# (19) 国家知识产权局



# (12) 发明专利



(10) 授权公告号 CN 112651880 B (45) 授权公告日 2022. 12. 30

审查员 程小梅

(21) 申请号 202011562381.4

(22) 申请日 2020.12.25

- (65) 同一申请的已公布的文献号 申请公布号 CN 112651880 A
- (43) 申请公布日 2021.04.13
- (73) **专利权人** 北京市商汤科技开发有限公司 地址 100080 北京市海淀区北四环西路58 号11层1101-1117室
- (72) 发明人 单钰皓 张宁 王权
- (74) 专利代理机构 北京林达刘知识产权代理事务所(普通合伙) 11277 专利代理师 刘新宇

(51) Int.CI.

**G06T 3/00** (2006.01)

GO6N 3/08 (2006.01)

权利要求书2页 说明书14页 附图4页

S11

S12

## (54) 发明名称

视频数据处理方法及装置、电子设备和存储介质

#### (57) 摘要

本公开涉及一种视频数据处理方法及装置、电子设备和存储介质,所述方法包括:获取待处理视频数据;通过图像处理网络对所述待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据;其中,在所述图像处理网络的训练过程中,通过确定相同语义的像素点在不同样本图像中的不同位置处,采用所述图像处理网络进行图像处理后所产生的差异,训练所述图像处理网络。本公开实施例可提高图像处理后的视频数据的稳定性。

获取待处理视频数据

通过图像处理网络对所述待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据;其中,在所述图像处理网络的训练过程中,通过确定相同语义的像素点在不同样本图像中的不同位置处,采用所述图像处理网络进行图像处理后所产生的差异,训练所述图像处理网络

CN 112651880 B

1.一种视频数据处理方法,其特征在于,所述方法包括:

获取待处理视频数据:

通过图像处理网络对所述待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据:

所述方法还包括:

对第一样本图像进行变换处理,得到第二样本图像,其中,所述第二样本图像与所述第一样本图像中具有相同语义的像素点,在所述第二样本图像及所述第一样本图像中的位置不同:

通过图像处理网络对所述第一样本图像及第二样本图像进行图像处理,得到图像处理 后的第一样本图像及图像处理后的第二样本图像;

对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像;

通过所述逆变换图像及所述图像处理后的第一样本图像之间的差异,以及所述图像处理后的第一样本图像和所述图像处理后的第二样本图像中具有相同语义的像素点的差异确定的稳定损失,训练所述图像处理网络:

在所述变换处理为随机仿射变换的情况下,所述对第一样本图像进行变换处理,得到第二样本图像,包括:

根据仿射变换矩阵对所述第一样本图像进行变换处理,得到所述第二样本图像;

所述对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像,包括:

根据逆仿射变换矩阵对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像,其中,所述仿射变换矩阵与所述逆仿射变换矩阵的乘积为1。

- 2.根据权利要求1所述的方法,其特征在于,所述图像处理网络为风格化网络,所述风格化网络用于对待处理视频数据进行风格化处理,得到风格化视频数据,所述风格化视频数据的风格与所述风格化网络对应的风格相同。
- 3.根据权利要求1所述的方法,其特征在于,所述通过所述逆变换图像及所述图像处理 后的第一样本图像之间的差异,训练所述图像处理网络,包括:

通过所述第一样本图像及所述逆变换图像得到蒙版图像;

根据所述图像处理后的第一样本图像、所述逆变换图像及所述蒙版图像,确定所述图像处理网络的稳定损失,其中,所述稳定损失包括所述逆变换图像及所述图像处理后的第一样本图像之间的差异;

根据所述图像处理网络的稳定损失,调整所述图像处理网络的网络参数。

4.根据权利要求3所述的方法,其特征在于,所述根据所述图像处理后的第一样本图像、所述逆变换图像及所述蒙版图像,确定所述图像处理网络的稳定损失,包括:

根据所述图像处理后的第一样本图像与所述蒙版图像,得到第一融合图像,及根据所述逆变换图像与所述蒙版图像,得到第二融合图像;

根据所述第一融合图像与所述第二融合图像,确定所述图像处理网络的稳定损失。

- 5.根据权利要求1至4中任一项所述的方法,其特征在于,所述变换处理包括随机仿射变换、非均匀采样操作、深度信息透视变化操作中的任一项,所述随机仿射变换包括平移、缩放、翻转、旋转、剪切中的至少一项操作。
  - 6.一种视频数据处理装置,其特征在于,包括:

获取模块,用于获取待处理视频数据:

第一处理模块,用于通过图像处理网络对所述待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据;

所述装置还包括:

第二处理模块,用于对第一样本图像进行变换处理,得到第二样本图像,其中,所述第二样本图像与所述第一样本图像中具有相同语义的像素点,在所述第二样本图像及所述第一样本图像中的位置不同;

第三处理模块,用于通过图像处理网络对所述第一样本图像及第二样本图像进行图像 处理,得到图像处理后的第一样本图像及图像处理后的第二样本图像;

逆变换处理模块,用于对所述图像处理后的第二样本图像进行逆变换处理,得到逆变 换图像:

训练模块,用于通过所述逆变换图像及所述图像处理后的第一样本图像之间的差异,以及所述图像处理后的第一样本图像和所述图像处理后的第二样本图像中具有相同语义的像素点的差异确定的稳定损失,训练所述图像处理网络:

在所述变换处理为随机仿射变换的情况下,所述第二处理模块,还用于:

根据仿射变换矩阵对所述第一样本图像进行变换处理,得到所述第二样本图像;

所述逆变换处理模块,还用于:

根据逆仿射变换矩阵对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换 图像,其中,所述仿射变换矩阵与所述逆仿射变换矩阵的乘积为1。

7.一种电子设备,其特征在于,包括:

处理器:

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为调用所述存储器存储的指令,以执行权利要求1至5中任意一项所述的方法。

8.一种计算机可读存储介质,其上存储有计算机程序指令,其特征在于,所述计算机程序指令被处理器执行时实现权利要求1至5中任意一项所述的方法。

# 视频数据处理方法及装置、电子设备和存储介质

#### 技术领域

[0001] 本公开涉及计算机技术领域,尤其涉及一种视频数据处理方法及装置、电子设备和存储介质。

### 背景技术

[0002] 随着科技的发展和进步,对多媒体数据的处理方式越来越多。例如:可以对多媒体数据进行美化、添加特效、添加滤镜等处理,对多媒体数据处理处理方式的丰富性以及处理质量,有着越来越高的需求。

#### 发明内容

[0003] 本公开提出了一种用于对视频数据进行图像处理的技术方案。

[0004] 根据本公开的一方面,提供了一种视频数据处理方法,所述方法包括:

[0005] 获取待处理视频数据;通过图像处理网络对所述待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据;其中,在所述图像处理网络的训练过程中,通过确定相同语义的像素点在不同样本图像中的不同位置处,采用所述图像处理网络进行图像处理后所产生的差异,训练所述图像处理网络。

[0006] 本公开实施例提供的视频数据处理方法,由于采用的图像处理网络在训练过程中考虑了具有相同语义的像素点在不同位置,采用所述图像处理网络进行图像处理后所产生的差异,故而训练得到的图像处理网络对待处理视频数据中的每一视频帧进行图像处理后,得到的图像处理后的视频数据的稳定性较高,可以缓解图像处理后的视频数据的闪烁问题。

[0007] 在一种可能的实现方式中,所述图像处理网络为风格化网络,所述风格化网络用于对待处理视频数据进行风格化处理,得到风格化视频数据,所述风格化视频数据的风格与所述风格化网络对应的风格相同。

[0008] 本公开实施例提供的视频数据处理方法,可以缓解风格化网络处理得到的风格化视频数据的闪烁问题,提高风格化视频数据的稳定性。

[0009] 在一种可能的实现方式中,所述方法还包括:对第一样本图像进行变换处理,得到第二样本图像,其中,所述第二样本图像与所述第一样本图像中具有相同语义的像素点,在所述第二样本图像及所述第一样本图像中的位置不同;通过所述图像处理网络对所述第一样本图像及第二样本图像进行图像处理,得到图像处理后的第一样本图像及图像处理后的第二样本图像;对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像;通过所述逆变换图像及所述图像处理后的第一样本图像之间的差异,训练所述图像处理网络。

[0010] 本公开实施例提供的视频数据处理方法,通过对第一样本图像进行变换处理,可以得到第二样本图像,进而通过计算第一样本图像及第二样本图像进行图像处理网络进行处理后,相同语义的像素点进行图像处理后的差异,训练图像处理网络,可以得到稳定性较高的图像处理网络。

[0011] 在一种可能的实现方式中,所述通过所述逆变换图像及所述图像处理后的第一样本图像之间的差异,训练所述图像处理网络,包括:通过所述第一样本图像及所述逆变换图像得到蒙版图像;根据所述图像处理后的第一样本图像、所述逆变换图像及所述蒙版图像,确定所述图像处理网络的稳定损失,其中,所述稳定损失包括所述逆变换图像及所述图像处理后的第一样本图像之间的差异;根据所述图像处理网络的稳定损失,调整所述图像处理网络的网络参数。

[0012] 本公开实施例提供的视频数据处理方法,可以通过蒙版图像确定逆变换图像与图像处理后的第一样本图像之间具有相同语义的像素点,进而可以通过计算这些像素点之间的差异,精准的得到图像处理网络的稳定损失。

[0013] 在一种可能的实现方式中,所述根据所述图像处理后的第一样本图像、所述逆变换图像及所述蒙版图像,确定所述图像处理网络的稳定损失,包括:根据所述图像处理后的第一样本图像与所述蒙版图像,得到第一融合图像,及根据所述逆变换图像与所述蒙版图像,得到第二融合图像;根据所述第一融合图像与所述第二融合图像,确定所述图像处理网络的稳定损失。

[0014] 本公开实施例提供的视频数据处理方法,可以通过蒙版图像确定逆变换图像与图像处理后的第一样本图像之间具有相同语义的像素点,进而可以通过计算这些像素点之间的差异,精准的得到图像处理网络的稳定损失。

[0015] 在一种可能的实现方式中,所述变换处理包括随机仿射变换、非均匀采样操作、深度信息透视变化操作中的任一项,所述随机仿射变换包括平移、缩放、翻转、旋转、剪切中的至少一项操作。

[0016] 本公开实施例提供的视频数据处理方法,通过对第一样本图像进行变换处理,可以得到第二样本图像,进而通过计算第一样本图像及第二样本图像进行图像处理网络进行处理后,相同语义的像素点进行图像处理后的差异,训练图像处理网络,可以得到稳定性较高的图像处理网络。

[0017] 在一种可能的实现方式中,在所述变换处理为随机仿射变换的情况下,所述对第一样本图像进行变换处理,得到第二样本图像,包括:

[0018] 根据仿射变换矩阵对所述第一样本图像进行变换处理,得到所述第二样本图像:

[0019] 所述对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像,包括:

[0020] 根据逆仿射变换矩阵对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像,其中,所述仿射变换矩阵与所述逆仿射变换矩阵的乘积为1。

[0021] 本公开实施例提供的视频数据处理方法,通过能够完全恢复图像的仿射变换,可以给出最大程度的损失计算域,并较大程度的减少计算时间,提高了图像处理网络的训练速度。

[0022] 根据本公开的一方面,提供了一种视频数据处理装置,包括:

[0023] 获取模块,用于获取待处理视频数据;

[0024] 第一处理模块,用于通过图像处理网络对所述待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据:

[0025] 其中,在所述图像处理网络的训练过程中,通过确定相同语义的像素点在不同样本图像中的不同位置处,采用所述图像处理网络进行图像处理后所产生的差异,训练所述

图像处理网络。

[0026] 在一种可能的实现方式中,所述图像处理网络为风格化网络,所述风格化网络用于对待处理视频数据进行风格化处理,得到风格化视频数据,所述风格化视频数据的风格与所述风格化网络对应的风格相同。

[0027] 在一种可能的实现方式中,所述装置还包括:

[0028] 第二处理模块,用于对所述第一样本图像进行变换处理,得到第二样本图像,其中,所述第二样本图像与所述第一样本图像中具有相同语义的像素点,在所述第二样本图像及所述第一样本图像中的位置不同:

[0029] 第三处理模块,用于通过所述图像处理网络对所述第一样本图像及第二样本图像 进行图像处理,得到图像处理后的第一样本图像及图像处理后的第二样本图像;

[0030] 逆变换处理模块,用于对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像:

[0031] 训练模块,用于通过所述逆变换图像及所述图像处理后的第一样本图像之间的差异,训练所述图像处理网络。

[0032] 在一种可能的实现方式中,所述训练模块,还用于:

[0033] 通过所述第一样本图像及所述逆变换图像得到蒙版图像;

[0034] 根据所述图像处理后的第一样本图像、所述逆变换图像及所述蒙版图像,确定所述图像处理网络的稳定损失,其中,所述稳定损失包括所述逆变换图像及所述图像处理后的第一样本图像之间的差异:

[0035] 根据所述图像处理网络的稳定损失,调整所述图像处理网络的网络参数。

[0036] 在一种可能的实现方式中,所述训练模块,还用于:

[0037] 根据所述图像处理后的第一样本图像与所述蒙版图像,得到第一融合图像,及根据所述逆变换图像与所述蒙版图像,得到第二融合图像;

[0038] 根据所述第一融合图像与所述第二融合图像,确定所述图像处理网络的稳定损失。

[0039] 在一种可能的实现方式中,所述变换处理包括随机仿射变换、非均匀采样操作、深度信息透视变化操作中的任一项,所述随机仿射变换包括平移、缩放、翻转、旋转、剪切中的至少一项操作。

[0040] 在一种可能的实现方式中,在所述变换处理为随机仿射变换的情况下,所述第二处理模块,还用于:

[0041] 根据仿射变换矩阵对所述第一样本图像进行变换处理,得到所述第二样本图像;

[0042] 所述逆变换处理模块,还用干:

[0043] 根据逆仿射变换矩阵对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像,其中,所述仿射变换矩阵与所述逆仿射变换矩阵的乘积为1。根据本公开的一方面,提供了一种电子设备,包括:处理器;用于存储处理器可执行指令的存储器;其中,所述处理器被配置为调用所述存储器存储的指令,以执行上述方法。

[0044] 根据本公开的一方面,提供了一种计算机可读存储介质,其上存储有计算机程序指令,所述计算机程序指令被处理器执行时实现上述方法。

[0045] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,而非

限制本公开。根据下面参考附图对示例性实施例的详细说明,本公开的其它特征及方面将变得清楚。

#### 附图说明

[0046] 此处的附图被并入说明书中并构成本说明书的一部分,这些附图示出了符合本公开的实施例,并与说明书一起用于说明本公开的技术方案。

[0047] 图1示出根据本公开实施例的视频数据处理方法的流程图:

[0048] 图2示出根据本公开实施例的视频数据处理方法的示意图;

[0049] 图3示出根据本公开实施例的视频数据处理方法的示意图;

[0050] 图4示出根据本公开实施例的视频数据处理装置的框图;

[0051] 图5示出根据本公开实施例的一种电子设备800的框图;

[0052] 图6示出根据本公开实施例的一种电子设备1900的框图。

#### 具体实施方式

[0053] 以下将参考附图详细说明本公开的各种示例性实施例、特征和方面。附图中相同的附图标记表示功能相同或相似的元件。尽管在附图中示出了实施例的各种方面,但是除非特别指出,不必按比例绘制附图。

[0054] 在这里专用的词"示例性"意为"用作例子、实施例或说明性"。这里作为"示例性" 所说明的任何实施例不必解释为优于或好于其它实施例。

[0055] 本文中术语"和/或",仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。另外,本文中术语"至少一种"表示多种中的任意一种或多种中的至少两种的任意组合,例如,包括A、B、C中的至少一种,可以表示包括从A、B和C构成的集合中选择的任意一个或多个元素。

[0056] 另外,为了更好地说明本公开,在下文的具体实施方式中给出了众多的具体细节。本领域技术人员应当理解,没有某些具体细节,本公开同样可以实施。在一些实例中,对于本领域技术人员熟知的方法、手段、元件和电路未作详细描述,以便于凸显本公开的主旨。

[0057] 风格化滤镜是视频处理过程中高频使用的功能,它通过端到端的处理对待处理图像进行特定的风格变化,例如:将待处理图像或者待处理图像中的目标人物处理成油画、漫画、素描等风格。视频数据的风格化处理对于帧间的稳定性要求较高,稳定性低的情况下,风格化后的视频数据会发生闪烁等现象,降低视觉效果及用户体验。

[0058] 为了提高风格化后的视频数据帧间的稳定性,本公开实施例提供了一种视频数据处理方法。

[0059] 图1示出根据本公开实施例的视频数据处理方法的流程图,所述视频数据处理方法可以由终端设备或服务器等电子设备执行,终端设备可以为用户设备(User Equipment, UE)、移动设备、用户终端、终端、蜂窝电话、无绳电话、个人数字助理(Personal Digital Assistant,PDA)、手持设备、计算设备、车载设备、可穿戴设备等,所述方法可以通过处理器调用存储器中存储的计算机可读指令的方式来实现。或者,可通过服务器执行所述方法。如图1所示,所述视频数据处理方法可以包括:

[0060] 在步骤S11中,获取待处理视频数据;

[0061] 在步骤S12中,通过图像处理网络对所述待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据;其中,在所述图像处理网络的训练过程中,通过确定相同语义的像素点在不同样本图像中的不同位置处,采用所述图像处理网络进行图像处理后所产生的差异,训练所述图像处理网络。

[0062] 在本公开实施例中,电子设备可以通过采集视频数据、上传视频数据或者下载视频数据等方式中的任一种方式获取待处理视频数据。在获取待处理视频数据后,可以根据相应的图像处理网络对待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据。其中,图像处理网络可以为对待处理视频数据进行美化、添加特效、添加滤镜、风格化等处理的神经网络。

[0063] 实际上,本公开实施例提供的图像处理网络也可以对图像进行图像处理,本公开实施例对此不作限定。在本公开实施例中,图像处理网络可例如为卷积神经网络,本公开对该图像处理网络的具体网络结构不作限制。

[0064] 在图像处理网络的训练过程中,不同的两个样本图像中具有相同语义的像素点在该两个样本图像中位置不同,则可以通过计算该两个不同样本图像中具有相同语义的像素点,经过图像处理网络进行图像处理后所产生的差异,训练图像处理网络,直至不同样本图像中相同语义的像素点,在不同位置经过图像处理网络进行图像处理后所产生的差异满足训练要求时,完成训练,得到训练后的图像处理网络。

[0065] 举例来说,在训练图像处理网络的过程中,第一样本图像与第二样本图像中存在 具有相同语义的像素点,该相同语义的像素点在第一样本图像及第二样本图像中位置不 同,采用所述图像处理网络分别对第一样本图像及第二样本图像进行图像处理后,确定该 具有相同语义的像素点在经图像处理后所产生的差异作为图像处理网络的稳定损失,并将 稳定损失作为图像处理网络的网络损失中的一部分,进而通过网络损失训练图像处理网 络。或者,可以将图像处理网络的训练过程分为两部分,在第一部分训练过程中,可以预训 练图像处理网络,在第二部分训练过程中,在完成图像处理网络的预训练后,可以进一步计 算图像处理网络的稳定损失,进而根据图像处理网络的稳定损失调整图像处理网络的网络 参数。

[0066] 这样一来,根据本公开实施例提供的视频数据处理方法,由于采用的图像处理网络在训练过程中考虑了具有相同语义的像素点在不同位置,采用所述图像处理网络进行图像处理后所产生的差异,故而训练得到的图像处理网络对待处理视频数据中的每一视频帧进行图像处理后,得到的图像处理后的视频数据的稳定性较高,可以缓解图像处理后的视频数据的闪烁问题。

[0067] 在一种可能的实现方式中,所述图像处理网络可以为风格化网络,所述风格化网络用于对待处理视频数据进行风格化处理,得到风格化视频数据,所述风格化视频数据的风格与所述风格化网络对应的风格相同。

[0068] 例如:用户可以选定一种风格(例如油画风格),电子设备可以采用该风格对应的风格化网络对待处理视频数据进行风格化处理,得到风格化视频数据,该风格化视频数据可以表现为油画风格。这样,可以缓解风格化网络处理得到的风格化视频数据的闪烁问题,提高风格化视频数据的稳定性。

[0069] 在一种可能的实现过程中,所述方法还可以包括:

[0070] 对所述第一样本图像进行变换处理,得到第二样本图像,其中,所述第二样本图像与所述第一样本图像中具有相同语义的像素点,在所述第二样本图像及所述第一样本图像中的位置不同;

[0071] 通过所述图像处理网络对所述第一样本图像及第二样本图像进行图像处理,得到图像处理后的第一样本图像及图像处理后的第二样本图像:

[0072] 对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像;

[0073] 通过所述逆变换图像及所述图像处理后的第一样本图像之间的差异,训练所述图像处理网络。

[0074] 举例来说,可以通过训练集训练图像处理网络,训练集中可以包括多个第一样本图像。在图像处理网络的训练过程中,针对任意第一样本图像,可以对第一样本图像进行变换处理,该变换处理可以对第一样本图像中的像素点的位置进行改变的同时,不改变该像素点的语义,从而得到第二样本图像,也即第二样本图像与第一样本图像中具有相同语义的像素点,在第二样本图像和第一样本图像中的位置不同。

[0075] 示例性的,以第一样本图像中的任一像素点p为例,其在第一样本图像中的位置为 X,在进行变换处理后,其在第二样本图像中的位置为T,也即第一样本图像中位置X的相素 点p的语义与第二样本图像中的位置T的相素点p的语义相同。针对第一样本图像中的其他 像素点以此类推进行变换,可以得到第二样本图像,本公开实施例对此过程不再赘述。

[0076] 在一种可能的实现过程中,所述变换处理可以包括随机仿射变换、非均匀采样操作、深度信息透视变化操作中的任一项,所述随机仿射变换可以包括平移、缩放、翻转、旋转、剪切中的至少一项操作。

[0077] 可以将第一样本图像及第二样本图像分别输入图像处理网络中进行图像处理,得到图像处理后的第一样本图像及图像处理后的第二样本图像。并对图像处理后的第二样本图像进行逆变化处理,得到逆变换图像。示例性的,以第二样本图像中位于位置T的像素点p为例,第二样本图像进行图像处理后,在得到的图像处理后的第二样本图像中位置T的像素点为p1,对图像处理后的第二样本图像进行逆变换处理后,在得到的逆变换图像中像素点p1对应的位置为X,针对其他像素以此类推进行逆变换,以得到逆变换图像,本公开实施例对此过程不再赘述。

[0078] 这样一来,本公开实施例通过对第一样本图像进行变换处理,可以得到第二样本图像,进而通过计算第一样本图像及第二样本图像进行图像处理网络进行处理后,相同语义的像素点进行图像处理后的差异,训练图像处理网络,可以得到稳定性较高的图像处理网络。

[0079] 在一种可能的实现过程中,在所述变换处理为随机仿射变换的情况下,所述对所述第一样本图像进行变换处理,得到所述第二样本图像,可以包括:

[0080] 根据仿射变换矩阵对所述第一样本图像进行变换处理,得到所述第二样本图像:

[0081] 本公开实施例中,所述对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像,可以包括:

[0082] 根据逆仿射变换矩阵对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像,其中,所述仿射变换矩阵与所述逆仿射变换矩阵的乘积为1。

[0083] 举例来说,可以根据欲对第一样本图像进行的变换处理,确定对应的仿射变换矩

阵。例如:如图2所示,对第一样本图像进行旋转,平移及缩放,其对应的仿射变换矩阵

$$\mathbf{M}$$
=[ $A$   $B$ ]= $\begin{bmatrix} a_{00} & a_{01} & b_{00} \\ a_{10} & a_{11} & b_{10} \end{bmatrix}_{2\times 3}$ ,其中, $\mathbf{A}$ = $\begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}_{2\times 2}$ , $\mathbf{B}$ = $\begin{bmatrix} b_{00} \\ b_{10} \end{bmatrix}_{2\times 1}$ ,其中矩阵A控制对第一

样本图像进行旋转和缩放,矩阵B控制对第一样本图像进行平移。假设像素点p的位置表示为 $\mathbf{X}=\left[egin{array}{c} x \\ y \end{array}\right]$ ,通过仿射变换矩阵对该像素点N进行变换处理后,得到像素点p变换后的位置

$$\mathbf{T} = \mathbf{M} \times [x, y, 1]^T = \begin{bmatrix} a_{00}x + a_{01}y + b_{00} \\ a_{10}x + a_{11}y + b_{10} \end{bmatrix}.$$

[0084] 实际上,在仿射变换过程中,为了避免第一样本图像变换过小无法区别原图,变换过大降低计算损失的计算空间,需要对各变换进行超参设置。例如:以上述示例为例,对各项变换的超参设置可以为,设置平移变换的平移因子Translation=random\_float(0,01),旋转角度Rotation=random\_int(-2,2),缩放因子Scale=random\_float(1,1.01),其中,random表示括号范围内随机取值,int表示整数,float表示浮点数。

[0085] 仿射变换矩阵与逆仿射变换矩阵的乘积为1,故根据仿射变换矩阵可以得到对应的逆仿射变换矩阵。得到逆仿射变换矩阵后,可以根据逆仿射变换矩阵对图像处理后的第二样本图像进行逆变换处理,得到逆变换图像。示例性的,像素点p由位置X变换到位置T后,进行图像处理,得到图像处理后的像素点p1,将图像处理后的像素点p1由位置T逆变换到位置X,进而得到逆变换图像。

[0086] 这样一来,本公开实施例通过能够完全恢复图像的仿射变换,可以给出最大程度的损失计算域,并较大程度的减少计算时间,提高了图像处理网络的训练速度。

[0087] 在得到逆变换图像后,可以确定逆变换图像与图像处理后的第一样本图像之间的差异,进而根据逆变换图像与图像处理后的第一样本图像之间的差异训练图像处理网络,直至逆变换图像与图像处理后的第一样本图像之间的差异满足训练要求,得到训练后的图像处理网络。在一种可能的实现过程中,所述通过所述逆变换图像及所述图像处理后的第一样本图像之间的差异,训练所述图像处理网络,可以包括:

[0088] 通过所述第一样本图像及所述逆变换图像得到蒙版图像:

[0089] 根据所述图像处理后的第一样本图像、所述逆变换图像及所述蒙版图像,确定所述图像处理网络的稳定损失,其中,所述稳定损失包括所述逆变换图像及所述图像处理后的第一样本图像之间的差异;

[0090] 根据所述图像处理网络的稳定损失,调整所述图像处理网络的网络参数。

[0091] 举例来说,比对第一样本图像与逆变换图像,确定未进行图像处理的像素点的位置,在蒙版图像中该像素点的位置对应的像素值可以设置为0,对于进行图像处理的像素点,在蒙版图像中可以将该像素点的位置对应的像素值设置为1,进而得到蒙版图像。

[0092] 进而通过蒙板图像,确定图像处理后的第一样本图像与逆变换图像中相同位置进行过图像处理后的像素点,进而根据这些图像处理后的像素点的差异确定图像处理网络的稳定损失。也即,稳定损失可以包括所述稳定损失用于表征所述逆变换图像及所述图像处理后的第一样本图像之间的差异。稳定损失可以用于表征图像处理网络得到的图像处理后的视频数据的视频帧的稳定性,并可以根据稳定损失调整图像处理网络的网络参数,直至图像处理网络的稳定损失满足精度要求。

[0093] 这样一来,本公开实施例可以通过蒙版图像确定逆变换图像与图像处理后的第一样本图像之间具有相同语义的像素点,进而可以通过计算这些像素点之间的差异,可以精准的得到图像处理网络的稳定损失。

[0094] 在一种可能的实现过程中,所述根据所述图像处理后的第一样本图像、所述逆变换图像及所述蒙版图像,确定所述图像处理网络的稳定损失,可以包括:

[0095] 根据所述图像处理后的第一样本图像与所述蒙版图像,得到第一融合图像,及根据所述逆变换图像与所述蒙版图像,得到第二融合图像;

[0096] 根据所述第一融合图像与所述第二融合图像,确定所述图像处理网络的稳定损失。

[0097] 举例来说,可以将图像处理后的第一样本图像与蒙版图像相乘,得到第一融合图像,将逆变换图像与蒙版图像相乘,得到第二融合图像。根据第一融合图像及第二融合图像进行差异计算,得到图像处理网络的稳定损失。例如:可以确定第一融合图像及第二融合图像的均方误差,参照公式(一),将该均方误差作为图像处理网络的稳定损失。

[0100] 其中,E可用于标识错误率,D用于标识方差, $\theta$ 用于标识第一样本图像和第二样本图像, $\hat{\theta}$ 用于标识第一融合图像和第二融合图像。

[0101] 需要说明的是,采用均方误差确定稳定损失仅作为本公开实施例中的一种实现方式,实际上,可以计算差异的方式均可以,例如:1范式、2范式等方式,本公开实施例对此不作限定。

[0102] 为使本领域技术人员更好的理解本公开实施例,以下通过具体示例对本公开实施例加以说明。

[0103] 示例性的,参照图3。对第一样本图像进行缩放和旋转等仿射变换处理后,得到第二样本图像。通过图像处理网络分别对第一样本图像及第二样本图像进行图像处理后,可以得到图像处理后的第一样本图像和图像处理后的第二样本图像。进一步的,对图像处理后的第二样本图像进行逆变换后,可以得到逆变换图像。根据逆变换图像及第一样本图像可以得到蒙版图像,进而通过图像处理后的第一样本图像及蒙版图像进行相乘,得到第一融合图像,通过逆变换图像与蒙版图像进行相乘,得到第二融合图像。通过计算第一融合图像及第二融合图像之间的差异,可以得到图像处理网络的稳定损失,进而可以根据稳定损失训练图像处理网络。

[0104] 这样一来,本公开实施例提供的视频处理方法中涉及的图像处理网络,在训练过程中考虑了稳定损失,使得训练得到的图像处理网络稳定性较高,通过该图像处理网络处理得到的图像处理视频数据稳定性较高,缓解了视频闪烁的问题,并且由于是在图像处理网络的训练过程中实现了稳定性,故在后续图像处理处理过程不会产生任何其余的时空负载。

[0105] 本公开实施例提供的图像处理网络可嵌于深度学习网络来提高视频效果,并适用

于不同学习网络。

[0106] 可以理解,本公开提及的上述各个方法实施例,在不违背原理逻辑的情况下,均可以彼此相互结合形成结合后的实施例,限于篇幅,本公开不再赘述。本领域技术人员可以理解,在具体实施方式的上述方法中,各步骤的具体执行顺序应当以其功能和可能的内在逻辑确定。

[0107] 此外,本公开还提供了视频数据处理装置、电子设备、计算机可读存储介质、程序,上述均可用来实现本公开提供的任一种视频数据处理方法,相应技术方案和描述和参见方法部分的相应记载,不再赘述。

[0108] 图4示出根据本公开实施例的视频数据处理装置的框图,如图4所示,所述装置包括:

[0109] 获取模块41,可以用于获取待处理视频数据;

[0110] 第一处理模块42,可以用于通过图像处理网络对所述待处理视频数据中的每一视频帧进行图像处理,得到图像处理后的视频数据;

[0111] 其中,在所述图像处理网络的训练过程中,通过确定相同语义的像素点在不同样本图像中的不同位置处,采用所述图像处理网络进行图像处理后所产生的差异,训练所述图像处理网络。

[0112] 根据本公开实施例提供的视频数据处理装置,由于采用的图像处理网络在训练过程中考虑了具有相同语义的像素点在不同位置,采用所述图像处理网络进行图像处理后所产生的差异,故而训练得到的图像处理网络对待处理视频数据中的每一视频帧进行图像处理后,得到的图像处理后的视频数据的稳定性较高,可以缓解图像处理后的视频数据的闪烁问题。

[0113] 在一种可能的实现方式中,所述图像处理网络可以为风格化网络,所述风格化网络用于对待处理视频数据进行风格化处理,得到风格化视频数据,所述风格化视频数据的风格与所述风格化网络对应的风格相同。

[0114] 在一种可能的实现方式中,所述装置还可以包括:

[0115] 第二处理模块,可以用于对所述第一样本图像进行变换处理,得到第二样本图像,其中,所述第二样本图像与所述第一样本图像中具有相同语义的像素点,在所述第二样本图像及所述第一样本图像中的位置不同:

[0116] 第三处理模块,可以用于通过所述图像处理网络对所述第一样本图像及第二样本图像进行图像处理,得到图像处理后的第一样本图像及图像处理后的第二样本图像;

[0117] 逆变换处理模块,可以用于对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像;

[0118] 训练模块,可以用于通过所述逆变换图像及所述图像处理后的第一样本图像之间的差异,训练所述图像处理网络。

[0119] 在一种可能的实现方式中,所述训练模块,还可以用于:

[0120] 通过所述第一样本图像及所述逆变换图像得到蒙版图像;

[0121] 根据所述图像处理后的第一样本图像、所述逆变换图像及所述蒙版图像,确定所述图像处理网络的稳定损失,其中,所述稳定损失包括所述逆变换图像及所述图像处理后的第一样本图像之间的差异;

- [0122] 根据所述图像处理网络的稳定损失,调整所述图像处理网络的网络参数。
- [0123] 在一种可能的实现方式中,所述训练模块,还可以用于:
- [0124] 根据所述图像处理后的第一样本图像与所述蒙版图像,得到第一融合图像,及根据所述逆变换图像与所述蒙版图像,得到第二融合图像;
- [0125] 根据所述第一融合图像与所述第二融合图像,确定所述图像处理网络的稳定损失。
- [0126] 在一种可能的实现方式中,所述变换处理包括随机仿射变换、非均匀采样操作、深度信息透视变化操作中的任一项,所述随机仿射变换包括平移、缩放、翻转、旋转、剪切中的至少一项操作。
- [0127] 在一种可能的实现方式中,在所述变换处理为随机仿射变换的情况下,所述第二处理模块,还可以用于:
- [0128] 根据仿射变换矩阵对所述第一样本图像进行变换处理,得到所述第二样本图像;
- [0129] 所述逆变换处理模块,还可以用于:
- [0130] 根据逆仿射变换矩阵对所述图像处理后的第二样本图像进行逆变换处理,得到逆变换图像,其中,所述仿射变换矩阵与所述逆仿射变换矩阵的乘积为1。在一些实施例中,本公开实施例提供的装置具有的功能或包含的模块可以用于执行上文方法实施例描述的方法,其具体实现可以参照上文方法实施例的描述,为了简洁,这里不再赘述。
- [0131] 本公开实施例还提出一种计算机可读存储介质,其上存储有计算机程序指令,所述计算机程序指令被处理器执行时实现上述方法。计算机可读存储介质可以是非易失性计算机可读存储介质。
- [0132] 本公开实施例还提出一种电子设备,包括:处理器;用于存储处理器可执行指令的存储器;其中,所述处理器被配置为调用所述存储器存储的指令,以执行上述方法。
- [0133] 本公开实施例还提供了一种计算机程序产品,包括计算机可读代码,当计算机可读代码在设备上运行时,设备中的处理器执行用于实现如上任一实施例提供的视频数据处理方法的指令。
- [0134] 本公开实施例还提供了另一种计算机程序产品,用于存储计算机可读指令,指令被执行时使得计算机执行上述任一实施例提供的视频数据处理方法的操作。
- [0135] 电子设备可以被提供为终端、服务器或其它形态的设备。
- [0136] 图5示出根据本公开实施例的一种电子设备800的框图。例如,电子设备800可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等终端。
- [0137] 参照图5,电子设备800可以包括以下一个或多个组件:处理组件802,存储器804,电源组件806,多媒体组件808,音频组件810,输入/输出(I/0)的接口812,传感器组件814,以及通信组件816。
- [0138] 处理组件802通常控制电子设备800的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理组件802可以包括一个或多个处理器820来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件802可以包括一个或多个模块,便于处理组件802和其他组件之间的交互。例如,处理组件802可以包括多媒体模块,以方便多媒体组件808和处理组件802之间的交互。

[0139] 存储器804被配置为存储各种类型的数据以支持在电子设备800的操作。这些数据的示例包括用于在电子设备800上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器804可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0140] 电源组件806为电子设备800的各种组件提供电力。电源组件806可以包括电源管理系统,一个或多个电源,及其他与为电子设备800生成、管理和分配电力相关联的组件。

[0141] 多媒体组件808包括在所述电子设备800和用户之间的提供一个输出接口的屏幕。在一些实施例中,屏幕可以包括液晶显示器 (LCD) 和触摸面板 (TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界,而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件808包括一个前置摄像头和/或后置摄像头。当电子设备800处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0142] 音频组件810被配置为输出和/或输入音频信号。例如,音频组件810包括一个麦克风(MIC),当电子设备800处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器804或经由通信组件816发送。在一些实施例中,音频组件810还包括一个扬声器,用于输出音频信号。

[0143] I/0接口812为处理组件802和外围接口模块之间提供接口,上述外围接口模块可以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0144] 传感器组件814包括一个或多个传感器,用于为电子设备800提供各个方面的状态评估。例如,传感器组件814可以检测到电子设备800的打开/关闭状态,组件的相对定位,例如所述组件为电子设备800的显示器和小键盘,传感器组件814还可以检测电子设备800或电子设备800一个组件的位置改变,用户与电子设备800接触的存在或不存在,电子设备800方位或加速/减速和电子设备800的温度变化。传感器组件814可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件814还可以包括光传感器,如互补金属氧化物半导体(CMOS)或电荷耦合装置(CCD)图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件814还可以包括加速度传感器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0145] 通信组件816被配置为便于电子设备800和其他设备之间有线或无线方式的通信。电子设备800可以接入基于通信标准的无线网络,如无线网络(WiFi),第二代移动通信技术(2G)或第三代移动通信技术(3G),或它们的组合。在一个示例性实施例中,通信组件816经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信组件816还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0146] 在示例性实施例中,电子设备800可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述方法。

[0147] 在示例性实施例中,还提供了一种非易失性计算机可读存储介质,例如包括计算机程序指令的存储器804,上述计算机程序指令可由电子设备800的处理器820执行以完成上述方法。

[0148] 图6示出根据本公开实施例的一种电子设备1900的框图。例如,电子设备1900可以被提供为一服务器。参照图6,电子设备1900包括处理组件1922,其进一步包括一个或多个处理器,以及由存储器1932所代表的存储器资源,用于存储可由处理组件1922的执行的指令,例如应用程序。存储器1932中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外,处理组件1922被配置为执行指令,以执行上述方法。

[0149] 电子设备1900还可以包括一个电源组件1926被配置为执行电子设备1900的电源管理,一个有线或无线网络接口1950被配置为将电子设备1900连接到网络,和一个输入输出 (I/0) 接口1958。电子设备1900可以操作基于存储在存储器1932的操作系统,例如微软服务器操作系统 (Windows Server ),苹果公司推出的基于图形用户界面操作系统 (Mac OS  $X^{TM}$ ),多用户多进程的计算机操作系统 (Unix ),自由和开放原代码的类Unix操作系统 (Linux ),开放原代码的类Unix操作系统 (FreeBSD )或类似。

[0150] 在示例性实施例中,还提供了一种非易失性计算机可读存储介质,例如包括计算机程序指令的存储器1932,上述计算机程序指令可由电子设备1900的处理组件1922执行以完成上述方法。

[0151] 本公开可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质,其上载有用于使处理器实现本公开的各个方面的计算机可读程序指令。

[0152] 计算机可读存储介质可以是可以保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是(但不限于)电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、静态随机存取存储器(SRAM)、便携式压缩盘只读存储器(CD-ROM)、数字多功能盘(DVD)、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身,诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波(例如,通过光纤电缆的光脉冲)、或者通过电线传输的电信号。

[0153] 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备,或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令,并转发该计算机可读程序指令,以供存储在各个计算/处理设备中的计算机可读存储介质中。

[0154] 用于执行本公开操作的计算机程序指令可以是汇编指令、指令集架构(ISA)指令、

机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码,所述编程语言包括面向对象的编程语言一诸如Smalltalk、C++等,以及常规的过程式编程语言一诸如"C"语言或类似的编程语言。计算机可读程序指令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一包括局域网(LAN)或广域网(WAN)一连接到用户计算机,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)。在一些实施例中,通过利用计算机可读程序指令的状态信息来个性化定制电子电路,例如可编程逻辑电路、现场可编程门阵列(FPGA)或可编程逻辑阵列(PLA),该电子电路可以执行计算机可读程序指令,从而实现本公开的各个方面。

[0155] 这里参照根据本公开实施例的方法、装置(系统)和计算机程序产品的流程图和/或框图描述了本公开的各个方面。应当理解,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合,都可以由计算机可读程序指令实现。

[0156] 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器,从而生产出一种机器,使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时,产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中,这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作,从而,存储有指令的计算机可读介质则包括一个制造品,其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

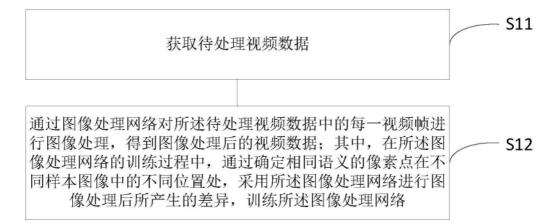
[0157] 也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上,使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤,以产生计算机实现的过程,从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

[0158] 附图中的流程图和框图显示了根据本公开的多个实施例的系统、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、程序段或指令的一部分,所述模块、程序段或指令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。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,两个连续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。也要注意的是,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动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。

[0159] 该计算机程序产品可以具体通过硬件、软件或其结合的方式实现。在一个可选实施例中,所述计算机程序产品具体体现为计算机存储介质,在另一个可选实施例中,计算机程序产品具体体现为软件产品,例如软件开发包(Software Development Kit,SDK)等等。

[0160] 以上已经描述了本公开的各实施例,上述说明是示例性的,并非穷尽性的,并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下,对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择,旨

在最好地解释各实施例的原理、实际应用或对市场中的技术的改进,或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。



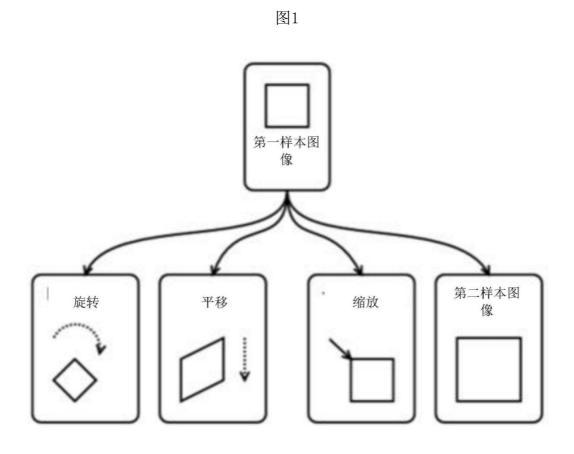


图2

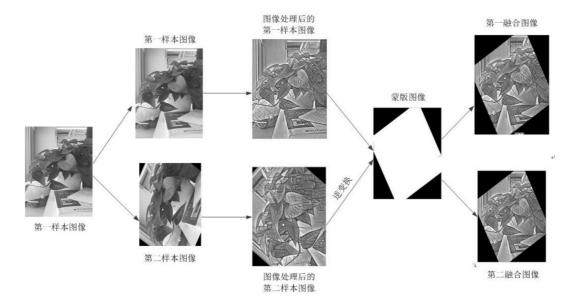


图3

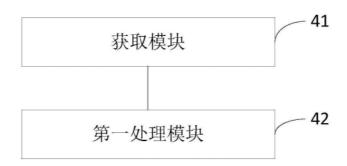
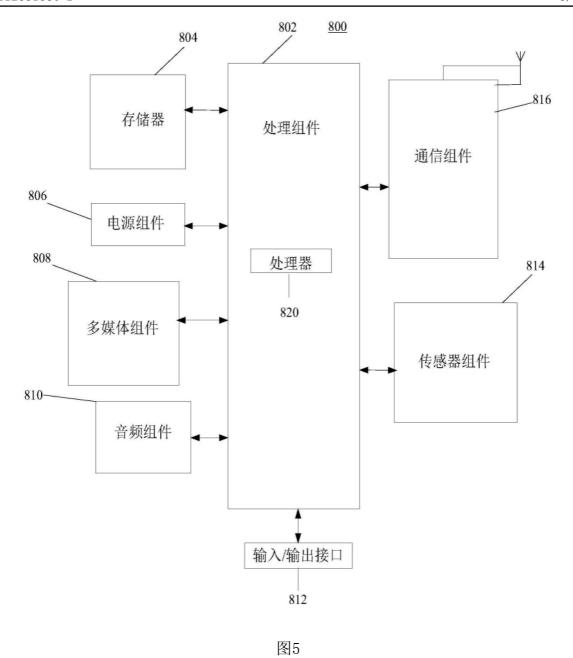


图4



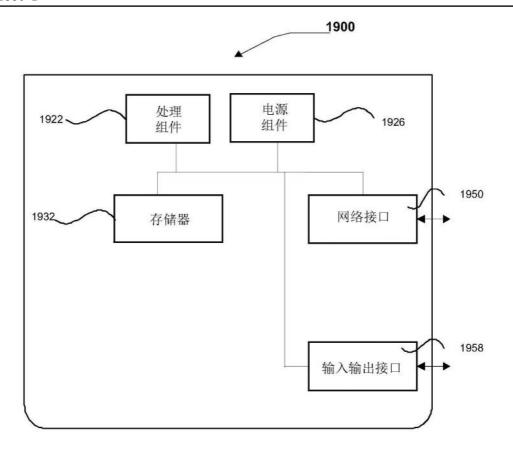


图6