

(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
2020 年 12 月 3 日 (03.12.2020)



(10) 国际公布号
WO 2020/238374 A1

- (51) 国际专利分类号:
G06K 9/00 (2006.01) **G06N 3/04** (2006.01)
- (21) 国际申请号: PCT/CN2020/081262
- (22) 国际申请日: 2020 年 3 月 26 日 (26.03.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910473174.2 2019年5月31日 (31.05.2019) CN
- (71) 申请人: 广州市百果园信息技术有限公司 (GUANGZHOU BAIGUOYUAN INFORMATION TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省广州市番禺区市桥街兴泰路 274 号 C 栋西塔 5-13 层, Guangdong 511402 (CN)。
- (72) 发明人: 项伟 (XIANG, Wei); 中国广东省广州市番禺区市桥街兴泰路 274 号 C 栋西塔 5-13 层, Guangdong 511402 (CN)。 张小伟 (ZHANG,

Xiaowei); 中国广东省广州市番禺区市桥街兴泰路 274 号 C 栋西塔 5-13 层, Guangdong 511402 (CN)。

(74) 代理人: 北京品源专利代理有限公司 (BEYOND ATTORNEYS AT LAW); 中国北京市海淀区莲花池东路 39 号西金大厦 6 层, Beijing 100036 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,

(54) Title: METHOD, APPARATUS, AND DEVICE FOR FACIAL KEY POINT DETECTION, AND STORAGE MEDIUM

(54) 发明名称: 人脸关键点检测方法、装置、设备及存储介质

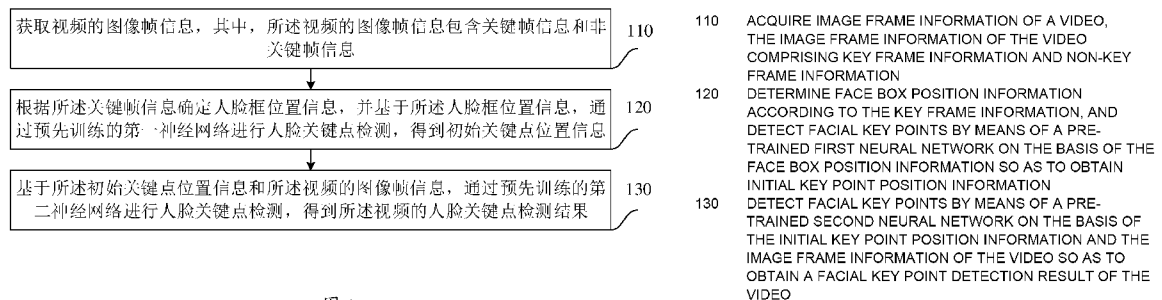


图 1

(57) Abstract: A method, apparatus and device for facial key point detection, and a storage medium, the method comprising: acquiring image frame information of a video, the image frame information of the video comprising key frame information and non-key frame information; determining face box position information according to the key frame information; detecting facial key points by means of a pre-trained first neural network on the basis of the face box position information so as to obtain initial key point position information; and detecting facial key points by means of a pre-trained second neural network on the basis of the initial key point position information and the image frame information of the video so as to obtain a facial key point detection result of the video, the facial key point detection result comprising facial key point position information corresponding to the key frame information and facial key point position information corresponding to the non-key frame information.

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, 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))。

(57) 摘要：一种人脸关键点检测方法、装置、设备以及存储介质，该方法包括：获取视频的图像帧信息，其中，视频的图像帧信息包含关键帧信息和非关键帧信息；根据关键帧信息确定人脸框位置信息；基于人脸框位置信息，通过预先训练的第一神经网络进行人脸关键点检测，得到初始关键点位置信息；基于初始关键点位置信息和视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到视频的人脸关键点检测结果，其中，人脸关键点检测结果包含关键帧信息对应的人脸关键点位置信息和非关键帧信息对应的人脸关键点位置信息。

人脸关键点检测方法、装置、设备及存储介质

本申请要求在2019年05月31日提交中国专利局、申请号为201910473174.2的中国专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本申请涉及计算机视觉技术领域，例如涉及一种人脸关键点检测方法、装置、设备及存储介质。

背景技术

在计算机视觉领域，基于视频数据的算法开发一直以来都受到学术界和工业界的广泛关注。其中，人脸视频数据由于其在生物信息验证、监控安防、视频直播等领域拥有非常现实的应用场景而占有非常重要的地位。人脸关键点的检测属于人脸图像处理当中非常重要的一步，其主要功能是准确地定位出人脸上的关键点在图片上的位置，以为后续的操作做准备，如定位眼睛、鼻子、嘴角、脸部轮廓点等人脸关键点在图片上的位置，以为后续的人脸对齐、人脸识别等操作做准备。

在实现中，人脸关键点检测通常是位于人脸检测之后的一个环节。人脸检测器通常将检测到的人脸位置信息以及相应的人脸图片输入到关键点检测算法，得到人脸的关键点位置，如将以一个矩形框或者正方形框的形式给出的人脸位置信息输入到关键点检测算法中进行计算，以将计算得到的结果确定为人脸的关键点位置。基于深度卷积网络的人脸关键点检测算法在精度上相比传统人脸关键点算法有了很大的提高。然而，基于一个深度卷积网络实现的人脸关键点检测方法通常计算量大，需要对深度卷积网络的网络结构进行精心设计安排，否则很难在计算资源有限的平台上达到实时处理的效果，如很难在手机等移动设备上达到实时处理的效果。

发明内容

本申请实施例提供一种新的人脸关键点检测方法、系统、设备以及存储介质，以解决人脸关键点检测方法在移动端中受计算能力有限、存储空间较小及实时性要求高等限制的问题。

本申请实施例提供了一种人脸关键点检测方法，包括：获取视频的图像帧信息，其中，所述视频的图像帧信息包含关键帧信息和非关键帧信息；根据所述关键帧信息确定人脸框位置信息；基于所述人脸框位置信息，通过预先训练

的第一神经网络进行人脸关键点检测，得到初始关键点位置信息；基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，其中，所述人脸关键点检测结果包含所述关键帧信息对应的人脸关键点位置信息和所述非关键帧信息对应的人脸关键点位置信息。

本申请实施例还提供了一种人脸关键点检测装置，包括：

视频图像帧获取模块，设置为获取视频的图像帧信息，其中，所述视频的图像帧信息包含关键帧信息和非关键帧信息；

第一人脸关键点检测模块，设置为根据所述关键帧信息确定人脸框位置信息，并基于所述人脸框位置信息，通过预先训练的第一神经网络进行人脸关键点检测，得到初始关键点位置信息；

第二人脸关键点检测模块，设置为基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，其中，所述人脸关键点检测结果包含所述关键帧信息对应的人脸关键点位置信息和所述非关键帧信息对应的人脸关键点位置信息。

本申请实施例还提供了一种设备，包括：处理器和存储器；所述存储器中存储有至少一条指令，所述指令由所述处理器执行，使得所述设备执行上述人脸关键点检测方法。

本申请实施例还提供了一种计算机可读存储介质，所述存储介质中的指令由设备的处理器执行时，使得设备能够执行上述人脸关键点检测方法。

附图说明

图 1 是本申请实施例中的一种人脸关键点检测方法的步骤流程示意图；

图 2 是本申请一个可选实施例中的一种人脸关键点检测方法的步骤流程示意图；

图 3 是本申请一个示例中的一个视频中的人脸关键点检测与追踪流程示意图；

图 4 是本申请一个示例中的对上一帧的人脸关键点进行修正的流程示意图；

图 5 是本申请实施例中的一种人脸关键点检测装置实施例的结构方框示意图；

图 6 是本申请一个示例中的一种设备的结构方框示意图。

具体实施方式

下面结合附图和实施例对本申请进行说明。此处所描述的具体实施例仅仅用于解释本申请，而非对本申请的限定。为了便于描述，附图中仅示出了与本申请相关的部分而非全部结构或组成。

大部分的人脸关键点检测算法都是基于单张静态图像去设计的，而对于视频中的人脸关键点的检测，通常是逐帧处理或者运用一般物体追踪算法来追踪人脸再检测人脸关键点。人脸关键点追踪方案可以大致可以分为两类：一类是通过逐帧进行人脸检测和人脸关键点检测；另一类是对第一个图像帧进行人脸检测，再以检测到的人脸为目标运用一般物体追踪方法对后续图像帧进行人脸框追踪，在每一个追踪到的人脸上使用关键点检测算法，如果在一个图像帧追踪失败没有找到人脸，则重新运用人脸检测器检测人脸。其中，第一类方案由于需要对每个图像帧都进行人脸检测和关键点检测，没有充分利用相邻帧之间的关联信息，速度上受到限制；另外，由于每个图像帧都是独立处理，容易出现关键点抖动的问题，影响到后续的依赖关键点稳定性的模块，如影响后续依靠检测出的人脸关键点进行人脸贴纸特效设置的模块，降低用户体验。第二类方案在一个图像帧追踪失败没有找到人脸时重新运用人脸检测器检测人脸，虽然在关键点稳定性方面相比第一类方案要高，但是一方面一般物体追踪方法通常比较耗时，另一方面存在两个潜在问题。一个潜在问题是，由于在视频中人脸会经常出现快速的姿态、尺度、遮挡、表情等变化，如手机视频中人脸会经常出现快速的姿态、尺度、遮挡、表情等变化，导致物体追踪方法失效并且重新使用人脸检测器；另一个问题是，一般人脸关键点算法对于输入的人脸在人脸框内的相对位置比较敏感，即对于扰动输入的人脸框，关键点检测算法在扰动前后输出的结果会有很大差异，而通过追踪得到的人脸框相比通过检测器得到的人脸框人脸贴合程度要低，这会导致关键点检测出现误差。可见，这些人脸关键点检测方法存在计算复杂度高、易丢失追踪目标等问题。另外，大部分人脸关键点检测的应用场景都在诸如手机等移动端，人脸关键点检测方案存在计算能力有限、存储空间较小、实时性要求较高等限制。

为了实现快速、稳定地人脸关键点检测及追踪，本申请实施例提出了一种人脸关键点检测方法。本申请实施例在获取到视频信息后，可以根据该视频信息中的关键帧信息确定出人脸框位置信息，以根据该人脸框位置信息通过第一神经网络确定出初始关键点位置信息，随后可基于该初始关键点位置信息，通过第二神经网络进行人脸关键点检测，得到该视频的人脸关键点检测结果，即利用了两级的神经网络来实现视频信息中的人脸关键点检测，从而能够高效处理视频中人脸关键点的检测。

参照图 1, 示出了本申请实施例中的一种人脸关键点检测方法的步骤流程示意图。该人脸关键点检测方法可用于诸如人脸识别、人脸上的特效贴纸、换脸特效等人脸视觉应用中, 可以包括如下步骤:

步骤 110, 获取视频的图像帧信息, 其中, 所述视频的图像帧信息包含关键帧信息和非关键帧信息。

一个视频可以包含一个或多个视频帧; 每个视频帧可以包含用于显示视频画面的图像帧和/或用于播放视频声音的音频帧等。本实施例中的视频的图像帧信息可以表征视频中的图像帧, 如可以是指视频帧中的图像信息, 可以用于显示视频画面, 使得用户可以观看到视频的播放画面。

本申请实施例在检测视频中的人脸关键点时, 可以获取需要检测的视频的图像帧信息, 以根据该图像帧信息中的人脸图片信息进行人脸关键点检测。其中, 人脸图片信息可以用于表征视频帧中所包含的人脸图片, 如在一视频帧包含一个人的人脸图片时, 可以基于图像帧信息中的人脸图片信息确定这个人在该视频帧中显示的人脸图片; 又如, 在一视频帧包含多个人脸图片的情况下, 可以基于图像帧信息中的人脸图片信息确定在该视频帧中显示的多个人的脸图片等。

本申请实施例在获取到视频的图像帧信息后, 可以将获取到的图像帧信息分为关键帧信息和非关键帧信息, 以基于关键帧信息检测出人脸框位置, 即执行步骤 120。其中, 关键帧信息可以表征视频中关键图像帧 (简称关键帧), 而非关键帧信息可以表征视频中非关键图像帧 (简称非关键帧)。

步骤 120, 根据所述关键帧信息确定人脸框位置信息, 并基于所述人脸框位置信息, 通过预先训练的第一神经网络进行人脸关键点检测, 得到初始关键点位置信息。

本实施例在获取到表征关键帧的关键帧信息后, 可以利用预先设置的人脸检测器, 如作为人脸检测器的联合人脸定位及对齐的多任务级联卷积网络 (Multi-Task Convolutional Neural Network, MTCNN), 对该关键帧信息进行检测, 产生人脸框位置信息。该人脸框位置信息可以表征人脸框位置, 可以确定人脸框在视频的图像帧中的显示位置。随后, 可以基于该人脸框位置信息从该关键帧中裁剪出人脸框图片, 即按照人脸框位置从作为视频关键帧的图像帧中裁剪出包含人脸的人脸框图片, 并可生成对应的人脸框图片信息, 以采用人脸框图片信息表征裁剪出的人脸图片, 然后可以将生成的人脸框图片信息输入到预先训练的第一神经网络进行人脸关键点检测, 以初步检测出人脸关键点的位置, 如可以将第一神经网络的输出信息作为初始关键点位置信息, 以便后续可以采用该初始关键点位置信息初步确定人脸关键点位置, 如确定人脸关键点在

当前关键帧的大概位置。

步骤 130, 基于所述初始关键点位置信息和所述视频的图像帧信息, 通过预先训练的第二神经网络进行人脸关键点检测, 得到所述视频的人脸关键点检测结果。

其中, 所述人脸关键点检测结果包含所述关键帧信息对应的人脸关键点位置信息和所述非关键帧信息对应的人脸关键点位置信息。

本实施例中的视频的人脸关键点检测结果可以用于确定该视频中每个图像帧的人脸关键点位置, 可以包括视频中每个图像帧信息对应的人脸关键点位置信息, 如可以包括关键帧信息对应的人脸关键点位置信息、所述非关键帧信息对应的人脸关键点位置信息等。其中, 图像帧信息对应的人脸关键点位置信息可以用于表征图像帧的人脸关键点位置, 如关键帧信息对应的人脸关键点位置信息可以用于表征关键帧中的人脸关键点位置, 又如非关键帧信息对应的人脸关键点位置信息可以用于表征非关键帧中的人脸关键点位置等。

在实现中, 本申请实施例在确定出初始关键点位置信息后, 可以基于该初始关键点位置信息, 按照人脸关键点在当前关键帧的大概位置生成图片裁剪框, 随后可以采用该图片裁剪框对当前关键帧中人脸图片进行裁剪, 即可以采用该图片裁剪框对该视频的图像帧进行裁剪处理, 得到关键帧人脸图片信息, 并可采用该关键帧人脸图片信息表征此次裁剪处理后得到的图片。随后, 可以将得到的关键帧人脸图片信息输入到预先训练的第二神经网络进行人脸关键点检测, 并可将第二神经网络检测后输出的信息确定为当前关键帧的人脸关键点信息, 以及可以基于该关键帧的人脸关键点信息对视频中的非关键帧进行人脸关键点检测及追踪, 得到非关键帧的人脸关键点信息, 从而可以基于该关键帧的人脸关键点信息和/或非关键帧的人脸关键点信息生成视频的人脸关键点检测结果。

综上, 本申请实施例在获取视频的图像帧信息后, 可以根据该视频的图像帧信息中的关键帧信息确定出人脸框位置信息, 以基于该人脸框位置信息通过第一神经网络进行人脸关键点检测, 得到初始关键点位置信息, 随后可基于该初始关键点位置信息和视频的图像帧信息通过第二神经网络进行人脸关键点检测, 即利用两级的神经网络进行人脸关键点检测, 从而解决了相关技术采用单个深度卷积网络实现人脸关键点检测方案中存在的计算复杂度高、计算量大、实时处理效果差等问题, 能够快速地、稳定地检测到人脸的关键点位置, 达到快速、稳定处理视频中的人脸关键点检测及追踪问题的目的。

在实际处理中, 本实施例在获取视频的图像帧后, 可以按照预设规则从该视频图像帧中选择其中一帧或多帧作为关键帧, 如从视频的每 N 帧图像帧中选

择其中一帧作为关键帧，并可将其余图像帧作为非关键帧，即将与该关键帧相邻的连续 $(N-1)$ 帧的图像帧确定为该关键帧对应的非关键帧，其中， N 的取值可以根据不同的应用场景确定，即， N 的取值可根据不同的应用场景变化。随后，可在该关键帧中利用人脸检测器检测出人脸框位置，以按照该人脸框位置从关键帧信息中裁剪出人脸框图像信息，进而可以通过第一神经网络对该人脸框图像信息进行人脸关键点检测，以检测出该关键帧中大概的人脸关键点。

在上述实施例的基础上，可选的，本实施例提供的人脸关键点检测方法在根据所述关键帧信息确定人脸框位置信息之前，还可以包括：从所述视频的图像帧信息中，选取出关键帧信息和所述关键帧信息对应的非关键帧信息。随后，可以根据该关键帧信息确定出人脸框位置信息，以基于该人脸框位置信息通过第一神经网络确定出人脸关键点大概位置。例如，可以将视频中第 t 帧图片确定为关键帧信息，并将该视频中的第 $(t+1)$ 帧到第 $(t+N-1)$ 帧的图片确定为非关键帧信息，并将这些非关键帧信息与作为关键帧信息的第 t 帧图片相关联，以将这些非关键帧信息确定为上述关键帧信息对应的非关键帧信息。其中， t 可以是大于0的整数。

可选的，上述根据关键帧信息确定出人脸框位置信息，可以包括：将所述关键帧信息输入到人脸检测器中，其中，所述人脸检测器用于检测人脸框位置；将所述人脸检测器的输出信息确定为所述人脸框位置信息。从而，可以基于该人脸框位置信息确定出人脸框位置，以按照人脸框位置从关键帧中裁剪出人脸框图片，生成对应的人脸框图片信息，如使用作为人脸检测器的MTCNN检测出人脸框位置，并对每个人脸框位置对应的方框做正方形扩充处理，即可以以框的中心为正方形中心，框的长边为正方形的边，进行扩充处理，得到这个正方形所裁剪到的人脸框图片信息，并可该人脸框图片信息输入到第一神经网络中进行人脸关键点检测。该第一神经网络可以作为人脸关键点检测过程中的第一级的人脸关键点检测网络，可以用于对视频关键帧的人脸框图片进行关键点检测，输出初始关键点位置信息，以便后续流程也依据该初始关键点位置信息对关键帧和/或非关键帧进行人脸关键点检测，从而能够快速、稳定地检测出视频中人脸关键点位置。其中，初始关键点位置信息可以用于初步确定出人脸关键点大概位置。

在本申请的一个可选实施例中，上述基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，包括：根据所述初始关键点位置信息生成图片裁剪框；通过所述图片裁剪框对所述关键帧信息进行裁剪处理，得到关键帧人脸图片信息，并将所述关键帧人脸图片信息输入到所述第二神经网络进行人脸关键点检测，得到所述关键帧信息对应的人脸关键点位置信息。

本申请实施例在确定出初始关键点位置信息后，可以基于该初始关键点位置信息生成一个图片裁剪框，以采用该图片裁剪框按照人脸关键点大概位置裁剪出当前关键帧对应的人脸图片信息，即对关键帧信息进行裁剪处理，得到关键帧人脸图片信息。随后，可以将关键帧人脸图片信息作为第二神经网络的输入，输入到第二神经网络中进行人脸关键点检测，以精确地确定出该关键帧的人脸关键点位置，并可将第二神经网络的输出信息确定为该关键帧信息对应的人脸关键点位置信息，以便后续可以基于该人脸关键点位置信息对该关键帧信息对应的非关键帧信息进行人脸关键点进行检查及追踪，即利用视频相邻帧间的信息进行人脸关键点追踪，生成非关键帧信息对应的人脸关键点位置信息，从而可以基于关键帧信息对应的人脸关键点位置信息和/非关键帧信息对应的人脸关键点信息生成视频的人脸关键点检测结果，达到高速处理视频中人脸关键点的检测的目的。

在上述实施例的基础上，可选的，本申请实施例中基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，还可以包括：通过所述图片裁剪框对所述关键帧信息对应的非关键帧信息进行裁剪处理，得到非关键帧图片信息；当所述非关键帧图片信息包含人脸图片信息时，依据所述关键帧信息对应的人脸关键点位置信息生成非关键帧人脸图片信息，并将所述非关键帧人脸图片信息输入到所述第二神经网络进行人脸关键点检测，得到所述非关键帧信息对应的人脸关键点位置信息。

参照图 2，示出了本申请一个可选实施例中的一种人脸关键点检测方法的步骤流程图示意图。该人脸关键点检测方法可以包括如下步骤：

步骤 210，获取视频的图像帧信息。

其中，所述视频的图像帧信息包含关键帧信息和非关键帧信息。

步骤 220，从所述视频的图像帧信息中，选取出关键帧信息和所述关键帧信息对应的非关键帧信息。

步骤 230，将所述关键帧信息输入到人脸检测器中。

其中，所述人脸检测器用于检测人脸框位置。

步骤 240，将所述人脸检测器的输出信息确定为所述人脸框位置信息。

本申请实施例在从视频中选出关键帧信息后，可以将该关键帧信息输入到人脸检测器中，以通过人脸检测器检测出关键帧的人脸框位置，即可基于人脸检测器的输出信息确定为人脸框位置信息，从而基于该人脸框位置信息按照人脸框位置从关键帧裁剪出人脸框图片信息进行初步的人脸关键点检测，即执行

步骤 250。

步骤 250, 基于所述人脸框位置信息, 通过预先训练的第一神经网络进行人脸关键点检测, 得到初始关键点位置信息。

步骤 260, 根据所述初始关键点位置信息生成图片裁剪框。

步骤 270, 通过所述图片裁剪框对所述关键帧信息进行裁剪处理, 得到关键帧人脸图片信息, 并将所述关键帧人脸图片信息输入到所述第二神经网络进行人脸关键点检测, 得到所述关键帧信息对应的人脸关键点位置信息。

本申请实施例在确定出人脸框位置信息后, 可以基于该人脸框位置信息按照人脸框位置从关键帧中裁剪出人脸框图片信息, 以将该裁剪出的人脸图片信息输入到第一神经网络中进行人脸关键点检测, 得到初始关键点位置信息, 从而可以基于初始关键点位置信息生成一个图片裁剪框, 以采用该图片裁剪框按照人脸关键点大概的位置对关键帧进行裁剪处理, 得到关键帧人脸图片信息。关键帧人脸图片信息可以用于表征视频关键帧中的人脸图片。随后, 可以将关键帧人脸图片信息作为第二神经网络的输入, 输入到第二神经网络中进行人脸关键点检测, 以基于第二神经网络所输出的信息精确、稳定地确定出关键帧中人脸关键点位置, 如可将该第二神经网络所输出的信息确定为关键帧信息对应的人脸关键点位置信息, 以便后续可以基于该关键帧信息对应的人脸关键点位置信息对非关键帧的人脸关键点进行检测及追踪。

作为本申请的一个示例, 在从视频的每 N 帧选取一帧作为关键帧, 其余帧作为非关键帧的情况下, 如图 3 所示, 可以将视频中第 t 帧、第 $t+N$ 帧确定为关键帧信息, 并可以在每个关键帧中, 使用作为人脸检测器的 MTCNN 检测出人脸框位置, 并可对该 MTCNN 确定出的每一个框作正方形扩充处理, 以根据这个正方形裁剪人脸图片, 如图 3 中所示的裁剪人脸图片 I 模块, 并且可以将裁剪得到人脸图片缩放到宽、高均为 70 像素后输入到作为第一神经网络的人脸关键点检测网络 C 进行处理, 得到作为初始关键点位置信息的 106 个人脸关键点坐标。随后, 可以根据该 106 个人脸关键点坐标构成最小正方形框裁剪一个人脸图片, 如图 3 中所示的裁剪人脸图片 II 模块, 即可将构成的最小正方形框作为图片裁剪框, 以采用该图片裁剪框对关键帧信息进行裁剪处理, 得到关键帧人脸图片信息, 如可将裁剪得到的人脸图片缩放到宽、高均为 70 像素后输入到作为第二神经网络的人脸关键点检测网络 F 中进行处理, 以得到更准确的 106 个人脸关键点坐标, 作为关键帧信息对应的人脸关键点位置信息, 从而可以基于关键帧信息对应的人脸关键点位置信息生对应视频的人脸关键点检测结果, 以便后续可以基于该关键帧信息对应的人脸关键点位置信息进行关键点处理, 并可根椐得到的关键帧信息对应的人脸关键点位置信息对非关键帧的人脸关键

点进行检测及追踪，执行步骤 280，以利用相邻帧间的信息去直接追踪视频中的人脸关键点，达到高效处理视频中人脸关键点的检测的目的。

本示例中的人脸关键点检测网络 C 与人脸关键点检测网络 F 均可以通过多个卷积层（Convolution Layer）和特征池化层（Pooling Layer）去提取特征，并且可以通过全连接层（Fully Connected Layer）来回归关键点的相对位置。虽然这两个人脸关键点检测网络的网络结构一样，但是在人脸关键点检测网络 C 的每一层中使用了更少的通道（Channel）数目，因此作为第一神经网络的人脸关键点检测网络 C 比作为第二神经网络的人脸关键点检测网络 F 更轻量。另外，这两个人脸关键点检测网络的输入图片的裁剪方式不同，人脸关键点检测网络 C 的输入图片可以是通过人脸框裁剪得到的，而人脸关键点检测网络 F 的输入图片可以根据 106 个人脸关键点裁剪得到的，且根据这 106 个人脸关键点位置裁剪得到的输入图片会更加紧贴人脸。此外，这两个人脸关键点检测网络可以是分别独立训练的，且每个卷积层的权重都可以是不同的，以降低由于人脸框不够紧贴人脸而导致关键点不准的影响。由此可见，本申请实施例可以通过一个两级神经网络进行人脸关键点检测的渐进方法，得到更准确的关键点位置，人脸关键点检测网络 C 可以回归出关键点的粗略位置，而人脸关键点检测网络 F 则改善得到更准确的关键点。

步骤 280，通过所述图片裁剪框对所述关键帧信息对应的非关键帧信息进行裁剪处理，得到非关键帧图片信息。

本申请实施例在处理非关键帧时，可以利用图片裁剪框对当前帧进行裁剪处理，如对图 3 中所示的第 $t+1$ 帧图片进行裁剪处理，以基于裁剪出的图片生成对应的非关键帧图片信息。该非关键帧图片信息可以用于表征根据关键帧的人脸关键点位置从视频非关键帧中裁剪出的图片。随后，可以将该非关键帧图片信息作为人脸检测追踪网络的输入，以通过该人脸检测追踪网络对该非关键帧图片信息中的人脸进行检测及追踪，如确定非关键帧图片信息是否包含人脸图片信息。其中，人脸检测追踪网络可以作为非关键帧的人脸检测器，如可以是图 3 中所示的人脸检测器追踪网络（Tracking Net, TNet），该人脸检测器 TNet 可以判断非关键帧图片信息是否包含人脸图片信息，以判断输入图片是否是人脸图片，并可在判断出输入图片是人脸图片时，输出人脸框的相对位置和人脸关键点的相对位置信息等。人脸图片信息可以包括用于表征人脸图片的多种信息，如可以是人脸图片对应的图像信息等，本实施例对此不作限制。

步骤 290，依据所述关键帧信息对应的人脸关键点位置信息以及所述关键帧信息对应的非关键帧信息生成非关键帧人脸图片信息，并将所述非关键帧人脸图片信息输入到所述第二神经网络进行人脸关键点检测，得到所述非关键帧信

息对应的人脸关键点位置信息。

可选地，本申请实施例在得到非关键帧图片信息之后，在所述非关键帧图片信息包含人脸图片信息的情况下，依据所述关键帧信息对应的人脸关键点位置信息以及所述关键帧信息对应的非关键帧信息生成非关键帧人脸图片信息之前，还包括：将所述非关键帧图片信息输入到人脸检测跟踪网中，得到人脸检测跟踪网的输出信息，所述输出信息包含人脸概率信息；基于所述人脸概率信息确定所述非关键帧图片信息是否包含人脸图片信息。人脸概率信息可以用于确定非关键帧是否包含人脸图片，如可以表征该非关键帧包含人脸图片的概率；当人脸概率信息的值超过一定阈值时，可以确定非关键帧图片信息包含人脸图片信息；相应的，在人脸概率信息的值不超过上述阈值时，可以确定非关键帧图片信息不包含人脸图片信息。若非关键帧图片信息不包含人脸图片信息，则可以确定当前的非关键帧不包含人脸图片，进而可以忽略该非关键帧，不对该非关键帧激进行人脸关键点检测。若非关键帧图片信息包含人脸图片信息，则可以确定当前的非关键帧包含人脸图片，并可通过人脸检测追踪网络，利用关键帧的人脸关键点位置对该非关键帧进行人脸检测，以检测出非关键帧中人脸关键点大概的位置，以及生成对应的非关键帧人脸图片信息，即执行步骤 290。其中，非关键帧人脸图片信息可以表征非关键帧的人脸图片信息，可以包括非关键帧信息的人脸关键点信息，如可以包括非关键帧中 5 个人脸关键点坐标等。这 5 个人脸关键点坐标可以分别是左眼中心、右眼中心、鼻尖、嘴巴左角、嘴巴右角的位置坐标。

在本申请的一个可选实施例中，人脸检测跟踪网的输出信息还可以包含人脸框相对位置信息和关键点相对位置信息。在得到非关键帧图片信息之后，在所述非关键帧图片信息包含人脸图片信息的情况下，依据所述关键帧信息对应的人脸关键点位置信息以及所述关键帧信息对应的非关键帧信息生成非关键帧人脸图片信息之前，还可以包括：依据所述人脸框相对位置信息和关键点相对位置信息，确定所述非关键帧信息的人脸关键点信息。其中，人脸框相对位置信息可以表示回归人脸框的相对位置，如可以是人脸检测跟踪网通过输出层输出的一个 4 维向量；关键点相对位置信息可以用于表示 5 个人脸关键点的相对位置，如可以是人脸检测跟踪网通过输出层输出的一个 10 维向量等。

在实际处理中，本申请实施例中的人脸检测跟踪网在接收到输入的非关键帧图片信息后，可以基于该非关键帧图片信息判断该非关键帧所显示的图片是否是人脸图片；如果是人脸图片，则回归该人脸图片在当前帧人脸框的位置，并且可输出其中 5 个人脸关键点左眼中心、右眼中心、鼻尖、嘴巴左角、嘴巴右角的位置坐标，即输出非关键帧信息的人脸关键点信息，以作为该人脸检测跟踪网的输出信息。

结合上述示例,人脸检测跟踪网输出的 5 个人脸关键点信息可以是人脸关键点检测网络 C 和人脸关键点检测网络 F 输出的 106 个关键点当中的一部分,如在人脸检测跟踪网的输出层为全连接层 FC 的情况下,可以通过该全连接层 FC,输出一个 2 维向量(p_0, p_1),作为人脸概率信息,以表示输入图片是/不是人脸的概率,如在 p_0 表示非人脸的概率且 p_1 表示人脸的概率的情况下, p_0 与 p_1 的和可以为 1,即 $p_0+p_1=1$,当 p_1 超过预设的阈值时可以判断为检测到了人脸,当 p 不超过预设的阈值时可以判断为当前输入的图片是非人脸图片;并可输出一个 4 维向量(x_0, y_0, w, h),作为人脸框相对位置信息,以表示回归的人脸框的相对位置,其中, (x_0, y_0)可以是人脸框的左上角在图片中的坐标, (w, h)可以是人脸框的宽和高,如在输入 TNet 的一个作为非关键帧图片信息的方框信息是(x_0, y_0, w, h),输出的 4 维向量是(dx_0, dy_0, dx_1, dy_1)时,可以采用输出的 4 维向量中的这 4 个数表示检测出来的框相对于输入框的相对位置,其对应的检出框是($x_0+dx_0*w, y_0+dy_0*h, (dx_1-dx_0)*w, (dy_1-dy_0)*h$),以便后续可以将该检出框作为非关键帧的人脸框对非关键帧的图片进行裁剪处理;以及可输出一个 10 维向量($dx_0, dy_0, \dots, dx_4, dy_4$),作为关键点相对位置信息,以表示 5 个人脸关键点的相对位置,从而可以确定出该非关键帧的 5 个人脸关键点坐标是($x_0+dx_0*w, y_0+dy_0*h, \dots, x_0+dx_4*w, y_0+dy_4*h$),以作为非关键帧信息的人脸关键点信息,随后可以基于这 5 个人脸关键点坐标对上一帧的人脸关键点位置信息进行修正处理,得到输入到第二神经网络的非关键帧人脸图片信息,如图 3 中所示的裁剪人脸图片 III,进而可以通过第二神经网络对该非关键帧人脸图片信息进行人脸关键点检测,以产生非关键帧的人脸关键点位置信息。相比在关键帧上使用的人脸检测器 MTCNN, TNet 可以是一个计算量更小的网络。因为视频中相邻帧之间人脸在图片中的位置变化不大,由上一帧传递来的人脸关键点位置信息已经给出了人脸在当前帧的大概位置,因此只需要用一个简单的人脸检测追踪网络就可以回归出人脸框的位置。

由于将上一帧的信息作为关键帧信息,用在当前帧上,对于快速的人脸运动可能会出现偏差,因此引入一个关键点修改模块修正这些人脸关键点的坐标位置。在一种可选实施方式中,可以采用一个线性变换的方式,利用 TNet 学习到的当前帧的新信息来修正上一帧传递来的 106 个人脸关键点的位置,从而可以利用修正后的 106 个人脸关键点的坐标构成最小正方形框裁剪一个人脸图片,如图 3 中裁剪人脸图片 III,并可以将该裁剪出的人脸图片缩放到宽、高均为 70 像素后,作为非关键帧人脸图片信息,输入到人脸关键点检测网络 F 进行人脸关键点检测,得到当前帧的 106 个人脸关键点坐标,以作为非关键帧信息对应的人脸关键点位置信息。

可选的,本实施例依据所述关键帧信息对应的人脸关键点位置信息以及所

述关键帧信息对应的非关键帧信息生成非关键帧人脸图片信息，可以包括：基于所述非关键帧信息的人脸关键点信息，对所述关键帧信息对应的人脸关键点位置信息进行修正，得到关键点修正信息；根据所述关键点修正信息和所述初始关键点位置信息，确定关键点追踪位置信息；根据所述关键点追踪位置信息生成人脸图片裁剪框；通过所述人脸图片裁剪框对所述非关键帧信息和/或所述非关键帧图片信息进行裁剪处理，得到所述非关键帧人脸图片信息。

本实施例在人脸关键点的追踪过程中，针对非关键帧，可以采用上一帧的人脸关键点坐标作为当前帧对应人脸关键点的大致位置，以利用相邻帧的信息达到对非关键帧进行人脸关键点检测及追踪的目的。为了应对帧间的变化，增加了一个矫正的步骤，即是以 TNet 输出的 5 个人脸关键点坐标为基准，根据 TNet 输出的 5 个人脸关键点坐标与上一帧对应 5 个人脸关键点坐标的差异计算出线性变换信息 (A^* , b^*)，然后将该线性变化信息 (A^* , b^*) 作用在上一帧的所有 106 个人脸关键点上，得到修正后的 106 个人脸关键点信息，以依据修改后的 106 个人脸关键点信息裁人脸图片，使得裁剪得到的人脸图片会更贴合当前帧的人脸，在效果上起到了关键帧处理流程上人脸关键点检测网络 C 的作用。

例如，结合上述示例，可在非关键帧上，诸如 TNet 等人脸检测跟踪网可回归出 5 个人脸关键点的坐标，如可以记作 $\{(u_1', v_1'), \dots, (u_5', v_5')\}$ ，如图 4 所示，TNet 输出 5 个人脸关键点坐标；并可从上一帧的人脸关键点检测网络 F 输出的 106 个人脸关键点坐标中抽取这 5 个人脸关键点的坐标，记作 $\{(u_1, v_1), \dots, (u_5, v_5)\}$ ，以及可以将人脸关键点检测网络 F 输出的其余 101 个人脸关键点的坐标记作为 $\{(u_6, v_6), \dots, (u_{106}, v_{106})\}$ ，随后可以将抽取出的 5 个人脸关键点的坐标 $\{(u_1, v_1), \dots, (u_5, v_5)\}$ 作为关键帧信息对应的人脸关键点位置信息，以基于该关键帧信息对应的人脸关键点位置信息进行修正，得到关键点修正信息。

作为本申请的一个可选实施方式，可以通过计算公式 $(A^*, b^*) = \operatorname{argmin}_{A, b} \sum_{i=1, \dots, 5} \|(u_i', v_i') - (u_i, v_i) A - b\|_2^2$ 确定出作为关键点修正信息的线性变换信息 (A^* , b^*)。其中， A 可以通过计算公式 $A = S * R = S * \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ 来确定， b 可以根据公式 $b = (b_x, b_y)$ 来确定。 S 可以是表征缩放系数， R 可以是 2x2 的旋转变换矩阵， b 可以是 2 维的位移向量。

一实施例中，线性变换信息 (A^* , b^*) 可以由以下步骤得到：

步骤 S10, 分别根据公式 $(\bar{u}, \bar{v}) = \frac{1}{5} \sum_{i=1, \dots, 5} (u_i, v_i)$ 和公式 $(\bar{u}', \bar{v}') = \frac{1}{5} \sum_{i=1, \dots, 5} (u_i', v_i')$ 求两组人脸关键点的平均坐标, 并可以中心化计算出的这两组人脸关键点的坐标, 如可以按照公式 $(\hat{u}_i, \hat{v}_i) = (u_i, v_i) - (\bar{u}, \bar{v})$ 中心化作为上一帧的人脸关键点的坐标, 且可按照公式 $(\hat{u}_i', \hat{v}_i') = (u_i', v_i') - (\bar{u}', \bar{v}')$ 中心化作为当前帧的人脸关键点的坐标。

步骤 S20, 按照公式 $C = \sum_{i=1, \dots, 5} \begin{bmatrix} \hat{u}_i' & \hat{u}_i & \hat{u}_i' \hat{v}_i' \\ \hat{v}_i' & \hat{v}_i & \hat{v}_i' \hat{v}_i \end{bmatrix}$ 计算 2x2 矩阵 C , 并可按

照公式 $C = U \sum V^T$ 对矩阵 C 奇异值分解, 得到最优的 2x2 旋转矩阵 R^* , 且 $R^* = V^T U$ 。

步骤 S30, 根据最优的 2x2 旋转矩阵 R^* 计算数值 S^* , 如按照计算公式 $S^* = \frac{e}{d}$ 计算得到, 其中, $d = \sum_{i=1, \dots, 5} [\hat{u}_i'^2 + \hat{v}_i'^2]$, $e = \sum_{i=1, \dots, 5} [\hat{u}_i, \hat{v}_i] R^* (\hat{u}_i', \hat{v}_i')^T$ 。

步骤 S40, 根据最优的 2x2 旋转矩阵 R^* 和数值 S^* 确定 A^* 和 b^* , 其中, $A^* = S^* R^*$, $b^* = -A^*(\bar{u}, \bar{v}) + (\bar{u}', \bar{v}')$ 。

随后, 可以通过线性变换信息 (A^*, b^*) , 利用 TNet 学习到的当前帧的新信息来修正上一帧传递来的 106 个人脸关键点的位置, 如可以按照修正公式 $(\tilde{u}_i, \tilde{v}_i) = (u_i, v_i) A^* - b^*$, $i = 1, \dots, 106$ 。将线性变换线性信息 (A^*, b^*) 作用于上一帧所有的 106 个人脸关键点的坐标位置, 从而使得根据修正后 106 个人脸关键点裁剪出得人脸图片更加贴近当前帧的人脸。可以利用修正后的 106 个关键点的坐标构成最小正方形框裁剪一个人脸图片, 如图 3 中裁剪人脸图片 III, 并可将裁剪得到的人脸图片缩放到宽、高均为 70 像素后输入人脸关键点检测网络 F 得到当前帧的 106 个人脸关键点坐标, 即根据所述关键点追踪位置信息生成人脸图片裁剪框, 并通过所述人脸图片裁剪框对所述非关键帧信息和/或所述非关键帧图片信息进行裁剪处理, 得到非关键帧人脸图片信息, 以及将非关键帧人脸图片信息输入到作为第二神经网络的人脸关键点检测网络 F 进行

人脸关键点检测，得到所述非关键帧信息对应的人脸关键点位置信息，以便后续可以基于非关键帧信息对应的人脸关键点位置信息生成视频的人脸关键点检测结果，达到对视频中人脸关键点检测及追踪的目的。

对于方法实施例，为了简单描述，故将其都表述为一系列的动作组合，本申请实施例并不受所描述的动作顺序的限制，因为依据本申请实施例，一些步骤可以采用其他顺序或者同时进行。

参照图 5，示出了本申请实施例中的一种人脸关键点检测装置实施例的结构方框示意图，该人脸关键点检测装置可以包括如下模块：

视频图像帧获取模块 510，设置为获取视频的图像帧信息，其中，所述视频的图像帧信息包含关键帧信息和非关键帧信息；第一人脸关键点检测模块 520，设置为根据所述关键帧信息确定人脸框位置信息，并基于所述人脸框位置信息，通过预先训练的第一神经网络进行人脸关键点检测，得到初始关键点位置信息；第二人脸关键点检测模块 530，设置为基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，其中，所述人脸关键点检测结果包含所述关键帧信息对应的人脸关键点位置信息和所述非关键帧信息对应的人脸关键点位置信息。

在实现中，上述人脸关键点检测装置可以集成在设备中。该设备可以是两个或多个物理实体构成，也可以是一个物理实体构成，如设备可以是个人计算机（Personal Computer，PC）、电脑、手机、平板设备、个人数字助理、服务器、消息收发设备、游戏控制台等。

本申请实施例还提供一种设备，包括：处理器和存储器。存储器中存储有至少一条指令，且指令由所述处理器执行，使得所述设备执行如上述方法实施例中所述的人脸关键点检测方法。

参照图 6，示出了本申请一个示例中的一种设备的结构示意图。如图 6 所示，该设备可以包括：处理器 60、存储器 61、具有触摸功能的显示屏 62、输入装置 63、输出装置 64 以及通信装置 65。该设备的处理器 60、存储器 61、显示屏 62、输入装置 63、输出装置 64 以及通信装置 65 可以通过总线或者其他方式连接，图 6 中以通过总线连接为例。

存储器 61 作为一种计算机可读存储介质，可用于存储软件程序、计算机可执行程序以及模块，如本申请任意实施例所述的人脸关键点检测方法对应的程序指令/模块（例如，人脸关键点检测装置中的视频图像帧获取模块 510、第一人脸关键点检测模块 520 以及第二人脸关键点检测模块 530 等）。

处理器 60 通过运行存储在存储器 61 中的软件程序、指令以及模块，从而执行设备的多种功能应用以及数据处理，即实现上述的人脸关键点检测方法。

实施例中，处理器 60 执行存储器 61 中存储的一个或多个程序时，实现如下操作：获取视频的图像帧信息，其中，所述视频的图像帧信息包含关键帧信息和非关键帧信息；根据所述关键帧信息确定人脸框位置信息，并基于所述人脸框位置信息，通过预先训练的第一神经网络进行人脸关键点检测，得到初始关键点位置信息；基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，其中，所述人脸关键点检测结果包含所述关键帧信息对应的人脸关键点位置信息和所述非关键帧信息对应的人脸关键点位置信息。

本申请实施例还提供一种计算机可读存储介质，所述存储介质中的指令由设备的处理器执行时，使得设备能够执行如上述方法实施例所述的人脸关键点检测方法。示例性的，该人脸关键点检测方法包括：获取视频的图像帧信息，其中，所述视频的图像帧信息包含关键帧信息和非关键帧信息；根据所述关键帧信息确定人脸框位置信息，并基于所述人脸框位置信息，通过预先训练的第一神经网络进行人脸关键点检测，得到初始关键点位置信息；基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，其中，所述人脸关键点检测结果包含所述关键帧信息对应的人脸关键点位置信息和所述非关键帧信息对应的人脸关键点位置信息。

1、一种人脸关键点检测方法，包括：

获取视频的图像帧信息，其中，所述视频的图像帧信息包含关键帧信息和非关键帧信息；

根据所述关键帧信息确定人脸框位置信息，并基于所述人脸框位置信息，通过预先训练的第一神经网络进行人脸关键点检测，得到初始关键点位置信息；

基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，其中，所述人脸关键点检测结果包含所述关键帧信息对应的人脸关键点位置信息和所述非关键帧信息对应的人脸关键点位置信息。

2、根据权利要求1所述的方法，其中，所述基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，包括：

根据所述初始关键点位置信息生成图片裁剪框；

通过所述图片裁剪框对所述关键帧信息进行裁剪处理，得到关键帧人脸图片信息，并将所述关键帧人脸图片信息输入到所述第二神经网络进行人脸关键点检测，得到所述关键帧信息对应的人脸关键点位置信息。

3、根据权利要求2所述的方法，在所述根据所述关键帧信息确定人脸框位置信息之前，还包括：

从所述视频的图像帧信息中，选取出关键帧信息和所述关键帧信息对应的非关键帧信息；

所述基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，还包括：

通过所述图片裁剪框对所述关键帧信息对应的非关键帧信息进行裁剪处理，得到非关键帧图片信息；

在所述非关键帧图片信息包含人脸图片信息的情况下，依据所述关键帧信息对应的人脸关键点位置信息以及所述关键帧信息对应的非关键帧信息生成非关键帧人脸图片信息，并将所述非关键帧人脸图片信息输入到所述第二神经网络进行人脸关键点检测，得到所述非关键帧信息对应的人脸关键点位置信息。

4、根据权利要求3所述的方法，在所述得到非关键帧图片信息之后，在所述非关键帧图片信息包含人脸图片信息的情况下，依据所述关键帧信息对应的人脸关键点位置信息以及所述关键帧信息对应的非关键帧信息生成非关键帧人

脸图片信息之前，还包括：

将所述非关键帧图片信息输入到人脸检测跟踪网中，得到所述人脸检测跟踪网的输出信息，其中，所述输出信息包含人脸概率信息；

基于所述人脸概率信息确定所述非关键帧图片信息是否包含人脸图片信息。

5、根据权利要求4所述的方法，其中，所述输出信息还包含人脸框相对位置信息和关键点相对位置信息；

在所述得到非关键帧图片信息之后，在所述非关键帧图片信息包含人脸图片信息的情况下，依据所述关键帧信息对应的人脸关键点位置信息以及所述关键帧信息对应的非关键帧信息生成非关键帧人脸图片信息之前，所述方法还包括：

依据所述人脸框相对位置信息和所述关键点相对位置信息，确定所述非关键帧信息的人脸关键点信息。

6、根据权利要求5所述的方法，其中，所述依据所述关键帧信息对应的人脸关键点位置信息以及所述关键帧信息对应的非关键帧信息生成非关键帧人脸图片信息，包括：

基于所述非关键帧信息的人脸关键点信息，对所述关键帧信息对应的人脸关键点位置信息进行修正，得到关键点修正信息；

根据所述关键点修正信息和所述初始关键点位置信息，确定关键点追踪位置信息；

根据所述关键点追踪位置信息生成人脸图片裁剪框；

通过所述人脸图片裁剪框对所述非关键帧信息和所述非关键帧图片信息中的至少之一进行裁剪处理，得到所述非关键帧人脸图片信息。

7、根据权利要求1至6任一所述的方法，其中，所述根据所述关键帧信息确定人脸框位置信息，包括：

将所述关键帧信息输入到人脸检测器中，其中，所述人脸检测器用于检测人脸框位置；

将所述人脸检测器的输出信息确定为所述人脸框位置信息。

8、一种人脸关键点检测装置，包括：

视频图像帧获取模块，设置为获取视频的图像帧信息，其中，所述视频的图像帧信息包含关键帧信息和非关键帧信息；

第一人脸关键点检测模块，设置为根据所述关键帧信息确定人脸框位置信息，并基于所述人脸框位置信息，通过预先训练的第一神经网络进行人脸关键点检测，得到初始关键点位置信息；

第二人脸关键点检测模块，设置为基于所述初始关键点位置信息和所述视频的图像帧信息，通过预先训练的第二神经网络进行人脸关键点检测，得到所述视频的人脸关键点检测结果，其中，所述人脸关键点检测结果包含所述关键帧信息对应的人脸关键点位置信息和所述非关键帧信息对应的人脸关键点位置信息。

9、根据权利要求 8 所述的装置，其中，所述第二人脸关键点检测模块包括：

图片裁剪框生成子模块，设置为根据所述初始关键点位置信息生成图片裁剪框；

关键帧裁剪处理子模块，设置为通过所述图片裁剪框对所述关键帧信息进行裁剪处理，得到关键帧人脸图片信息；

关键帧人脸关键点检测子模块，设置为将所述关键帧人脸图片信息输入到所述第二神经网络进行人脸关键点检测，得到所述关键帧信息对应的人脸关键点位置信息。

10、一种设备，包括：处理器和存储器；

所述存储器中存储有至少一条指令，所述至少一条指令由所述处理器执行，使得所述设备执行如权利要求 1 至 7 任一项所述的人脸关键点检测方法。

11、一种计算机可读存储介质，其中，所述存储介质中的指令由设备的处理器执行时，使得所述设备执行如权利要求 1 至 7 任一项所述的人脸关键点检测方法。

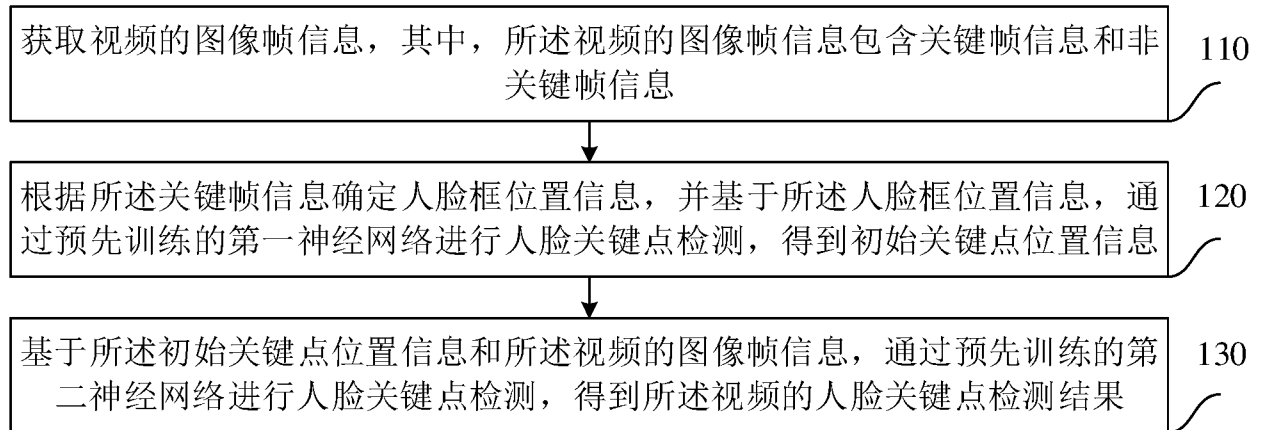


图 1

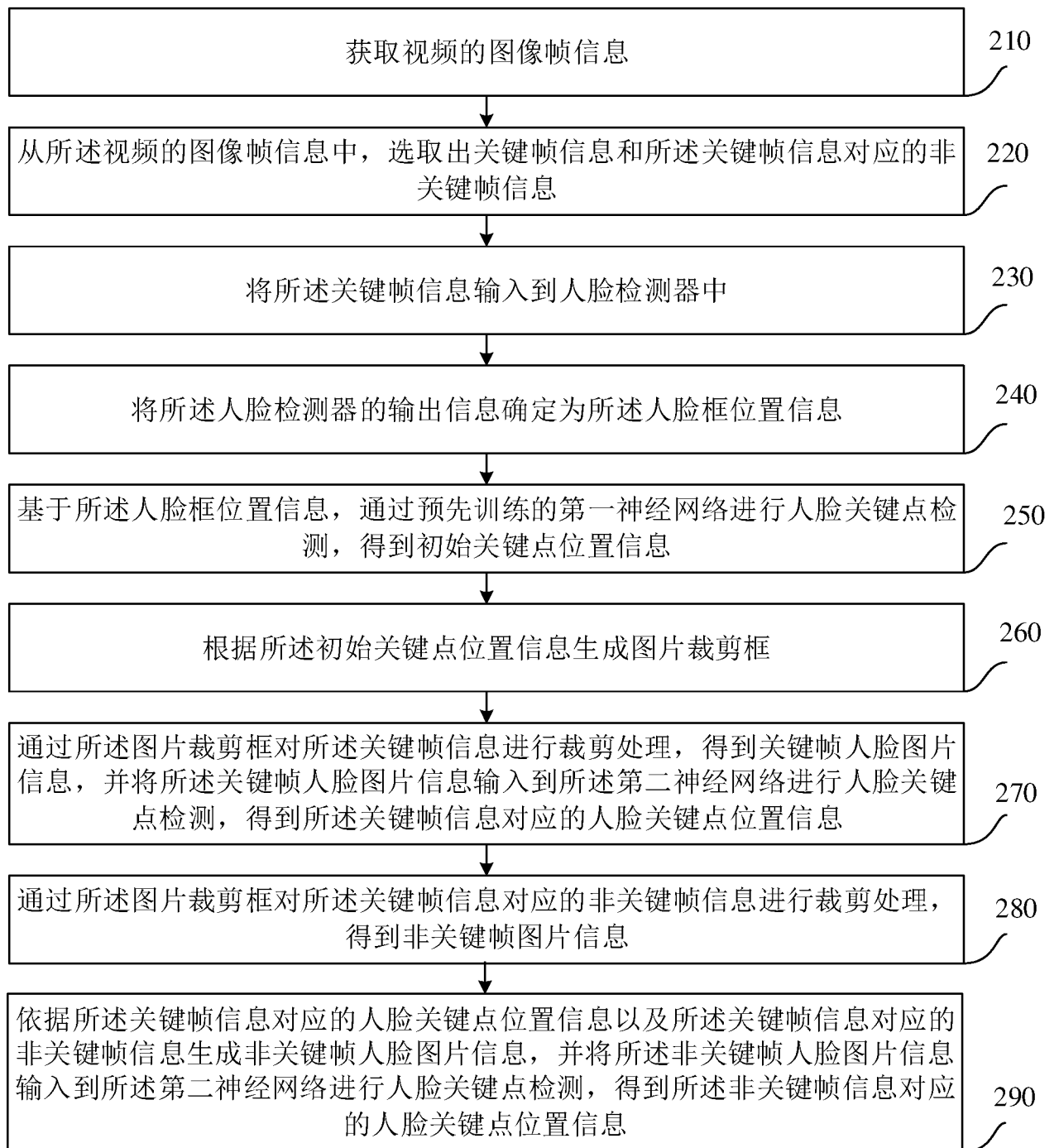


图 2

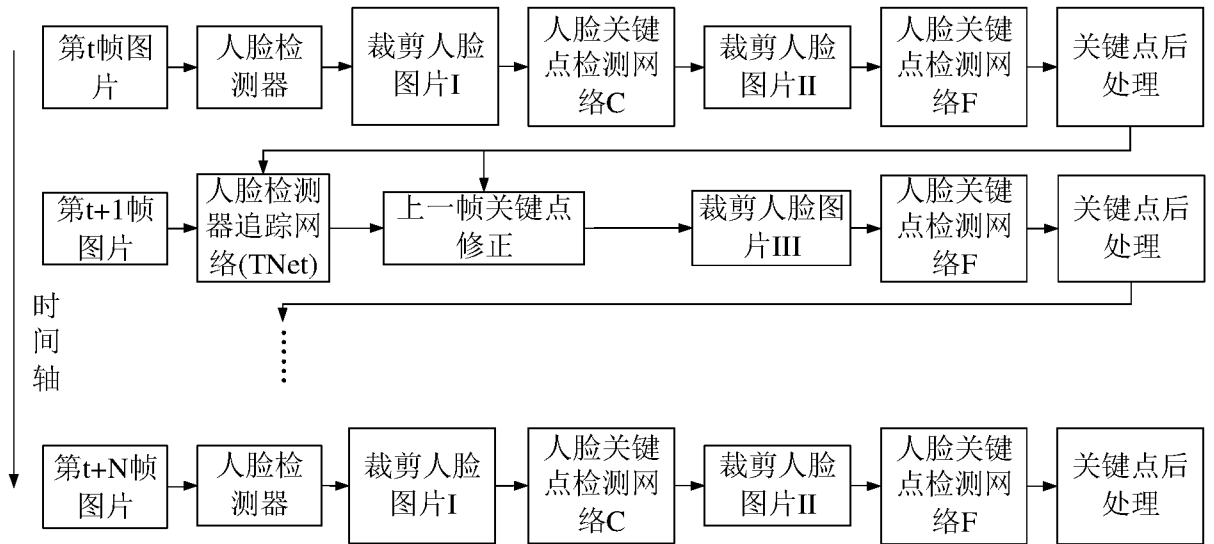


图 3

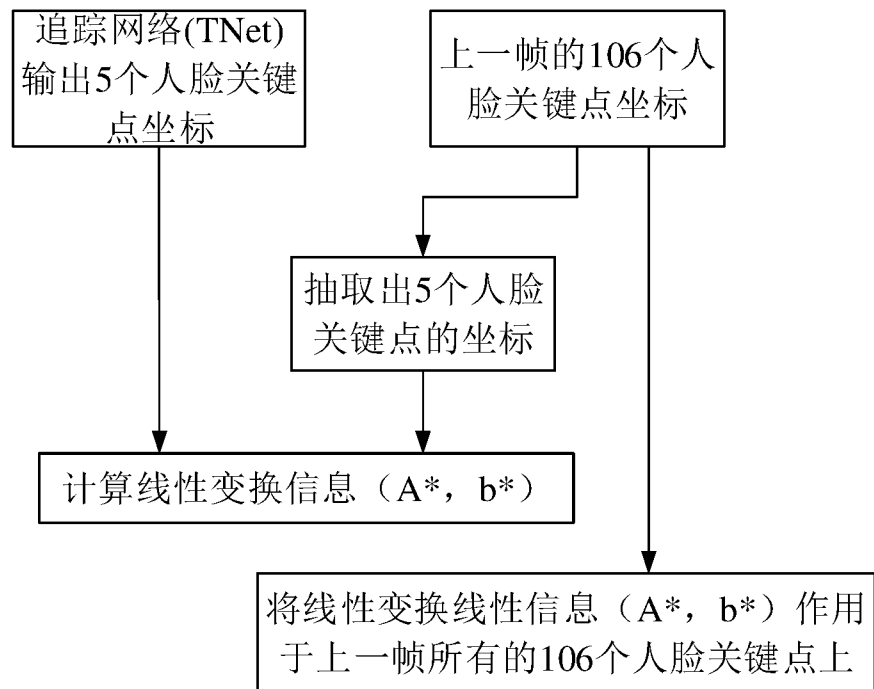


图 4

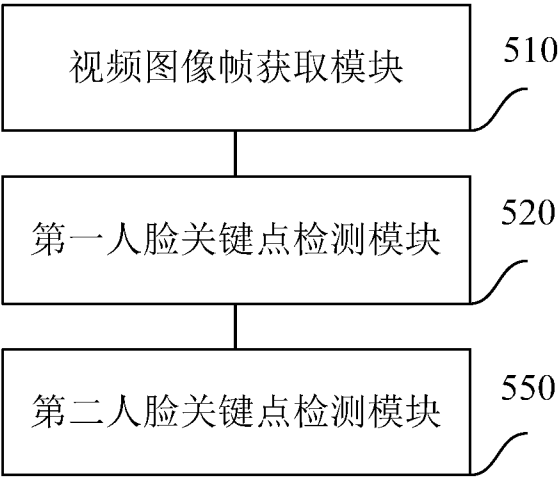


图 5

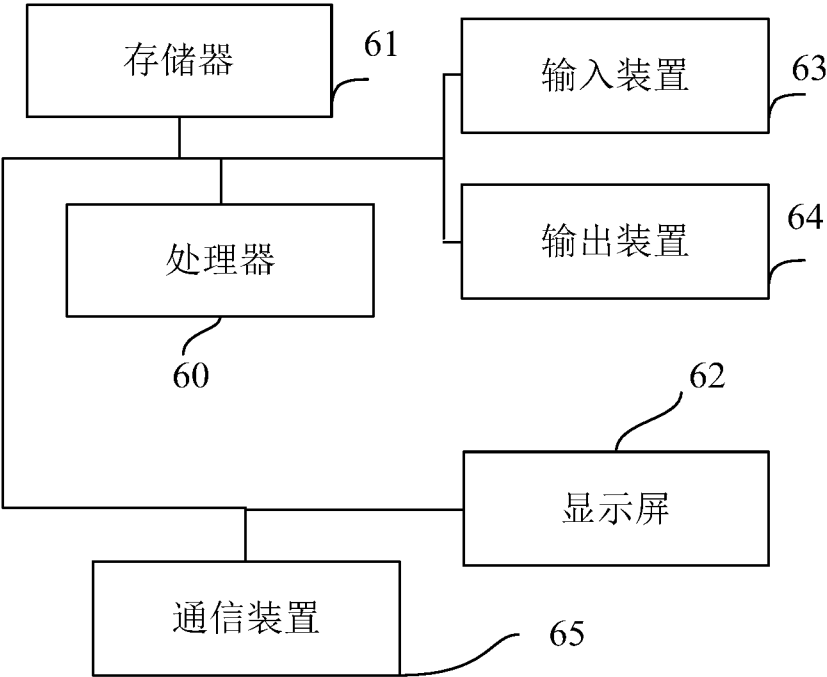


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/081262

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i; G06N 3/04(2006.01)n

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)

G06K; G06N

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; CNKI; DWPI; SIPOABS; WOTXT; USTXT; EPTXT: 广州市百果园信息技术有限公司, 人脸, 关键点, 关键帧, 特征点, 非关键帧, 非关键点, 检测, 监测, 探测, 侦测, 识别, 图像, landmark, face, facial, key point, feature point, image, detect+, recognit+, monitor

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 109657583 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 19 April 2019 (2019-04-19) description, paragraphs [0037]-[0186], figures 2, 3, 6, 7, 10	1-11
A	CN 109376684 A (GUANGZHOU BIGO INFORMATION TECHNOLOGY CO., LTD.) 22 February 2019 (2019-02-22) entire document	1-11
A	CN 109063679 A (DUOYI NETWORK CO., LTD.) 21 December 2018 (2018-12-21) entire document	1-11



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

“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

08 May 2020

Date of mailing of the international search report

22 June 2020

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/081262

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	109657583	A	19 April 2019	None	
CN	109376684	A	22 February 2019	None	
CN	109063679	A	21 December 2018	None	

国际检索报告

国际申请号

PCT/CN2020/081262

A. 主题的分类

G06K 9/00(2006.01)i; G06N 3/04(2006.01)n

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)

G06K; G06N

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; CNKI; DWPI; SIPOABS; WOTXT; USTXT; EPTXT: 广州市百果园信息技术有限公司, 人脸, 关键点, 关键帧, 特征点, 非关键帧, 非关键点, 检测, 监测, 探测, 侦测, 识别, 图像, landmark, face, facial, key point, feature point, image, detect+, recognit+, monitor

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
X	CN 109657583 A (腾讯科技深圳有限公司) 2019年 4月 19日 (2019 - 04 - 19) 说明书第[0037]-[0186]段, 附图2、3、6、7、10	1-11
A	CN 109376684 A (广州市百果园信息技术有限公司) 2019年 2月 22日 (2019 - 02 - 22) 全文	1-11
A	CN 109063679 A (广州多益网络股份有限公司 等) 2018年 12月 21日 (2018 - 12 - 21) 全文	1-11

☐ 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

“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

2020年 5月 8日

国际检索报告邮寄日期

2020年 6月 22日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

胡彦玲

电话号码 (86-512) 88995664

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/081262

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	109657583	A	2019年 4月 19日	无	
CN	109376684	A	2019年 2月 22日	无	
CN	109063679	A	2018年 12月 21日	无	