基于 GAN 生成网络中的图像解码器构建。

[104]图像之间的差异可以通过计算图像中逐像素值的差异获得，在一些实施例中训练模块还可以将需要比对的图像进一步处理为其他层面的图像，再计算逐像素值差异，例如提取待比对图像的特征图。优选地，训练模块通过多种途径计算目标活化图像与驱动图像之间的差异，再利用获得的多种损失训练模型有利于提高人脸活化生成模型的训练效率，加快模型性能的提升。

[105]在一些实施例中，人脸活化生成模型通过生成对抗网络实现，生成对抗网络包括判别器和生成器；训练模块具体包括：将驱动图像的人脸动作信息、源图像的人脸动作信息和源图像特征输入生成器，输出源图像对应的目标活化图像；利用判别器判断目标活化图像是否为驱动图像；根据目标活化图像与驱动图像计算生成损失；根据判别器的判别结果计算判别损失；分别以判别损失和生成损失最小化为训练目标，对生成对抗网络进行训练。

[106]生成对抗网络由判别器和生成器组成，在训练生成对抗网络的过程中，首先训练判别器，使其具有良好的判别能力；再训练生成器，使得生成器生成的结果能过骗过判别能力良好的判别器，拥有更强的逼真度。在上述实施例中，训练模块获得目标活化图像后将其输入判别器，以判断是真实的样本图像还是经过篡改的伪造图像，获得判别结果；再根据判别结果的准确性计算判别损失，通过使判别损失最小化训练判别器的判别能力。

[107]优选地，针对人脸活化生成模型的每一轮训练，采用的训练样本中源图像与驱动图像为同一个人。训练模块通过人脸活化将驱动图像中的人脸动作迁移至源图像中，与源图像中的人脸融合，所获得的目标活化图像在理想情况下应该与驱动图像一致，因此训练模块通过计算驱动图像与目标活化图像之间的差异确定生成损失，通过令生成损失最小化优化生成器的生成效果，使驱动图像与源图像融合得更加自然。

[108]随着生成器的不断优化，生成的目标活化图像越来越逼真，使得判别器越来越难以判断其真伪，当判别器已无法判断目标活化图像的真伪时，对人脸活化生成模型的生成对抗式训练完成。

[109]在不断地通过计算损失进行迭代训练的过程中，人脸活化生成模型还学习并迁移了源图像与驱动图像中的人脸纹理特征，即该人脸活化过程将表示人脸三维位置的人脸动作信息与人脸纹理特征进行关联与融合，实现三维空间下的人脸活化，改善了动作迁移后人脸动作发生畸变的问题。

[110]在一些更具体的实施例中，采用 GAN 网络模型构建生成对抗网络来实现人脸活化生成模型。

[111]在一些更具体的实施例中，训练模块用于计算目标活化图像与驱动图像之间基于像素的第一损失；计算目标活化图像与驱动图像之间基于特征图的第二损失；将第一损失和第二损失加权融合，作为生成损失。

[112]优选地，第一损失可以通过训练模块计算目标活化图像与驱动图像之间逐像素的均方差损失获得。

[113]优选地，第二损失通过训练模块计算目标活化图像与驱动图像之间的人脸相似度获得，可以分别提取出目标活化图像的特征向量与驱动图像的特征向量后，采用 Arcface（又称加性角度间隔损失函数）与交叉熵结合的方式计算人脸相似度。通过 Arcface 对人脸活化生成模型进行训练可以在保证类间距离的情况下有效地增大类间距离，同时达到减小类内距离的效果，使得人脸活化生成模型能够更高效地生成更准确的活化结果。

[114]在本说明书的另一个实施例中，提出了一种人脸活化装置，图 5 示例性地显示了本说明书实施例所述的人脸活化装置在一种实施方式下的结构示意图。

[115]如图 5 所示，包括：样本采集模块 40，用于获取待活化的源图像和包含目标动作的驱动视频，从驱动视频中抽取若干帧驱动图像；活化生成模块 42，用于针对获得的每帧驱动图像，将源图像和驱动图像输入人脸活化生成模型中，生成源图像对应的目标活化图像，其中，人脸活化生成模型是采用如上任一方法所述的步骤训练得到的；连接模块 44，用于将获得的所有目标活化图像进行连接，得到目标活化视频。

[116]在样本采集模块采集样本时，可以直接通过安装本地图像采集设备输入图像，也可以从云端调取异地的图像采集设备进行采集，或者从云端调取预先准备好的图像集直接作为模型输入。在样本采集模块中，源图像中的人脸与驱动视频中的人脸可以不是同一人，换句话说，使用该人脸活化装置可以将驱动视频中的人脸替换为源图像中的目标人脸。

[117]活化生成模块中的人脸活化生成模型通过提取源图像的人脸动作信息和驱动图像的人脸动作信息，在三维空间下将各自的人脸动作信息矫正为有利于进行动作迁移的姿态，从而生成逼真度极高的目标活化图像，能够更有效地用于对人脸识别算法进行攻击性测试，并解决动作畸变问题，较为完整地保留源图像中的人脸身份信息。

[118]连接模块基于抽取的驱动图像帧的顺序对生成的对应目标活化图像帧进行连接，即可得到人脸已替换为源图像中人脸的活化视频。

[119]在本说明书的一个实施例中，还提出了一种电子设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，处理器执行程序时实现如上任一方法所述的步骤。

[120]上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下，在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外，在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中，多任务处理和并行处理也是可以的或者可能是有利的。

[121]需要注意的是，以上列举的仅为本发明的具体实施例，显然本发明不限于以上实施例，随之有着许多的类似变化。本领域的技术人员如果从本发明公开的内容直接导出或联想到的所有变形，均应属于本发明的保护范围。

权利要求书

1.一种人脸活化生成模型训练方法，所述方法包括：

获取源图像和驱动图像；

将所述源图像和所述驱动图像输入所述人脸活化生成模型中，分别获取源图像的人脸动作信息和驱动图像的人脸动作信息；

对源图像进行特征编码，得到源图像特征；

在三维空间内将所述驱动图像的人脸动作信息投射到所述源图像的人脸动作信息上，以使驱动图像的动作迁移至源图像中；

根据所述驱动图像的人脸动作信息、所述源图像的人脸动作信息和所述源图像特征生成源图像对应的目标活化图像；

以所述目标活化图像与所述驱动图像的差异最小化为训练目标，对所述人脸活化生成模型进行训练。

2.如权利要求1所述的人脸活化生成模型训练方法，所述人脸动作信息包括人脸三维姿态信息以及人脸关键点信息。

3.如权利要求2所述的人脸活化生成模型训练方法，在三维空间内将所述驱动图像的人脸动作信息投射到所述源图像的人脸动作信息上，以使驱动图像的动作迁移至源图像中具体包括：

通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点；

将所述驱动关键点投射到所述源关键点上，以将驱动图像的动作迁移至源图像中。

4.如权利要求2所述的人脸活化生成模型训练方法，所述人脸三维姿态信息包括头部转动角度。

5.如权利要求1所述的人脸活化生成模型训练方法，所述人脸活化生成模型通过生成对抗网络实现，所述生成对抗网络包括判别器和生成器；

以所述目标活化图像与所述驱动图像的差异最小化为训练目标，对所述人脸活化生成模型进行训练具体包括：

将所述驱动图像的人脸动作信息、所述源图像的人脸动作信息和所述源图像特征输入所述生成器，输出源图像对应的目标活化图像；

利用判别器判断所述目标活化图像是否为所述驱动图像；

根据所述目标活化图像与所述驱动图像计算生成损失；

根据所述判别器的判别结果计算判别损失；

分别以所述判别损失和所述生成损失最小化为训练目标，对所述生成对抗网络进行训练。

6.如权利要求5所述的人脸活化生成模型训练方法，根据所述目标活化图像与所述驱动图像计算生成损失具体包括：

计算所述目标活化图像与所述驱动图像之间基于像素的第一损失；

计算所述目标活化图像与所述驱动图像之间基于特征图的第二损失；

将所述第一损失和所述第二损失加权融合，作为所述生成损失。

7.如权利要求3所述的人脸活化生成模型训练方法，通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点具体包括：

通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点。

8.如权利要求7所述的人脸活化生成模型训练方法，通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点具体包括：

通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，利用矩阵变化将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点。

9.一种人脸活化方法，所述方法包括：

获取待活化的源图像和包含目标动作的驱动视频，从所述驱动视频中抽取若干帧驱动图像；

针对获得的每帧驱动图像，将所述源图像和所述驱动图像输入所述人脸活化生成模型中，生成源图像对应的目标活化图像，其中，所述人脸活化生成模型是采用如权利要求1至8任一所述的方法训练得到的；

将获得的所有目标活化图像进行连接，得到目标活化视频。

10.一种人脸活化生成模型训练装置，所述装置包括：

样本获取模块，用于获取源图像和驱动图像；

信息提取模块，用于将所述源图像和所述驱动图像输入所述人脸活化生成模型中，

分别获取源图像的人脸动作信息和驱动图像的人脸动作信息；对源图像进行特征编码，得到源图像特征；

生成模块，用于在三维空间内将所述驱动图像的人脸动作信息投射到所述源图像的人脸动作信息上，以使驱动图像的动作迁移至源图像中；根据所述驱动图像的人脸动作信息、所述源图像的人脸动作信息和所述源图像特征生成源图像对应的目标活化图像；

训练模块，用于以所述目标活化图像与所述驱动图像的差异最小化为训练目标，对所述人脸活化生成模型进行训练。

11.如权利要求 10 所述的人脸活化生成模型训练装置，所述人脸动作信息包括人脸三维姿态信息以及人脸关键点信息。

12.如权利要求 11 所述的人脸活化生成模型训练装置，

所述生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息，将源图像的人脸关键点信息和驱动图像的人脸关键点信息矫正到同一姿态下，得到三维空间下的源关键点和驱动关键点；

将所述驱动关键点投射到所述源关键点上，以将驱动图像的动作迁移至源图像中。

13.如权利要求 11 所述的人脸活化生成模型训练装置，所述人脸三维姿态信息包括头部转动角度。

14.如权利要求 10 所述的人脸活化生成模型训练装置，所述人脸活化生成模型通过生成对抗网络实现，所述生成对抗网络包括判别器和生成器；

所述训练模块具体包括：

将所述驱动图像的人脸动作信息、所述源图像的人脸动作信息和所述源图像特征输入所述生成器，输出源图像对应的目标活化图像；

利用判别器判断所述目标活化图像是否为所述驱动图像；

根据所述目标活化图像与所述驱动图像计算生成损失；

根据所述判别器的判别结果计算判别损失；

分别以所述判别损失和所述生成损失最小化为训练目标，对所述生成对抗网络进行训练。

15.如权利要求 14 所述的人脸活化生成模型训练装置，

所述训练模块计算所述目标活化图像与所述驱动图像之间基于像素的第一损失；

计算所述目标活化图像与所述驱动图像之间基于特征图的第二损失；

将所述第一损失和所述第二损失加权融合，作为所述生成损失。

16.如权利要求 12 所述的人脸活化生成模型训练装置，

所述生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息,将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下,得到三维空间下的源关键点和驱动关键点。

17.如权利要求 16 所述的人脸活化生成模型训练装置,

所述生成模块通过比对源图像的人脸三维姿态信息和驱动图像的人脸三维姿态信息,利用矩阵变化将源图像的人脸关键点信息和驱动图像的人脸关键点信息经过旋转或者/以及偏置矫正到同一姿态下,得到三维空间下的源关键点和驱动关键点。

18.一种人脸活化装置,所述装置包括:

样本采集模块,用于获取待活化的源图像和包含目标动作的驱动视频,从所述驱动视频中抽取若干帧驱动图像;

活化生成模块,用于针对获得的每帧驱动图像,将所述源图像和所述驱动图像输入所述人脸活化生成模型中,生成源图像对应的目标活化图像,其中,所述人脸活化生成模型是采用如权利要求 1 至 8 任一所述的方法训练得到的;

连接模块,用于将获得的所有目标活化图像进行连接,得到目标活化视频。

19.一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述程序时实现上述权利要求 1-9 中任一所述方法的步骤。

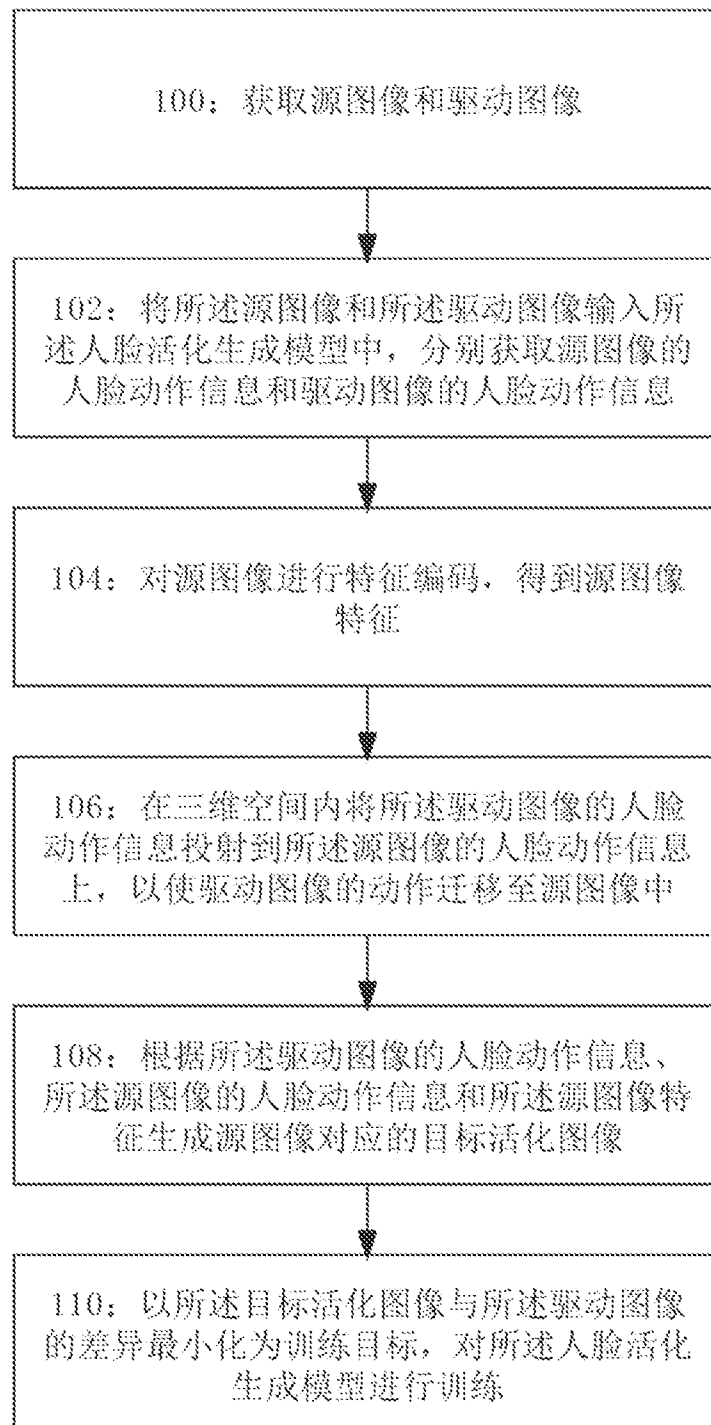


图 1

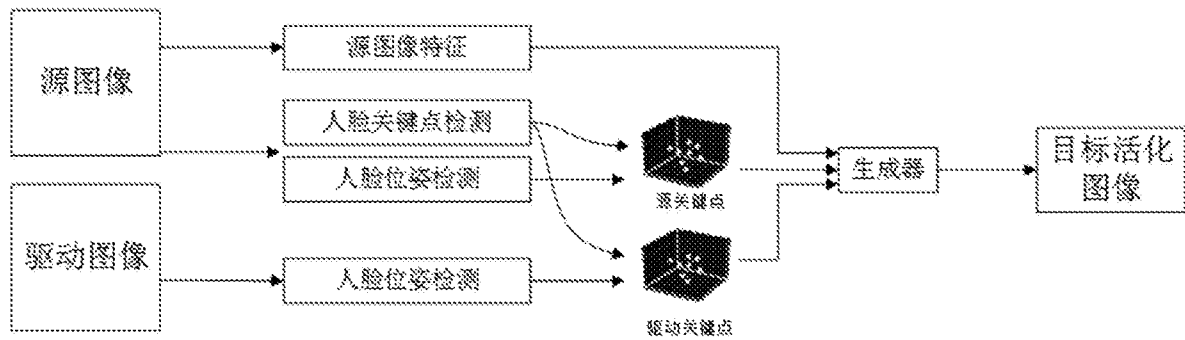


图 2

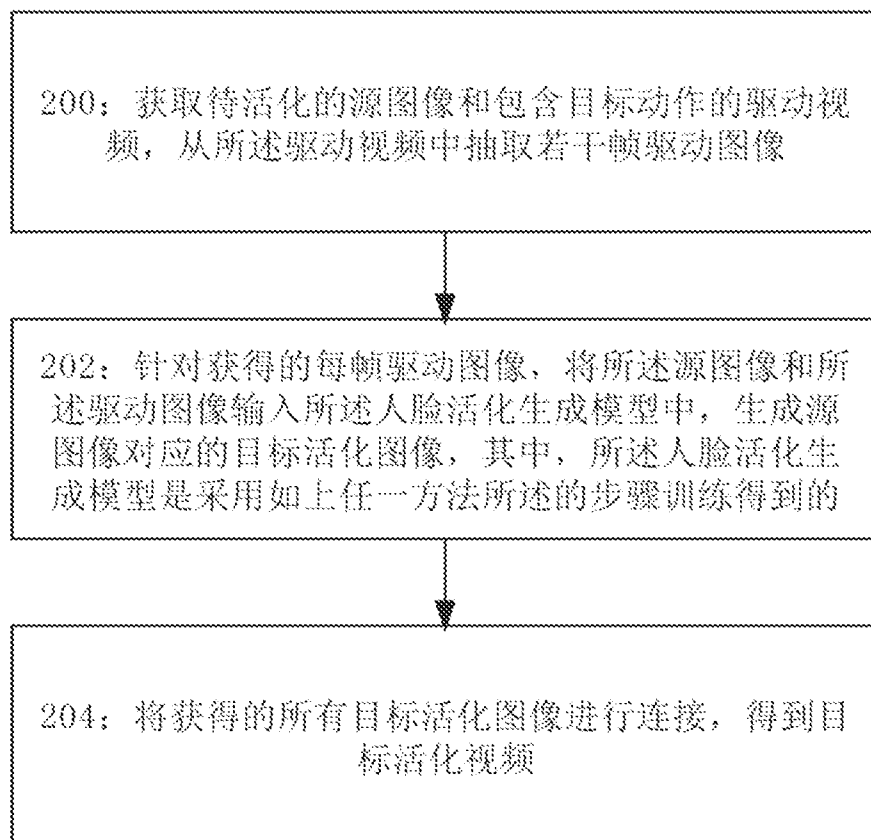


图 3

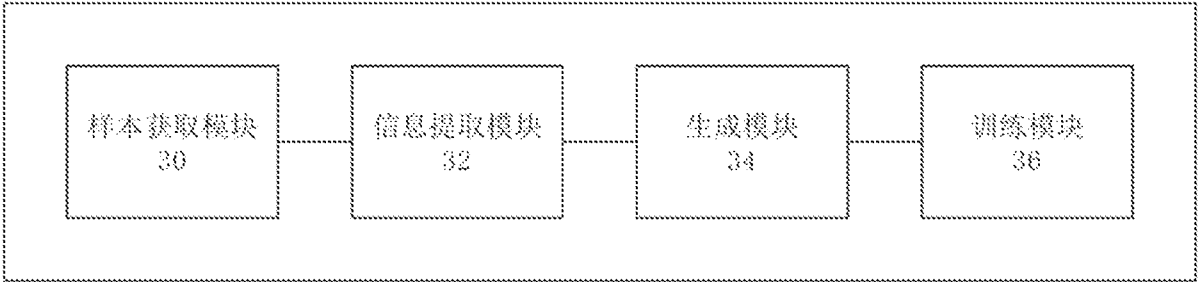


图 4

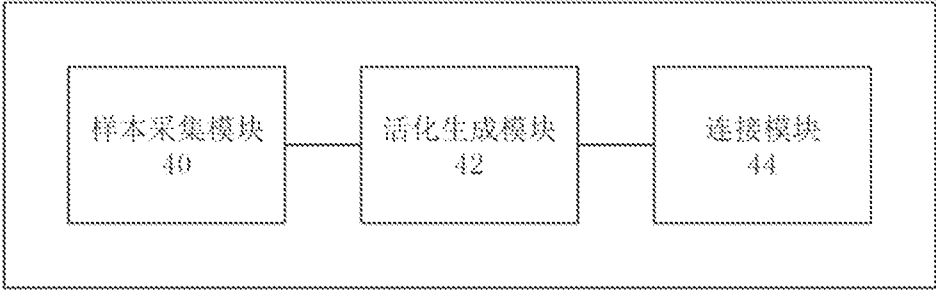


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/103675

A. CLASSIFICATION OF SUBJECT MATTER

G06V 10/774(2022.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC:G06V

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; VEN; WOTXT; EPTXT; CNKI; USTXT; IEEE: 人脸, 合成, 生成, 换脸, 活体, 伪造, 姿态, 姿势, 动作, 关键点, 特征点, 生成对抗网络, face, synthetic, generated, face swapping, live, counterfeit, gesture, pose, action, key point, feature point, GAN

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 116935156 A (ALIPAY (HANGZHOU) INFORMATION TECHNOLOGY CO., LTD.) 24 October 2023 (2023-10-24) description, paragraphs 6-72	1-19
X	CN 116311460 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 23 June 2023 (2023-06-23) description, paragraphs 70-244	1-19
X	CN 113807265 A (SHANDONG UNIVERSITY OF FINANCE AND ECONOMICS) 17 December 2021 (2021-12-17) description, paragraphs 6-33	1-19
A	CN 112233012 A (SHANGHAI JIAO TONG UNIVERSITY) 15 January 2021 (2021-01-15) entire document	1-19
A	US 2023162376 A1 (VIRNECT INC.) 25 May 2023 (2023-05-25) entire document	1-19



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

14 October 2024

Date of mailing of the international search report

17 October 2024

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/103675

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	116935156	A	24 October 2023	None			
CN	116311460	A	23 June 2023	None			
CN	113807265	A	17 December 2021	CN	113807265	B	06 May 2022
CN	112233012	A	15 January 2021	CN	112233012	B	31 October 2023
US	2023162376	A1	25 May 2023	EP	4184437	A1	24 May 2023
				KR	20230076048	A	31 May 2023

A. 主题的分类

G06V 10/774(2022.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC:G06V

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS;CNTXT;VEN;WO TXT;EPTXT;CNKI;USTXT;IEEE:人脸,合成,生成,换脸,活体,伪造,姿态,姿势,动作,关键点,特征点,生成对抗网络,face,synthetic,generated,face swapping,live,counterfeit,gesture,pose,action,key point,feature point,GAN

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 116935156 A (支付宝(杭州)信息技术有限公司) 2023年10月24日 (2023 - 10 - 24) 说明书第6-72段	1-19
X	CN 116311460 A (平安科技(深圳)有限公司) 2023年6月23日 (2023 - 06 - 23) 说明书第70-244段	1-19
X	CN 113807265 A (山东财经大学) 2021年12月17日 (2021 - 12 - 17) 说明书第6-33段	1-19
A	CN 112233012 A (上海交通大学) 2021年1月15日 (2021 - 01 - 15) 全文	1-19
A	US 2023162376 A1 (VIRNECT INC) 2023年5月25日 (2023 - 05 - 25) 全文	1-19

☐ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“p” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2024年10月14日

国际检索报告邮寄日期

2024年10月17日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

洪汇隆

电话号码 (+86) 020-28957215

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/103675

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	116935156	A	2023年10月24日	无	
CN	116311460	A	2023年6月23日	无	
CN	113807265	A	2021年12月17日	CN	113807265 B 2022年5月6日
CN	112233012	A	2021年1月15日	CN	112233012 B 2023年10月31日
US	2023162376	A1	2023年5月25日	EP	4184437 A1 2023年5月24日
				KR	20230076048 A 2023年5月31日