



# (12)发明专利申请

(10)申请公布号 CN 110555360 A

(43)申请公布日 2019.12.10

(21)申请号 201910443199.8

(22)申请日 2019.05.27

(30)优先权数据

62/679,850 2018.06.03 US

16/119,842 2018.08.31 US

(71)申请人 苹果公司

地址 美国加利福尼亚

(72)发明人 A·库马尔 T·戈尔诺斯 汤锋

(74)专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 吴信刚

(51)Int.Cl.

G06K 9/00(2006.01)

G06N 3/08(2006.01)

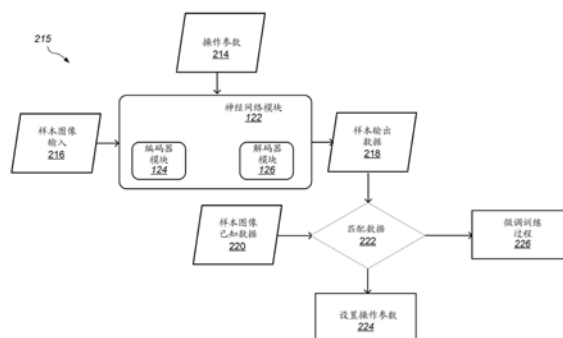
权利要求书2页 说明书17页 附图11页

(54)发明名称

鲁棒的面部检测

(57)摘要

本发明涉及鲁棒的面部检测。神经网络可在移动设备或计算机系统上实施面部检测过程。可使用面部检测过程来处理使用设备上的相机捕获的图像。面部检测过程可为在图像中检测到的面部提供边界框。在图像中检测到的面部可具有任何朝向。在图像中检测到的面部可包括用户的整个面部或仅包括用户的局部面部。面部检测过程可提供针对所捕获图像中的曝光大体上不变的面部检测。还可在相机和用户的面部之间的宽泛的距离范围上检测面部。



1. 一种方法,包括:

使用位于设备上的相机捕获至少一个图像,所述设备包括计算机处理器和存储器,其中所述至少一个图像包括在利用位于所述设备上的照明器对所述图像中的被摄体照明的同时所捕获的图像;

对所述至少一个图像进行编码以生成特征向量,所述特征向量限定特征空间中的所述图像中的被摄体的特征;

评估所述特征空间中的所生成的特征向量以检测所述图像中的面部,其中所述面部的至少某部分存在于所述图像中,并且其中所述面部在所述图像中具有任何朝向;

围绕所述图像中检测到的所述面部形成边界框,其中如果所检测到的面部的一部分位于图像边界之外,则所述边界框围绕检测到存在于所述图像中的所述面部的至少某部分形成;以及

提供输出,所述输出包括针对所述图像中的所检测到的面部形成的所述边界框。

2. 根据权利要求1所述的方法,其中所述计算机处理器包括多尺度卷积神经网络模块。

3. 根据权利要求1所述的方法,其中评估所生成的特征向量以检测所述图像中的所述面部包括确定表示所述图像中的所述面部的所述特征向量。

4. 根据权利要求3所述的方法,其中形成所述边界框包括围绕被确定为表示所述图像中的面部的所述特征向量形成所述边界框。

5. 根据权利要求1所述的方法,其中评估所生成的特征向量以检测所述图像中的所述面部包括对所述特征向量进行分类以确定表示所述图像中的所述面部的所述特征向量。

6. 根据权利要求1所述的方法,其中所述输出包括在所述图像中检测到所述面部的指示。

7. 根据权利要求1所述的方法,其中所述输出包括所述面部在所述图像中的评估姿态以及所述图像中的所述面部距所述相机的评估距离。

8. 一种设备,包括:

相机;

照明器;

存储器,所述存储器被配置为存储程序指令;和

处理器,所述处理器被配置为执行所述程序指令并使得所述设备:

使用所述相机捕获至少一个图像,其中所述至少一个图像包括在利用所述照明器对所述图像中的被摄体照明的同时所捕获的图像;

对所述至少一个图像进行编码以生成特征向量,所述特征向量限定特征空间中的所述图像中的被摄体的特征;

评估所述特征空间中的所生成的特征向量以检测所述图像中的面部,其中所述面部的至少某部分存在于所述图像中,并且其中所述面部被与所述面部相对于所述相机的朝向无关地检测;

围绕所述图像中检测到的所述面部形成边界框,其中如果所检测到的所述面部的一部分位于图像边界之外,则所述边界框围绕检测到存在于所述图像中的所述面部的所述至少某部分形成;以及

提供输出,所述输出包括针对所述图像中的所检测到的面部形成的所述边界框。

9. 根据权利要求8所述的设备,其中所述相机包括红外传感器。
10. 根据权利要求8所述的设备,其中所述程序指令包括在训练过程期间确定的分类参数,并且其中所述处理器被配置为执行所述程序指令并使得所述设备使用所述分类参数来评估所生成的特征向量。
11. 根据权利要求8所述的设备,其中所述处理器被进一步配置为执行所述程序指令并使得所述设备使用针对所述图像中的所检测到的面部形成的所述边界框来操作面部识别过程。
12. 根据权利要求8所述的设备,其中所述照明器包括泛光红外照明器。
13. 根据权利要求8所述的设备,其中所述处理器被配置为围绕在利用由所述设备上的所述相机的曝光极限设定的范围中的任何曝光捕获的图像中检测到的所述面部来形成所述边界框。
14. 一种方法,包括:
- 使用位于设备上的相机捕获至少一个图像,所述设备包括计算机处理器和存储器,其中所述至少一个图像包括在利用位于所述设备上的照明器对所述图像中的被摄体照明的同时所捕获的图像;
- 对所述至少一个图像进行编码以生成特征向量,所述特征向量限定特征空间中的所述图像中的被摄体的特征,其中所述至少一个图像的曝光处于由所述设备上的所述相机的曝光限制设定的曝光范围内;
- 评估所述特征空间中的所生成的特征向量以检测所述图像中的面部,其中所述面部的至少某部分存在于所述图像中;
- 围绕所述图像中检测到的所述面部形成边界框;以及
- 提供输出,所述输出包括针对所述图像中的所检测到的面部形成的所述边界框。
15. 根据权利要求14所述的方法,其中所述至少一个图像在具有挑战性的照明条件下被捕获。
16. 根据权利要求14所述的方法,其中所述计算机处理器包括多尺度卷积神经网络模块。
17. 根据权利要求14所述的方法,其中评估所生成的特征向量以检测所述图像中的所述面部包括确定表示所述图像中的所述面部的所述特征向量。
18. 根据权利要求17所述的方法,其中形成所述边界框包括围绕被确定为表示所述图像中的面部的所述特征向量形成所述边界框。
19. 根据权利要求14所述的方法,其中评估所生成的特征向量以检测所述图像中的所述面部包括对所述特征向量进行分类以确定表示所述图像中的所述面部的所述特征向量。
20. 根据权利要求14所述的方法,其中所述输出包括在所述图像中检测到所述面部的指示。

## 鲁棒的面部检测

### 技术领域

[0001] 本文所述的实施方案涉及用于在设备上的相机所捕获的图像中进行面部检测的方法和系统。更具体地讲,本文所述的实施方案涉及在所捕获的图像中检测面部时使用神经网络。

### 背景技术

[0002] 通常可以使用面部识别过程来识别图像中的个人。可以使用面部检测来检测图像中的面部,以用于面部识别过程。然而,在仅某些情况下,面部检测过程通常仅限于检测面部。例如,当前面部检测过程通常仅以某种朝向(例如,竖直(垂直)纵向或横向模式)检测面部。通常可基于其他传感器数据来旋转图像以提供用于面部检测的竖直图像,这可能是不可靠的和处理器密集型的。

[0003] 如果在图像中仅检测到面部的一部分,则当前面部检测过程通常还拒绝用于面部检测(和下游过程)的图像。此类图像常常被拒绝,因为面部检测在检测局部面部方面不可靠。面部检测过程也通常限于在具有挑战性的照明条件(低光照和/或明亮光照条件)中提供面部检测。使用者的面部和相机之间的距离也可不利地影响面部检测过程的有效性。

### 发明内容

[0004] 设备上的神经网络可对使用设备(例如,移动设备或计算机系统)上的相机所捕获的图像实施面部检测过程。面部检测过程可评估图像中是否有面部,并且如果检测到面部,则为图像中的面部提供边界框。面部检测过程可针对图像中的面部的任何朝向提供面部检测(例如,检测面部,而与面部在图像中的朝向无关)。另外,面部检测过程可为包括用户的整个面部或仅用户面部的一部分的图像提供面部检测。边界框还可针对整个面部或仅针对存在于图像中的局部面部而定位。

### 附图说明

[0005] 当与附图结合时,通过参考根据本公开中描述的实施方案的目前优选的但仅为示例性实施方案的以下详细描述,将更充分地理解在本公开中描述的实施方案的方法与装置的特征和优点,在附图中:

[0006] 图1示出了包括相机的设备的实施方案的表示。

[0007] 图2示出了相机的实施方案的表示。

[0008] 图3示出了设备上的处理器的实施方案的表示。

[0009] 图4示出了神经网络模块的实施方案的表示。

[0010] 图5示出了用于神经网络模块的训练过程的实施方案的流程图。

[0011] 图6示出了具有神经网络模块的处理器的实施方案的表示。

[0012] 图7示出了面部在训练图像中的不同朝向的示例。

[0013] 图8示出了呈现给训练过程的多个面部的示例,其中每个面部具有不同的局部部

分。

[0014] 图9示出了具有区域的特征空间的实施方案的示例。

[0015] 图10示出了在图像输入中围绕面部形成(例如,放置)的边界框的示例。

[0016] 图11示出了用于神经网络模块的测试过程的实施方案的流程图。

[0017] 图12示出了使用设备上的神经网络模块实现的面部检测过程的实施方案的流程图。

[0018] 图13示出了在图像输入中检测到的局部面部的示例。

[0019] 图14示出了示例性计算机系统的一个实施方案的框图。

[0020] 图15示出了计算机可访问存储介质的一个实施方案的框图。

[0021] 尽管本公开中所述的实施方案可受各种修改形式和另选形式的影响,但其具体实施方案在附图中以举例的方式示出并将在本文中详细描述。然而,应当理解,附图和对其的详细描述不旨在将实施方案限制为所公开的特定形式,而相反,本发明旨在涵盖落入所附权利要求书的实质和范围内的所有修改、等同物和另选方案。本文所使用的标题仅用于组织目的,并不旨在用于限制说明书的范围。如在整個本申请中所使用的那样,以允许的意义(即,意味着具有可能性)而非强制的意义(即,意味着必须)使用“可能”一词。类似地,字词“包括(“include”、“including”和“includes”)”是指包括但不限于。

[0022] 各种单元、电路或其他部件可被描述为“被配置为”执行一个或多个任务。在此类上下文中,“被配置为”为通常表示“具有”在操作期间执行一个或多个任务的“电路”的结构宽泛表述。如此,即使在单元/电路/部件当前未接通时,单元/电路/部件也可被配置为执行任务。一般来讲,形成与“被配置为”对应的结构的电路可包括硬件电路和/或存储可执行以实现该操作的程序指令的存储器。该存储器可包括易失性存储器,诸如静态或动态随机存取存储器和/或非易失性存储器,诸如光盘或磁盘存储装置、闪存存储器、可编程只读存储器等。硬件电路可包括组合式逻辑电路、时钟存储设备(诸如触发器、寄存器、锁存器等)、有限状态机、存储器(诸如静态随机存取存储器或嵌入式动态随机存取存储器)、定制设计电路、可编程逻辑阵列等的任意组合。类似地,为了描述方便,可将各种单元/电路/部件描述为执行一个或多个任务。此类描述应当被解释为包括短语“被配置为”。表述被配置为执行一个或多个任务的单元/电路/部件明确地旨在对该单元/电路/部件不调用35U.S.C. § 112(f)的解释。

[0023] 在一个实施方案中,可通过以硬件描述语言(HDL)诸如Verilog或VHDL对电路的描述进行编码来实现根据本公开的硬件电路。可针对为给定集成电路制造技术设计的单元库来合成HDL描述,并可出于定时、功率和其他原因而被修改,以获取可被传输到工厂以生成掩模并最终产生集成电路的最终的设计数据库。一些硬件电路或其部分也可在示意图编辑器中被定制设计并与合成电路一起被捕获到集成电路设计中。该集成电路可包括晶体管并还可包括其他电路元件(例如,无源元件,诸如电容器、电阻器、电感器等),以及晶体管和电路元件之间的互连件。一些实施方案可实现耦接在一起的多个集成电路,以实现硬件电路,和/或可在一些实施方案中使用离散元件。

[0024] 本公开的范围包括本文(明确或暗示)公开的任意特征或特征组合或其任意推广,而无论其是否减轻本文解决的任何或所有问题。因此,在本专利申请(或要求享有其优先权的专利申请)进行期间可针对特征的任何此类组合作出新的权利要求。具体地,参考所附权

利要求书,可将从属权利要求的特征与独立权利要求的特征进行组合,并可通过任何适当的方式而不是仅通过所附权利要求书中所列举的特定组合来组合来自相应独立权利要求的特征。

### 具体实施方式

[0025] 本说明书包括参考“一个实施方案”或“实施方案”。短语“在一个实施方案中”或“在实施方案中”的出现不一定是指相同的实施方案,尽管通常设想包括特征的任何组合的实施方案,除非在本文明确地否认。特定特征、结构或特性可以与本公开一致的任何合适的方式被组合。

[0026] 如上所述,本技术的一个方面在于收集和使用得自各种来源的数据,以改善对设备的操作和访问。本公开预期,在一些实例中,这些所采集的数据可包括唯一地识别或可用于联系或定位特定人员的个人信息数据。此类个人信息数据可包括图像数据(例如,来自用户的图像的数据)、人口数据、基于位置的数据、电话号码、电子邮件地址、家庭地址或任何其他识别信息。对于图像数据,个人信息数据可仅包括来自用户的图像的数据而不包括图像本身。

[0027] 本公开认识到在本发明技术中使用此类个人信息数据可用于使用户受益。例如,个人信息数据可用于使用面部识别来控制对设备的解锁和/或授权。因此,使用此类个人信息数据使得能够对设备的访问进行有计划的控制。此外,本公开还预期个人信息数据有益于用户的其他用途。

[0028] 本公开还设想负责此类个人信息数据的收集、分析、公开、传输、存储或其他用途的实体将遵守已确立的隐私政策和/或隐私做法。具体地,此类实体应当实行并坚持使用被公认为满足或超出对维护个人信息数据的隐私性和安全性的行业或政府要求的隐私政策和实践。例如,在使用面部识别来对设备进行解锁和/或授权的情况下,来自用户的个人信息应被收集用于实体的合法且合理的用途,同样用途仅与设备的操作有关,并且不在这些合法使用之外共享或出售。另外,此类收集应仅在接收用户知情同意之后发生并且用户个人信息数据可在收集个人信息的设备上保持安全。另外,此类实体应采取任何所需的步骤,以保障和保护对此类个人信息数据的访问,并且确保能够访问个人信息数据的其他人遵守他们的隐私政策和程序。另外,这种实体可使其本身经受第三方评估以证明其遵守广泛接受的隐私政策和做法。

[0029] 不管前述情况如何,本公开还预期用户选择性地阻止使用或访问个人信息数据的实施方案。即本公开预期可提供硬件元件和/或软件元件,以防止或阻止对此类个人信息数据的访问。例如,本发明的技术可被配置为在注册服务期间允许用户选择“加入”或“退出”参与对个人信息数据的收集。

[0030] 图1示出了包括相机的设备的实施方案的表示。在某些实施方案中,设备100包括相机102、处理器104、存储器106、和显示器108。设备100可为小型计算设备,在一些情况下,其可足够小以便手持(因此,还被统称为手持式计算机或被简称为手持式设备)。在某些实施方案中,设备100为各种类型的计算机系统设备的任一种类型的计算机系统设备,其为移动的或便携的并且使用WLAN通信来执行无线通信(例如,“移动设备”)。移动设备的示例包括移动电话或智能电话、以及平板电脑。如果包括无线或RF通信能力(例如,Wi-Fi、蜂窝和/

或蓝牙),则各种其他类型的设备可落在这个类别中,诸如膝上型计算机、便携式游戏设备、便携式上网设备、和其他手持式设备、以及可穿戴设备诸如智能手表、智能眼镜、耳机、吊坠、耳塞等。通常,可宽泛地定义术语“移动设备”以涵盖容易被用户转移并能够使用例如WLAN、Wi-Fi、蜂窝和/或蓝牙来进行无线通信的任何电子、计算和/或电信设备(或设备的组合)。在某些实施方案中,设备100包括由用户使用的具有处理器104、存储器106和显示器108的任何设备。例如,显示器108可为LCD屏或触摸屏。在一些实施方案中,显示器108包括用于设备100的用户输入界面(例如,显示器允许用户的交互输入)。

[0031] 相机102可用于捕获设备100的外部环境的图像。在某些实施方案中,相机102被定位成捕获显示器108前方的图像。相机102可被定位成在用户与显示器108进行交互时捕获用户的图像(例如,用户的面部)。图2示出了相机102的实施方案的表示。在某些实施方案中,相机102包括一个或多个镜头和用于捕获数字图像的一个或多个图像传感器103。由相机102所捕获的数字图像可包括例如静止图像、视频图像、和/或逐帧图像。

[0032] 在某些实施方案中,相机102包括图像传感器103。图像传感器103可为例如传感器阵列。传感器阵列中的传感器可包括但不限于用于捕获红外图像(IR)或其他不可见电磁辐射的电荷耦合器件(CCD)和/或互补金属氧化物半导体(CMOS)传感器元件。在一些实施方案中,相机102包括用于捕获多种类型的图像的多于一个图像传感器。例如,相机102可包括IR传感器和RGB(红、绿和蓝)传感器两者。在某些实施方案中,相机102包括用于利用由图像传感器103检测到的不同类型的光来对表面(或被摄体)进行照明的照明器105。例如,相机102可包括用于可见光的照明器(例如,“闪光灯”),用于RGB光的照明器,和/或用于红外光的照明器(例如,泛光IR光源和光斑图案投影仪)。在一些实施方案中,泛光IR光源和光斑图案投影仪为其他波长的光(例如,非红外)。在某些实施方案中,照明器105包括光源阵列,诸如但不限于VCSEL(垂直腔面发射激光器)。在一些实施方案中,图像传感器103和照明器105被包括在单个芯片封装中。在一些实施方案中,图像传感器103和照明器105位于独立芯片封装上。

[0033] 在某些实施方案中,图像传感器103为用于捕获红外图像以用于进行面部检测,面部识别认证和/或深度检测的IR图像传感器和图像传感器。也可设想图像传感器103(例如,RGB图像传感器)的其他实施方案用于如本文所述的面部检测,面部识别认证和/或深度检测。为了进行面部检测,照明器105A可提供泛光IR照明,以利用IR照明(例如,IR闪光灯)来对被摄体进行泛光,并且图像传感器103可捕获泛光IR照明被摄体的图像。泛光IR照明图像例如可为由IR光照明的被摄体的二维图像。为了进行深度检测或生成深度图图像,照明器105B可提供具有光斑图案的IR照明。光斑图案可为具有已知且可控配置的光点的图案(例如,点的图案)和投射到被摄体上的图案。照明器105B可包括被配置为形成光斑图案的VCSEL阵列或被配置为形成光斑图案的光源和图案化幻灯片。由照明器105B提供的光斑图案的构型和图案可例如基于被摄体处的所期望的光斑图案密度(例如,点密度)而被选择。图像传感器103可捕获由光斑图案照明的被摄体的图像。可由成像和处理系统(例如,本文所述的图像信号处理器(ISP))来评估(例如,分析和/或处理)所捕获的被摄体上的光斑图案的图像,以产生或估计被摄体的三维图(例如,被摄体的深度图或深度图图像)。在授予Freedman等人的美国专利No.8,150,142、授予Pesach等人的美国专利No.8,749,796和授予Shpunt等人的美国专利No.8,384,997(在此通过引用并入,如同本文完全阐述的一样)以及

授予Mor等人的美国专利申请公开No.2016/0178915 (在此通过引用并入,如同本文完全阐述的一样)中描述了深度图成像的示例。

[0034] 在某些实施方案中,由相机102捕获的图像包括具有用户的面部的图像(例如,用户面部被包括在图像中)。具有用户的面部的图像可包括具有在图像的帧内示出的用户的面部的至少一部分的任何数字图像。此类图像可仅包括用户的面部或可在图像的较小分部或部分中包括用户的面部。可在图像中以足够的分辨率来捕获用户的面部,以允许对图像中的用户的面部的一个或多个特征的图像处理。

[0035] 由相机102捕获的图像可由处理器104来处理。图3示出了被包括在设备100中的处理器104的实施方案的表示。处理器104可包括被配置为执行在由处理器实施的指令集架构中定义的指令的电路。处理器104可执行设备100的主控制软件,诸如操作系统。通常,由处理器104在使用期间执行的软件可控制设备100的其他部件,以实现设备的所期望的功能。处理器还可执行其他软件。这些应用程序可提供用户功能并可依赖于操作系统,以进行低级设备控制、调度、存储器管理等。

[0036] 在某些实施方案中,处理器104包括图像信号处理器 (ISP) 110。ISP110可包括适用于处理从相机102接收的图像的电路(例如,图像信号处理电路)。ISP 110可包括能够处理或分析由相机102捕获的图像的任何硬件和/或软件(例如,程序指令)。

[0037] 在某些实施方案中,处理器104包括安全区域处理器 (SEP) 112。在一些实施方案中,在面部识别认证过程中涉及SEP 112,该过程涉及由相机102捕获并由ISP 110处理的图像。SEP 112可为被配置为对被授权使用设备100的活动用户(例如,当前正在使用设备100的用户)进行认证的安全电路。“安全电路”可为保护被隔离的内部资源免受外部电路直接访问的电路。内部资源可为存储敏感数据诸如个人信息(例如,生物识别信息、信用卡信息等)、加密密钥、随机数生成器种子等的存储器(例如,存储器106)。内部资源也可执行与敏感数据相关联的服务/操作的电路。如本文所述,SEP 112可包括能够使用面部识别认证过程来认证用户的任何硬件和/或软件(例如,程序指令)。面部识别认证过程可通过利用相机102捕获用户的图像并且将所捕获的图像与设备100的授权用户的先前收集的图像进行比较来认证用户。在一些实施方案中,ISP 110和SEP 112的功能可由单个处理器来执行(例如,ISP 110或SEP 112可执行两种功能,并且另一个处理器可省略)。

[0038] 在某些实施方案中,处理器104执行注册过程(例如,图像注册过程或登记过程),以捕获设备100的授权用户的(例如,先前收集的图像)。在注册过程期间,相机模块102可从授权用户捕获(例如,收集)图像和/或图像数据,以便允许SEP 112(或另一安全过程)随后使用面部识别认证过程来认证用户。在一些实施方案中,来自注册过程的图像和/或图像数据(例如,来自图像的特征向量数据)用来生成设备100中的模板。模板可被存储在例如设备100的存储器106中的模板空间中。在一些实施方案中,可通过从模板空间添加和/或减去模板来更新模板空间。处理器104可执行模板更新过程,以从模板空间添加和/或减去模板。例如,模板空间可利用附加模板而被更新,以适应授权用户的外观随时间推移的变化和/或硬件性能随时间推移的变化。当用于存储模板的模板空间已满时,可从模板空间中减去模板以补偿模板的添加。

[0039] 在一些实施方案中,相机模块102捕获用于面部识别会话的多对图像。每对可包括使用二维捕获模式所捕获的图像(例如,泛光IR图像)和使用三维捕获模式所捕获的图像



(例如,深度图图像)。在某些实施方案中,在对用户作出最终认证决定之前,ISP 110和/或SEP 112彼此独立地处理泛光IR图像和深度图图像。例如,ISP 110可独立处理图像,以分别确定每个图像的特征。然后,SEP 112可将每个类型图像的单独图像特征与所存储的模板进行比较,以生成每个单独图像的认证分数(例如,所捕获的图像和所存储的模板中的用户之间的匹配分数或匹配的其他排名)。单独图像(例如,泛光IR图像和深度图图像)的认证分数可组合,以对用户的身份作出决定,并且如果被认证,则允许用户使用设备100(例如,对解锁进行设备)。

[0040] 在一些实施方案中,ISP 110和/或SEP 112组合每对中的图像,以提供用于面部识别的合成图像。在一些实施方案中,ISP 110处理合成图像,以确定图像的特征,SEP 112可将其与所存储的模板进行比较,以对用户的身份作出决定,并且如果被认证,则允许用户使用设备100。

[0041] 在一些实施方案中,泛光IR图像数据和深度图图像数据的组合可允许SEP 112在三维空间中比较面部。在一些实施方案中,相机模块102经由安全通道来将图像数据传送到SEP 112。例如,安全信道可为用于传送数据的专用路径(即,仅由预期的参与者共享的路径),或者可为用于使用仅预期的参与者知道的密码密钥来传送加密的数据的专用路径。在一些实施方案中,相机模块102和/或ISP 110可在将图像数据提供至SEP 112之前对图像数据执行各种处理操作,以便促进由SEP来执行该比较。

[0042] 在某些实施方案中,处理器104操作一个或多个机器学习模型。机器学习模型可使用位于处理器104和/或设备100上的硬件和/或软件的任何组合(例如,程序指令)来操作。在一些实施方案中,使用一个或多个神经网络模块114来操作设备100上的机器学习模型。神经网络模块114可位于ISP 110和/或SEP 112中。

[0043] 图4示出了神经网络模块114的实施方案的表示。神经网络模块114可包括位于处理器104中和/或设备100上的硬件和/或软件的任何组合(例如,程序指令)。在一些实施方案中,神经网络模块114为多尺度神经网络或者其中在网络中所使用的内核的尺度可变化的另一神经网络。在一些实施方案中,神经网络模块114为重复性神经网络(RNN),诸如但不限于门控重复性单元(GRU)重复性神经网络或长短期记忆(LSTM)重复性神经网络。

[0044] 神经网络模块114可包括安装或配置有已由该神经网络模块或类似的神经网络模块(例如,在不同的处理器或设备上操作的神经网络模块)学习的操作参数的神经网络电路。例如,可使用训练图像(例如,参考图像)和/或其他训练数据来训练神经网络模块,以生成神经网络电路的操作参数。随后可将从训练中生成的操作参数提供至被安装在设备100上的神经网络模块114。将从训练中生成的操作参数提供至允许神经网络模块使用编程到神经网络模块中的训练信息进行操作的设备100上的神经网络模块114(例如,训练生成的操作参数可被神经网络模块用于操作并评估由设备捕获的图像)。

[0045] 在某些实施方案中,神经网络模块114包括编码器模块116和解码器模块118。编码器模块116和解码器模块118可为在神经网络模块114内操作的机器学习模型(例如,在神经网络模块中执行编码器模块和解码器模块)。编码器模块116可对输入到编码器模块中的图像进行编码,并将图像中的特征定义为特征空间中的特征向量(如本文所述)。解码器模块118可对由编码器模块116生成的特征空间中的特征向量进行解码并提供输出(如本文所述)。

[0046] 图5示出了用于神经网络模块的训练过程200的实施方案的流程图。在某些实施方案中,使用位于除处理器104之外的计算机处理器上的神经网络模块(例如,训练神经网络模块)来实现训练过程200。图6示出了具有可用于训练的神经网络模块122(例如,训练神经网络模块)的处理器120的实施方案的表示。神经网络模块122可包括编码器模块124和解码器模块126。在某些实施方案中,先前已被捕获的图像作为训练图像而被提供至神经网络模块122。可将训练图像的已知属性连同训练图像一起提供至神经网络模块122(例如,训练图像可利用已知属性来扩充)。在一些实施方案中,相机102可耦接到处理器120和/或神经网络模块122。相机102可用于捕获训练图像的图像并将相机捕获的图像提供至神经网络模块122。

[0047] 编码器模块124和解码器模块126可分别与编码器模块116和解码器模块118基本相似或基本相同。编码器模块124和解码器模块126可位于处理器120上的神经网络模块122中,以便通过训练过程200来进行训练。然后,可将由“训练的”神经网络模块122生成的操作参数输出用于设备100上的神经网络模块114,以实现设备上的“训练的”神经网络模块。

[0048] 在一些实施方案中,处理器120为支持GPU的计算机处理器。支持GPU的计算机处理器上的训练神经网络模块122可使用浮点数表示模式来输出操作参数。例如,从“训练的”神经网络模块122生成的操作参数可包括使用浮点数描述的权重或内核。在此类实施方案中,在设备100上的神经网络模块114上使用浮点操作参数之前,可能需要将浮点操作参数转换成整数表示。本领域中已知的任何转换过程可用于将操作参数从浮点数表示模式转换为整数表示模式。

[0049] 在某些实施方案中,如图5所示,训练过程200开始于提供图像输入202。图像输入202可包括提供至训练过程200的训练图像(例如,将利用如上所述的已知信息扩充或注释的(例如,标记的)训练图像提供至训练过程)。在一些实施方案中,图像输入202包括借助相机102捕获或者以另外方式提供给训练过程200(例如,以数字方式提供给训练过程)的训练图像。训练图像可包括从图像的数据库获取的参考图像或其他样本图像。例如,可从ImageNet或另一类似图像数据库获取训练图像。在某些实施方案中,在泛光IR照明图像上实现训练过程200,以在此类图像中训练用于面部检测的神经网络模块。在一些实施方案中,在深度图图像上实现训练过程200,以在此类图像中训练用于面部检测的神经网络模块。

[0050] 图像输入202可包括图像中的具有各种不同用户和/或面部的多个训练图像。在一些实施方案中,图像中的面部可在图像中具有不同的位置和/或在图像中具有不同的姿态。训练图像中的面部的位置和/或姿态可为已知的(例如,该图像具有识别位置和姿态的已知信息的标签或其他标记)。可将位置和姿态的已知信息作为已知数据204提供至训练过程200中。在一些实施方案中,训练图像利用已知数据204来扩充。

[0051] 在某些实施方案中,图像输入202包括在不同情况下向训练过程200提供训练图像。例如,单个或独立训练图像(或选择训练图像)可在不同场景和/或不同操作条件下被提供至训练过程中多次(例如,图像的多次传递或呈现)。在训练图像到训练过程200的每次呈现中,图像可呈现在不同的情况下。不同的情况可包括但不限于图像中面部的不同朝向,不同曝光,不同距离和/或不同部分。在多个不同情况下呈现训练图像可提供更鲁棒的训练过程200。更鲁棒的训练过程可生成神经网络模块122的操作参数输出,其使用设备100上的神

神经网络模块114提供更鲁棒的面部检测(具有来自在神经网络模块114上实现的神经网络模块122的输出操作参数)。例如,神经网络模块114可能够基于更鲁棒的训练过程提供多种不同的操作条件和许多不同图像输入情况下的面部检测。

[0052] 在某些实施方案中,在训练过程中,在图像的每个呈现(例如,图像的每个捕获或处理)中以不同朝向向训练过程200提供训练图像。以不同朝向提供训练图像可将用户在图像中的面部以不同朝向提供给训练过程200。在训练图像的呈现中的不同朝向可包括任何朝向。在一些实施方案中,训练图像以四种不同的朝向(例如,面部处于四个不同朝向)提供,如图7-(a)垂直纵向;(b)右转横向;(c)左转横向;和(d)倒置纵向。然而,可以任何所需的朝向提供训练图像。例如,可将单个训练图像在图像的不同呈现之间以 $30^\circ$ 增量旋转到训练过程200。在一些实施方案中,在每个朝向中向神经网络模块122提供相等次数的训练图像。

[0053] 利用以不同朝向提供的训练图像训练神经网络模块122可产生用于神经网络模块的操作参数输出,该神经网络模块提供对图像中的朝向更鲁棒的面部检测。例如,可使用从神经网络模块122输出的操作参数来实现神经网络模块114。使用这些操作参数,神经网络模块114可提供在神经网络模块遇到具有不同朝向的图像时更鲁棒的面部检测。例如,神经网络模块114能够为以任何朝向呈现的面部提供面部检测(例如,可在纵向右侧向上,纵向倒置,左转横向或右转横向上检测面部)。

[0054] 在一些实施方案中,向训练过程200提供训练图像,其中用户的整个面部在呈现给训练过程的图像的帧内(例如,用户的整个面部位于由训练过程处理的图像的帧内)。在某些实施方案中,向训练过程200提供训练图像,其中在图像的帧内仅将用户的面部的一部分提供给训练过程。例如,训练图像可向上,向下,向左或向右移动,以使用户的面部朝提供给训练过程200的图像中的帧的边缘移动。此类提供的图像可具有提供移动到提供的图像的边缘之外的面部的至少一部分。将面部的至少一些部分移动到所提供图像的边缘之外可仅向训练过程200提供用户面部的一部分(例如,局部面部)(例如,仅局部面部通过训练过程进行处理)。

[0055] 在某些实施方案中,训练图像被提供给训练过程200多次(例如,训练图像到训练过程的多个展示),该训练图像在训练图像的每个单独呈现中具有不同的局部面部。例如,训练图像的每个单独呈现可将面部的不同局部部分提供给训练过程。面部的不同局部部分可包括被定位在呈现给训练过程200的图像的帧中的面部的不同区域。例如,图像的帧可具有面部前额被修剪或切除的一部分,图像的帧可具有下巴被修剪或切除的一部分,和/或图像的帧可具有面部被修剪或切除的一侧。

[0056] 图8示出了呈现给训练过程200的多个面部的示例,其中每个面部具有不同的局部部分。在帧272之外的面部270的部分不被提供给训练过程200(例如,从图像向训练过程的呈现中被修剪或切除)。因此,对训练过程200进行训练以检测所述训练图像中的面部,其中图像中仅存在面部的一部分。

[0057] 在某些实施方案中,操作训练过程200以允许在图像中小于用户的整个面部的面部检测(例如,仅需要面部的局部百分比来提供面部检测)。训练过程通常经过训练,以仅允许面部检测,其中具有1的联合交集(其中联合交集是将图像中的面部的预测边界框与实际/已知(地面真实)边界框进行比较的度量)。1的联合交集提供预测边界框与地面真实边

界框完全重叠(具有完全联合)。然而,可操作训练过程200以针对小于1但高于所选值(例如,0.8,0.85,0.9或更大)的联合交集提供面部检测。以联合交集小于1来操作训练过程200为训练图像中的局部面部提供面部检测。然而,联合交集值可在所选值的上方操作,以在面部检测过程中保持准确性(例如,防止误报)。

[0058] 利用针对每个面部的不同局部部分提供的训练图像来训练神经网络模块122并允许面部的局部百分比的面部检测可能为神经网络模块产生操作参数输出,该神经网络模块提供对图像中存在的局部面部更鲁棒的面部检测。例如,可使用从神经网络模块122输出的操作参数来实现神经网络模块114。通过使用这些操作参数,神经网络模块114可通过在图像帧中仅具有局部面部的捕获图像中提供面部检测而更鲁棒。另外,用提供有不同局部部分的训练图像训练神经网络模块122并且允许对面部的局部百分比的面部检测可改善图像中具有整个面部的图像的面部检测。

[0059] 在某些实施方案中,提供给训练过程200的训练图像包括在图像中没有面部的训练图像。没有面部的训练图像可用于训练在某些情况下未检测到面部的神经网络模块122。神经网络模块122的训练可与用于检测局部面部的神经网络模块的训练结合使用,以防止误接率的增加并且在面部检测中提供更好的准确性。

[0060] 在某些实施方案中,将训练图像多次提供给训练过程200(例如,训练图像到训练过程的多个展示),其中在训练图像的每个单独演示中具有不同的曝光。例如,可在第一展示中以正常曝光,在第二展示中以较暗曝光,以及在第三展示中以较亮曝光来提供训练图像。

[0061] 利用具有不同曝光的训练图像训练神经网络模块122可产生神经网络模块的操作参数输出,该神经网络模块提供对所捕获图像中的不同曝光更鲁棒的面部检测。例如,可使用从神经网络模块122输出的操作参数来实现神经网络模块114。通过使用这些操作参数,神经网络模块114可通过在较宽范围的曝光条件下提供面部检测而更鲁棒。神经网络模块114能够提供对图像中的曝光的变化基本上不变的面部检测。因此,使用由提供更鲁棒操作的操作参数实现的神经网络模块114,对于由设备100捕获的图像的曝光调节可能不是必要的。

[0062] 在一些实施方案中,在图像和相机之间的变化距离处提供了训练图像(例如,展示给训练过程200)。每个所提供的图像的距离的值可为已知的,并且已知的信息可与位置和姿态的已知信息一起被提供至已知数据204中。因此,这些属性的已知信息(面部的位置和/或姿态、以及面部和相机之间的距离)被包括在已知数据204中。

[0063] 利用在不同距离处提供的训练图像来训练神经网络模块122可产生神经网络模块的操作参数输出,该神经网络模块提供对所捕获图像中距相机的距离更鲁棒的面部检测。例如,可使用从神经网络模块122输出的操作参数来实现神经网络模块114。使用这些操作参数,神经网络模块114能够在用户的面部和设备100之间的更大范围的距离上提供面部检测。因此,神经网络模块114可相对于用户的面部和设备100之间的距离提供更鲁棒的面部检测。

[0064] 如图5所示,可将图像输入202提供至解码器过程206。编码器过程206可由例如图6所示的编码器模块124来执行。在编码器过程206中,如图5所示,编码器模块可编码被输入到编码器过程中的图像,并将图像中的特征定义为特征空间中的特征向量。例如,编码器模

块可将用户面部中的面部特征 (以及图像中的其他特征) 定义为特征空间中的特征向量。编码器过程206可输出特征向量208。特征向量208 (例如, 编码器过程206 (以及编码器模块) 的输出) 包括表示特征空间中的用户的面部特征 (和/或图像中的其他特征) 的特征向量。特征空间可为N维特征空间。特征向量可为数值的n维向量, 该数值针对特征空间中的与图像中的区域对应的区域来定义图像中的特征。例如, 在一些实施方案中, 特征向量可为1024维向量。在各种实施方案中可实现任何数量的维度。

[0065] 图9示出了具有区域132的特征空间130的实施方案的示例。特征空间130中的区域132可为例如栅格中的小区, 其中栅格表示特征空间。在图9的示例中, 特征空间130为区域132的 $8 \times 8$ 栅格。然而, 特征空间130可根据需要而具有不同维度的栅格。点134表示每个区域132中的特征向量。特征空间130中的区域132可对应于输入图像中的区域或区。因此, 在图9的示例中, 输入图像在特征空间130中被划分为64个区域 ( $8 \times 8$ 个区域), 其中每个区域132表示输入图像的不同区域。

[0066] 在某些实施方案中, 在编码器过程206中使用的编码器模块 (例如, 在图6中所示的编码器模块124) 为神经网络。例如, 编码器模块可为多尺度神经网络或者其中在网络中所使用的内核的尺度可变化的另一神经网络。在某些实施方案中, 编码器模块为多尺度卷积神经网络。使用多尺度卷积神经网络, 编码器过程206可在特征空间中生成具有高级特征向量的图像输入202的高级表示。例如, 编码器过程206可在栅格的每个区域 (小区) 中生成具有特征向量的 $32 \times 32$ 栅格表示, 而输入图像可具有更高的分辨率 (例如, 图像输入202可为 $256 \times 256$ 图像)。

[0067] 如图5所示, 可将特征向量208提供至解码器过程210中。解码器过程210可由例如图6所示的解码器模块126来执行。在解码器过程210中, 解码器模块可对在编码器过程206中生成的特征向量208的特征空间中的特征向量进行解码。解码特征向量可包括利用一个或多个分类器或支持分类的网络对特征向量进行操作, 以从图像输入202确定 (例如, 提取) 输出数据212。输出数据212可包括例如关于图像输入202中的面部的信息或属性。

[0068] 在某些实施方案中, 在解码器过程210 (例如, 解码器模块126) 中使用的解码器模块为神经网络。例如, 解码器模块可为重复性神经网络 (RNN)。在某些实施方案中, 所述复发神经网络包括门控复发单元 (GRU)。但是, 也可使用其他重复性神经网络, 诸如长期短期记忆 (LSTM) 重复性神经网络。

[0069] 在某些实施方案中, 解码器过程210包括对特征空间中的每个区域 (例如, 图9的示例中所示的特征空间130中的每个区域132) 的特征向量进行解码。来自特征空间的每个区域的特征向量可被解码成输出数据212中的不重叠的框。在某些实施方案中, 对某个区域的特征向量进行解码 (例如, 从特征向量提取信息) 包括确定 (例如, 检测) 该区域中是否存在面部。在训练过程200中, 图像输入202中的面部的存在为已知的并且可与经解码的特征向量相关。由于解码器过程210对特征空间中的每个区域进行操作, 因此解码器模块可针对特征空间中的每个区域提供面部检测分数 (例如, 基于关于在该区域中是否检测到/存在面部或面部的一部分的置信度分数的预测)。在一些实施方案中, 使用RNN, 可针对特征空间的每个区域提供关于是否存在面部的多个预测, 其中该预测包括关于区域内的面部部分和区域周围 (例如, 在相邻区域) 的面部部分两者的预测。这些预测可折叠成图像输入202中的面部的存在的最终决定 (例如, 面部及其在图像输入202中的位置的检测)。

[0070] 在某些实施方案中,预测用于在图像输入202中检测到的面部周围形成(例如,放置)边界框。图10示出了在图像输入202中围绕面部270形成(例如,放置)的边界框274的示例。输出数据212可包括对面部和围绕图像输入202中的面部形成的边界框的存在的决定。

[0071] 在某些实施方案中,在与相邻区域没有太多重叠的区域中检测到面部,因为该区域被解码为不重叠的框。然而,在一些实施方案中,在解码器处理210中解码的多个区域可检测相同的面部。如果在多个区域中检测到相同的面部,则可对这些区域的置信度进行排序。可使用多个预测来确定在每个区域中存在面部或面部的一部分的置信度(例如,预测可用于对区域的置信度进行排序)。然后可将对于所检测到的面部具有最高置信度的一个或多个区域选择作为在训练过程200中使用的区域。

[0072] 在某些实施方案中,当检测到面部的存在时,由解码器过程210生成的预测包括对所检测到的面部的一个或多个属性的评估(例如,确定)。所评估的属性可包括面部在图像中的位置(其可由图像中的面部的边界框表示),面部的姿态(例如,面部在边界框中的姿态),以及面部和相机之间的距离。面部的姿态可包括面部的倾斜、偏转、和/或转动。所评估的属性可连同关于在图像输入202中存在面部的决定一起被包括在输出数据212中。

[0073] 在训练过程200中,可通过将经解码的特征向量与已知数据204相关来确定面部的属性的值。例如,已知数据204可提供图像输入202中的一个或多个面部的已知属性,该图像输入具有定义由解码器过程210评估的属性的已知属性。在某些实施方案中,在训练过程200期间,使经解码的特征向量与已知数据204相关包括用于解码器过程210的解码器模块,该解码器过程评估经解码的特征向量与已知数据204之间的差异。例如,检测器模块可对经解码的特征向量和已知数据204之间的差异执行误差函数分析(或类似分析)并细化特征向量解码过程,直到特征向量解码过程准确地确定已知数据。因此,当在训练过程200中处理多个训练图像时,可通过图像输入202中的训练图像和已知数据204来训练解码器过程210(和编码器过程206),以准确地检测一个或多个面部的存在并评估一个或多个面部的属性的值。

[0074] 在某些实施方案中,面部的姿态和/或面部与相机之间的距离的输出被离散化(例如,被提供作为离散输出)。例如,倾斜、偏转和/或转动值可被解码为浮点值。在一些实施方案中,浮点值可为正浮点值或负浮点值。代替对浮点值执行回归,可通过选择最小范围和最大范围来将浮点输出离散化,并且然后将浮点输出分成K个区段,其中K为正整数。使用这些区段,如果输出落入一个区段,则其被赋值为1,如果输出不落入一个区段,则其被赋值为0。如果浮点值不在由该区段表示的范围中,则可能先将其剪裁到该范围内的最接近的值。因此,浮点输出可从浮点值转换为0和1的离散向量(例如,特征向量为0和1的离散向量)。然后可训练网络(例如,编码器模块),以预测K维向量,而不是单个浮点值。在运行时(例如,在设备上的操作期间),通过将每个区段的网络激活视为权重,可从这些K维输出中恢复单个浮点值。然后取每个区段的中心值的加权和可产生单个浮点值。

[0075] 例如,最小范围和最大范围可为0到10,并且存在十个区段。然后,如果浮点训练目标介于0和1之间,则将其分配给第一个区段,如果介于1和2之间,则分配给第二个区段,并且依此类推。将低于0的值分配给第一个区段,并且将高于10的值分配给最后一个区段。利用这个过程,训练值2.4将被转换成向量(0 0 1 0 0 0 0 0 0 0),训练值-1.3将被转换成向量(1 0 0 0 0 0 0 0 0 0),并且训练值11.9将被转换成向量(0 0 0 0 0 0 0 0 0 1)。

在运行时,如果网络输出向量为(0 0 1 1 0 0 0 0 0),则加权求和程序将得到值3.0。

[0076] 在一些实施方案中,在训练期间,K维向量可基于使用任何合适的算法或公式的“软”分配。例如,如上所述给定一个初始区段分配,则也可相邻区段给定与目标和区段的中心值之间的差异相关的值。作为示例,以上示例中的训练值2.4可基于简单的指数公式而被转换成向量(0.67 1.54 0 0 0 0 0 0)。

[0077] 将浮点值转换为离散向量允许解码器过程210(和解码器模块)通过将该值分类到哪个区段中而不是使用浮点值所需的回归解决方案来对面部的姿态和/或面部与相机之间的距离的值进行操作。在分类之后,解码器过程210可包括区段的中心所表示的浮点值的加权和(例如,区段的峰值的加权平均值)的映射。离散向量和区段的分类和映射可提供相对精确的姿态和/或距离评估的输出。

[0078] 使用对离散向量的分类而不是对浮点值的回归可允许解码器过程210更容易学习(例如,在训练过程200中被训练),因为神经网络在进行分类方面通常比回归更好。此外,回归的误差功能信号可相对较大,因为回归中的误差函数信号越大则差异越大,而无论误差中的差异有多大,离散向量和区段的误差函数信号基本上相同。因此,相比于使用浮点值,在解码器过程210中使用离散向量和区段来评估姿态和/或距离对于解码器过程学习可更有效。

[0079] 如上所述,训练过程200可包括在多个训练图像上训练编码器过程206和解码器过程210(及其对应的编码器模块和解码器模块),其具有图像中的各种不同的用户和/或面部、以及图像中的面部的不同属性和/或情况。在对一组训练图像完成训练过程200之后,操作参数214可由该训练过程基于经解码的特征向量和已知数据204之间的相关性来生成。操作参数214包括如图6所示的可用于神经网络模块122(例如,编码器模块124和解码器模块126)的参数,以检测从相机102输入到神经网络模块中的一个或多个面部并评估该一个或多个面部的属性的值(例如,面部的位置(如面部的边界框表示)、面部的姿态、以及面部和相机之间的距离)。在一些实施方案中,操作参数214包括用于解码器模块126中的分类参数。分类参数可包括用于对在训练过程200期间与已知数据204相关的解码特征向量进行分类的参数。然后,解码器模块126能够使用分类参数对由编码器模块124生成的所捕获的图像的特征向量进行分类。通过(使用分类参数)对特征向量进行分类来对所捕获的图像的特征向量进行解码可允许神经网络模块122对一个或多个面部的存在以及所捕获的图像中的一个或多个面部的属性的值进行评估。

[0080] 在一些实施方案中,可通过将操作参数输入到神经网络模块122中并且使用已知信息(例如,已知的面部位置(已知边界框位置)、已知的姿态、和已知的距离)在样本图像上操作该模块来测试操作参数214。图11示出了用于神经网络模块122的测试过程的实施方案的流程图。在测试过程215中,可将样本图像输入216连同操作参数214一起提供至神经网络模块122。神经网络模块122可通过使用操作参数214对样本输入图像216进行处理来提供样本输出数据218。可将样本输出数据218与样本图像已知数据220进行比较,以查看该数据是否在匹配数据222中匹配。

[0081] 如果样本输出数据218与样本图像已知数据220匹配,则在224中设置该操作参数(例如,如图3所示,操作参数214可被设置并用于对处理器104上的神经网络模块114进行编程,以用于本文所述的面部检测过程)。如果样本输出数据218不与样本图像已知数据220匹



配(在期望的公差内),则可在226中对训练过程(例如,在图5中示出的训练过程200)进行微调。对训练过程进行微调可包括向训练过程200提供附加训练图像和/或训练过程中的其他调整,以优化神经网络模块122的操作参数(或生成新的操作参数)。

[0082] 一旦在224中设置用于神经网络模块122的操作参数214,则可通过将该操作参数提供至如图1所示的设备100上的神经网络模块114而将该操作参数应用于该设备。在某些实施方案中,用于神经网络模块122的操作参数214处于与神经网络模块114用于操作的数值表示模式不同的数值表示模式中。例如,神经网络模块122可使用浮点数,而神经网络模块114使用整数。因此,在此类实施方案中,用于神经网络模块122的操作参数214从浮点操作参数被转换成用于神经网络模块114的整数操作参数。

[0083] 在将操作参数提供至神经网络模块114之后,神经网络模块可在设备100上进行操作,以在设备上实现面部检测过程。图12示出了使用设备100上的神经网络模块114实现的面部检测过程250的实施方案的流程图。图像输入252可包括使用设备100上的相机102捕获的图像。所捕获的图像可为泛光IR照明图像或深度图图像。面部检测过程250可用于检测图像中是否存在面部(例如,在面部周围放置边界框),并且如果检测到面部,则评估该面部的属性的值(例如,位置、姿态、和/或距离)。

[0084] 可将来自图像输入252的所捕获的图像提供至编码器过程254。编码器过程254可由如图4所示的编码器116来执行。在某些实施方案中,编码器模块116为多尺度卷积神经网络(例如,编码器模块116与编码器模块124为基本上相同的神经网络)。在编码器过程254中,编码器模块116可对图像输入252进行编码,以将图像中的特征表示为特征空间(例如,与图9所示的特征空间130基本类似的特征空间)中的特征向量。编码器过程254可输出特征向量256。特征向量256可为(例如)被表示作为向量的经编码的图像特征。

[0085] 可将特征向量256提供至解码器过程258中。解码器过程258可由图4所示的编码器模块118来执行。在某些实施方案中,解码器模块118为重复性神经网络(例如,解码器模块118与解码器模块126为基本上相同的神经网络)。在解码器过程258中,解码器模块可对特征向量256进行解码,以评估图像输入252的一个或多个属性,从而确定(例如,提取)来自该图像输入的输出数据260。对特征向量进行解码可包括使用在训练过程200期间确定的分类参数来对特征向量进行分类。对特征向量进行分类可包括利用一个或多个分类器或支持分类的网络来对特征向量进行操作。

[0086] 在某些实施方案中,解码器过程258包括对特征空间中的每个区域的特征向量进行解码。来自特征空间的每个区域的特征向量可被解码成输出数据260中的不重叠的框。在某些实施方案中,对某个区域的特征向量进行解码(例如,从特征向量提取信息)包括确定(例如,检测)该区域中是否存在面部。由于解码器过程258对特征空间中的每个区域进行操作,因此解码器模块可针对特征空间中的每个区域提供面部检测分数(例如,基于关于在该区域中是否检测到/存在面部或面部的一部分的置信度分数的预测)。在一些实施方案中,使用RNN,可针对特征空间的每个区域提供关于是否存在面部(或面部一部分)的多个预测,其中该预测包括关于区域内的面部部分和区域周围(例如,在相邻区域)的面部部分两者的预测。这些预测可折叠成图像输入252中的面部的存在的最终决定(例如,面部在图像输入中的检测)。在某些实施方案中,预测用于在图像输入252中检测到的面部周围形成(例如,放置)边界框。输出数据260可包括关于在图像输入252中(例如,在所捕获的图像中)存在面



部和围绕面部形成的边界框的决定。

[0087] 在某些实施方案中,面部检测过程250检测图像输入252中的面部是否存在,而与面部在图像输入中的朝向无关。例如,神经网络模块114可使用由训练过程实施的操作参数来操作面部检测过程250,所述训练过程被开发为检测图像输入中的任何朝向的面部,如上所述。因此,面部检测过程250可检测图像输入252中在图像输入中具有任何朝向的面部,而无需旋转图像和/或接收可提供关于图像的朝向的信息的任何其他传感器输入数据(例如,加速度计或陀螺仪数据)。以任何朝向检测面也增加边界框中的姿态估计的范围。例如,在边界框中滚动估计从 $-180^{\circ}$ 到 $+180^{\circ}$ (所有滚动朝向)。

[0088] 在某些实施方案中,面部检测过程250检测图像输入252中的局部面部的存在。在图像输入252中检测到的局部面部可包括存在于图像输入252中的面部的任何部分。要检测的面部在图像输入252中需要存在的面部的量和/或神经网络模块114检测图像输入252中的局部面部的能力可取决于用于生成神经网络模块的操作参数和/或图像输入中可检测的面部特征的训练。

[0089] 图13示出了在图像输入252中检测到的局部面部的示例。如图13所示,面部检测过程250可将边界框274放置在图像输入252中的面部270周围。如面部270在图像输入252中仅为局部,边界框274与图像输入的帧之外的边界框的部分一起布置。面部检测过程250不具有检测图像输入252中的局部面部的能力,否则可能拒绝图像输入252,因为边界框274的一部分超过图像输入之外。面部检测过程250能够在图像输入252的局部面部上提供面部检测的能力增加了在更具挑战性的检测案例中的面部检测,并且在面部检测过程中提供总体改善。

[0090] 在某些实施方案中,面部检测过程250针对捕获为图像输入252的图像中的其他变化属性提供鲁棒的面部检测。例如,面部检测过程250可在面部和相机之间的宽范围的距离上提供面部检测。面部和相机之间的距离的变化可使得面部在图像输入252中更小或更大。在图像输入252中,较小或较大的面部可能更难以检测。然而,神经网络模块114可使用由如上所述的训练过程200实现的操作参数来操作,所述操作参数允许神经网络模块检测面部和相机之间的更大距离范围内的面部。

[0091] 在一些实施方案中,面部检测过程250可在图像输入252中的多种曝光范围内提供面部检测。在较暗的光照条件下,相机可增加曝光以提供图像输入252。增加的曝光可在图像输入252的数据中产生更多的噪声。在较亮的光照条件下,相机可使用较小的孔径时间捕获图像以避免过分曝光图像。由于孔径时间较小,图像输入252中可存在用于处理的更少数数据。然而,神经网络模块114可使用由如上所述的训练过程200实现的操作参数来操作,所述操作参数允许神经网络模块在更大的曝光范围(并因此在更大范围的光条件)上检测面部。

[0092] 在一些实施方案中,可在解码器过程258中解码的多个区域中检测到相同的面部。区域的置信度排名也可由解码器过程258来确定。如果在多个区域中检测到相同的面部,则可使用这些区域的置信度排名来确定对于所检测到的面部具有最高置信度的区域。然后可选择具有最高置信度的区域,以提供输出数据260(包括所检测到的面部的属性的值的附加数据)。

[0093] 当检测到在特征空间存在面部时,由解码器过程258生成的预测包括对所检测到的面部的属性的一个或多个值的评估(例如,确定)。对所检测到的面部的属性的值进行评

估可包括使用与待评估的属性相关联的分类参数(从训练过程200获取)在特征向量的解码期间对特征向量进行分类。在某些实施方案中,属性的评估值包括面部在图像中的位置(由边界框表示),面部的姿态以及面部和相机之间的距离。在某些实施方案中,面部的姿态包括面部的倾斜、偏转、和/或转动。属性的评估值可连同关于在图像输入252中存在面部和面部的边界框的决定一起被包括在输出数据260中。

[0094] 在某些实施方案中,输出数据260被提供至下游过程262。下游过程262可包括能够使用面部检测过程输出的设备100上的面部检测过程250的任何下游过程。下游过程262的示例包括但不限于附加图像信号处理和安全飞地处理诸如面部识别处理或注意力检测处理。在一些实施方案中,输出数据260中的一个或多个值用于控制设备100的一个或多个操作。在一些实施方案中,输出数据260中的距离值可用于控制从设备100上的相机102输出的光斑图案照明的操作。例如,输出数据260中的距离值可以用于为来自相机102的光斑图案照明输出确定密度(或密度设置),如Fasel、Guo、Kumar和Gernoth的名称为“DETERMINING SPARSE VERSUS DENSE PATTERN ILLUMINATION”的共同未决美国临时专利申请No.62/556,832所述,该专利申请通过引用并入本文,如同在本文中完全阐述的一般。

[0095] 如图12所示,面部检测过程250可用于检测由设备100上的相机102捕获的图像中的一个或多个面部。输出数据260可包括在所捕获的图像中存在面部的决定、以及所检测到的面部的属性的值(例如,位置、姿态、和/或距相机的距离)的数据。面部检测过程250利用单个网络模块(例如,神经网络模块114)来提供面部检测输出、以及所检测到的面部的位置、姿态和距相机的距离。如本文所述,神经网络模块114实现提供改善和更鲁棒的面部检测的操作参数。例如,神经网络模块114可为捕获图像中任何朝向的面部和/或捕获图像中仅存在的局部面部提供面部检测。另外,神经网络模块114可提供面部检测,该面部检测鲁棒并且适应局部面部和整个面部情况下所捕获图像的属性的变化。例如,面部检测过程250对于相机与用户面部之间的曝光和/或距离的变化可能是鲁棒的。在一些实施方案中,面部检测过程250的曝光极限可通过用来捕获图像的相机的曝光的极限来设置(例如,在设备100上的相机102)。因此,面部检测过程250可以能够形成捕获的图像中的边界框,具有由照明器提供的任何照明(例如,泛光照明器105A和/或散斑照明器105B)。

[0096] 使用如本文所述的训练过程200来改善面部检测过程250可在面部检测过程中提供改善的误接率和改善的假排斥率。另外,面部检测过程可在捕获图像中提供对面部的更好的位置识别(对于局部面部或整个面部情况)。提供改进且更鲁棒的面部检测还可改善在面部检测之后实施的下游过程(例如,下游过程262)。例如,更鲁棒的面部检测,在面部检测的更具挑战性的情况下,可增大完成面部检测的可能性并且增大下游过程操作的情况的频率。

[0097] 在某些实施方案中,可由执行被存储在非暂态计算机可读介质上的指令的一个或多个处理器(例如,计算机处理器)来执行本文所述的一个或多个过程步骤。例如,图5中所示的过程200或图12中所示的过程250可以具有一个或多个由执行在计算机可读存储介质(例如,非暂态计算机可读存储介质)中被存储为程序指令的指令的一个或多个处理器执行的步骤。

[0098] 图14示出了示例性计算机系统510的一个实施方案的框图。示例性计算机系统510可用于实现本文所述的一个或多个实施方案。在一些实施方案中,计算机系统510可由用户

操作以实施本文描述的一个或多个实施方案,诸如图5中所示的过程200或图12中所示的过程250。在图14的实施方案中,计算机系统510包括处理器512、存储器514和各种外围设备516。处理器512耦接到存储器514和外围设备516。处理器512被配置为执行指令,该指令包括用于过程200或过程250的指令,其可以在软件中。在各种实施方案中,处理器512可实现任何期望的指令集(例如,Intel架构-32(IA-32,也被称为x86)、IA-32(具有64位扩展)、x86-64、PowerPC、Sparc、MIPS、ARM、IA-64等)。在一些实施方案中,计算机系统510可包括多于一个处理器。此外,处理器512可包括一个或多个处理器或一个或多个处理器内核。

[0099] 处理器512可以任何期望的方式而被耦接到存储器514和外围设备516。例如,在一些实施方案中,处理器512可经由各种互连件而被耦接到存储器514和/或外围设备516。另选地或除此之外,可使用一个或多个桥接芯片来耦接处理器512、存储器514和外围设备516。

[0100] 存储器514可包括任何类型的存储器系统。例如,存储器514可包括DRAM,并且更具体地包括双倍数据速率(DDR) SDRAM、RDRAM等。可包括存储器控制器以与存储器514接口连接,和/或处理器512可包括存储器控制器。存储器514可存储在使用期间由处理器512执行的指令、以及在使用期间由处理器操作的数据等。

[0101] 外围设备516可表示可被包括在计算机系统510中或与其耦接的任何种类的硬件设备(例如,存储设备,可选地包括图15所示的计算机可访问存储介质600、其他输入/输出(I/O)设备诸如视频硬件、音频硬件、用户界面设备、联网硬件等)。

[0102] 现在转向图15,其示出了计算机可访问存储介质600的一个实施方案的框图,该计算机可访问存储介质包括表示被包括在集成电路设计中的设备100(在图1中示出)的一个或多个数据结构、以及表示过程250(在图12中示出)的一个或多个代码序列。每个代码序列可包括在由计算机中的处理器执行时实现针对对应的代码序列所述的操作的一个或多个指令。一般来讲,计算机可访问存储介质可包括在使用期间可被计算机访问以向计算机提供指令和/或数据的任何存储介质。例如,计算机可访问存储介质可包括非暂态存储介质诸如磁性介质或光学介质,例如磁盘(固定的或可拆卸的)、磁带、CD-ROM、DVD-ROM、CD-R、CD-RW、DVD-R、DVD-RW、或蓝光。存储介质还可包括易失性或非易失性存储器介质,诸如RAM(例如,同步动态RAM(SDRAM)、Rambus DRAM(RDRAM)、静态RAM(SRAM)等)、ROM或闪存存储器。存储介质可被物理地包括在存储介质将指令/数据提供至其的计算机中。另选地,存储介质可连接至计算机。例如,存储介质可通过网络或无线链路诸如网络附接存储装置而被连接至计算机。存储介质可通过外围接口诸如通用串行总线(USB)而被连接。一般来讲,计算机可访问存储介质600可以非暂态方式存储数据,其中在该上下文中非暂态是指不通过信号传输指令/数据。例如,非暂态存储装置可为易失性的(并且响应于断电而可能会丢失所存储的指令/数据)或非易失性的。

[0103] 根据本说明书,对本公开所描述的实施方案的各个方面的其他修改和替代实施方案对于本领域的技术人员而言将为显而易见的。因此,将本说明书理解为仅为示例性的并且用于教导本领域的技术人员执行这些实施方案的一般方式。应当理解,本文所示和所述的实施方案的形式将被当作目前优选的实施方案。元素与材料可被本文所示和所述的那些元素与材料替代,可反转部件和工艺,并且可独立地利用这些实施方案的某些特征,在受益于本说明书之后,所有这些对于本领域的技术人员而言都将为显而易见的。可在不脱离以

下权利要求书的实质和范围的情况下对本文所述的元素作出修改。

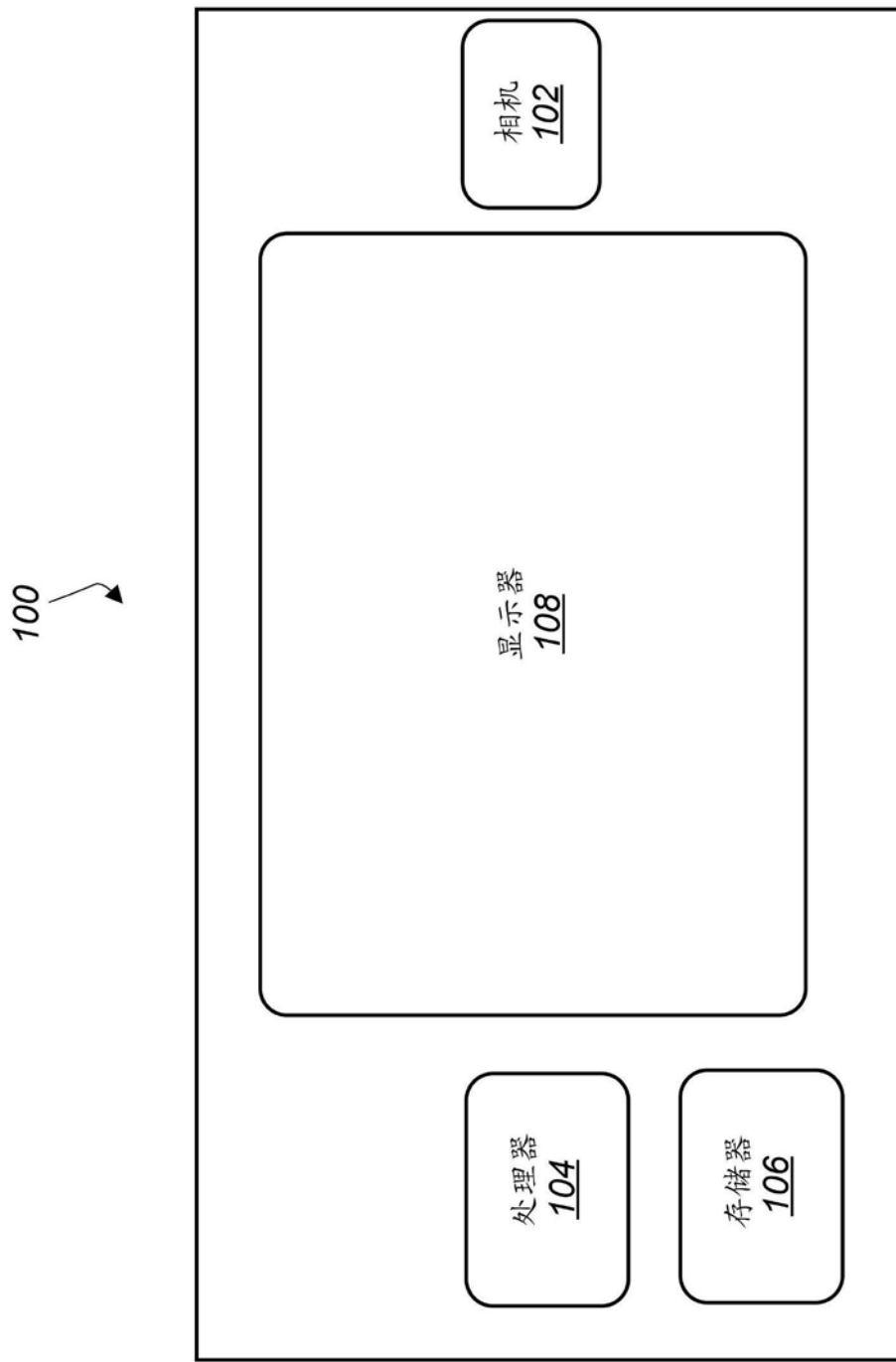


图1

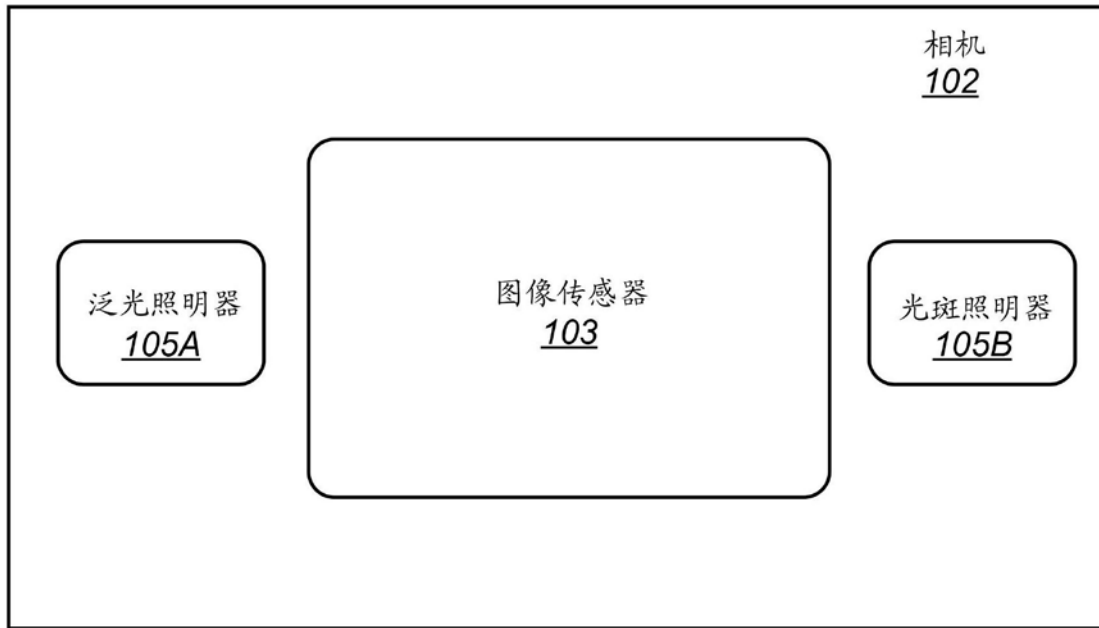


图2

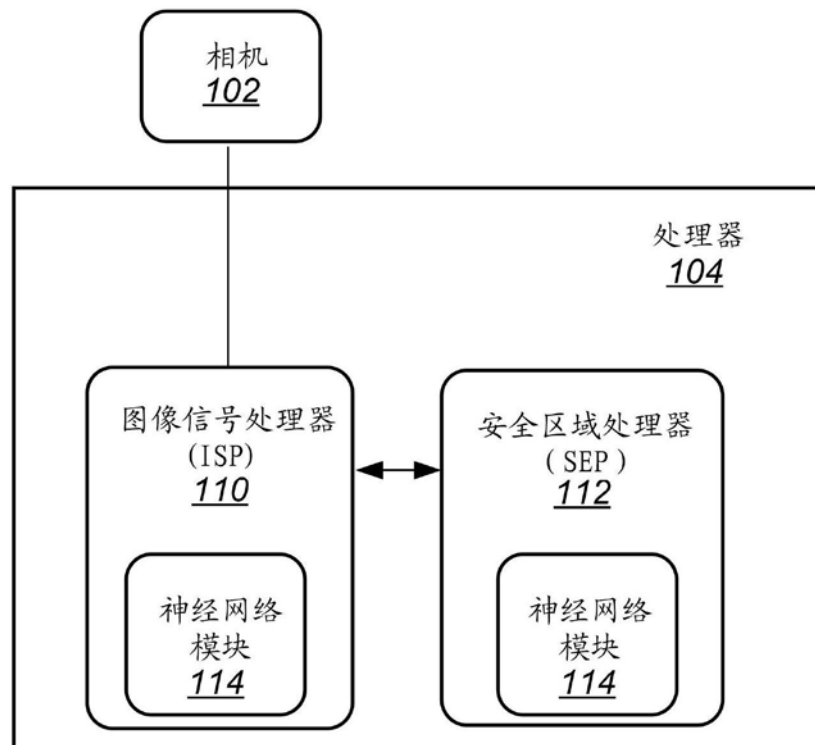


图3

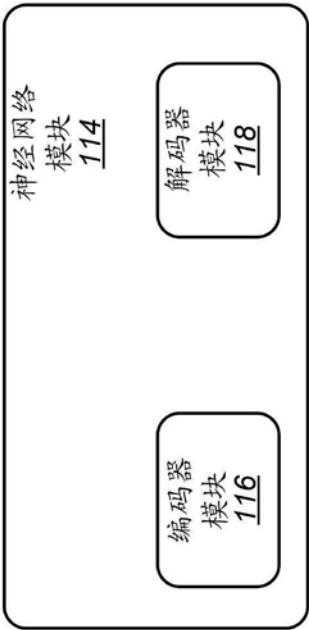


图4

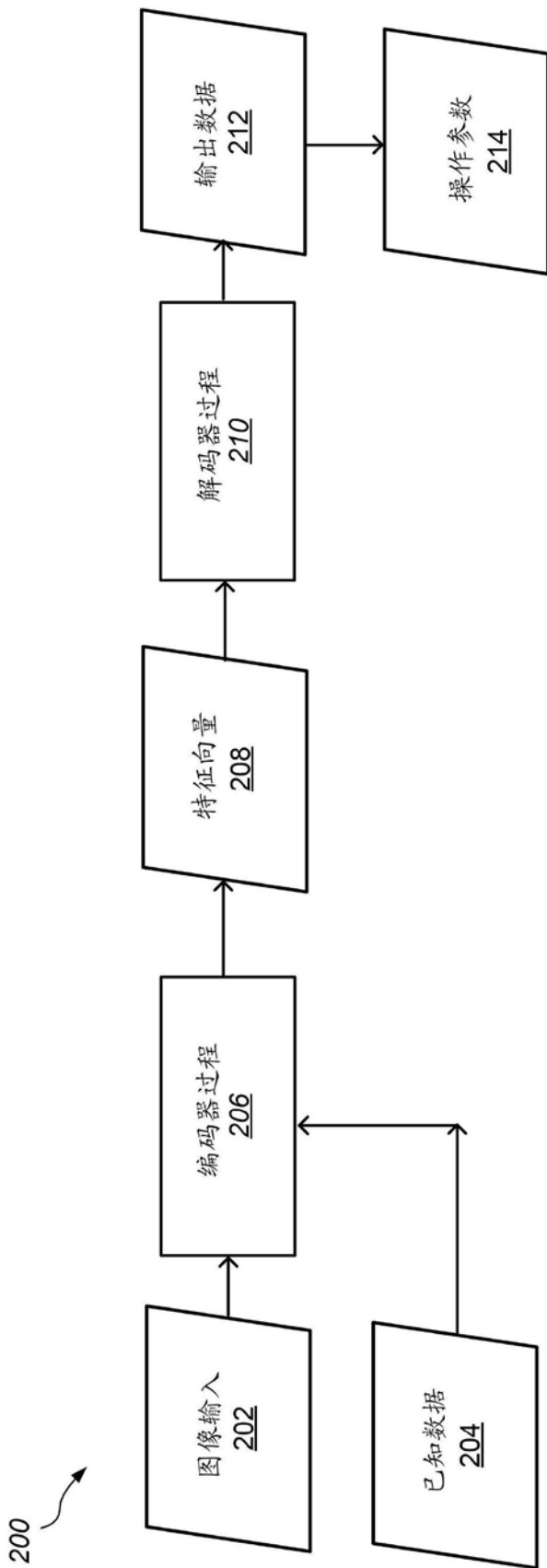


图5



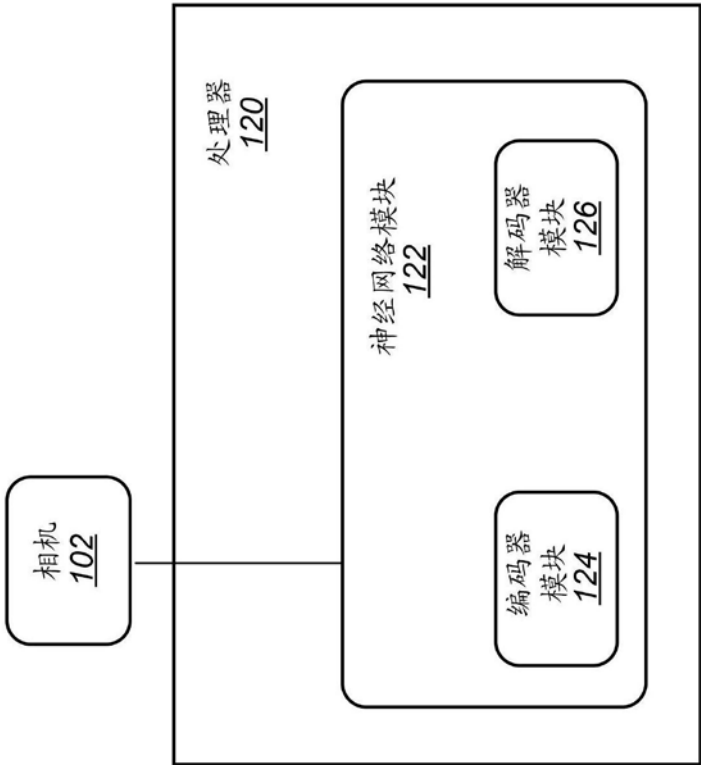


图6

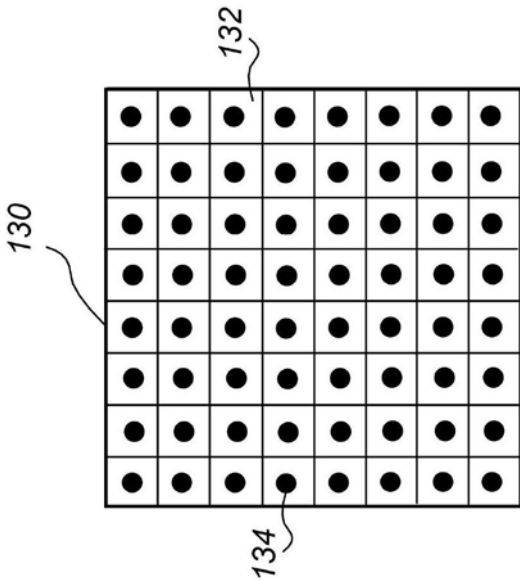


图9

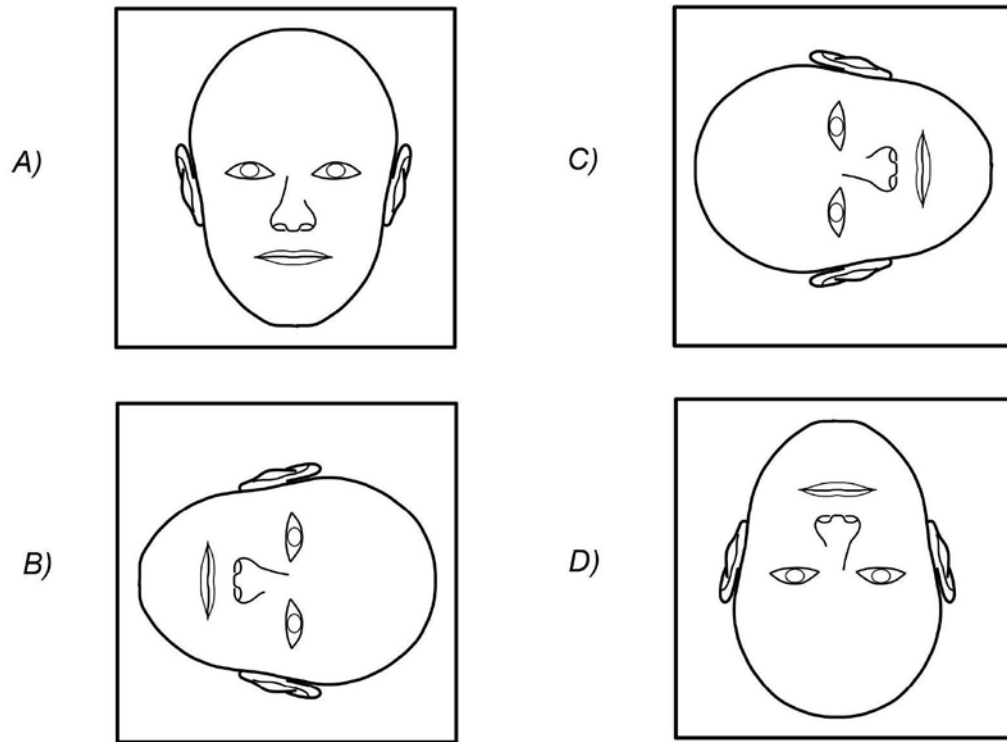


图7

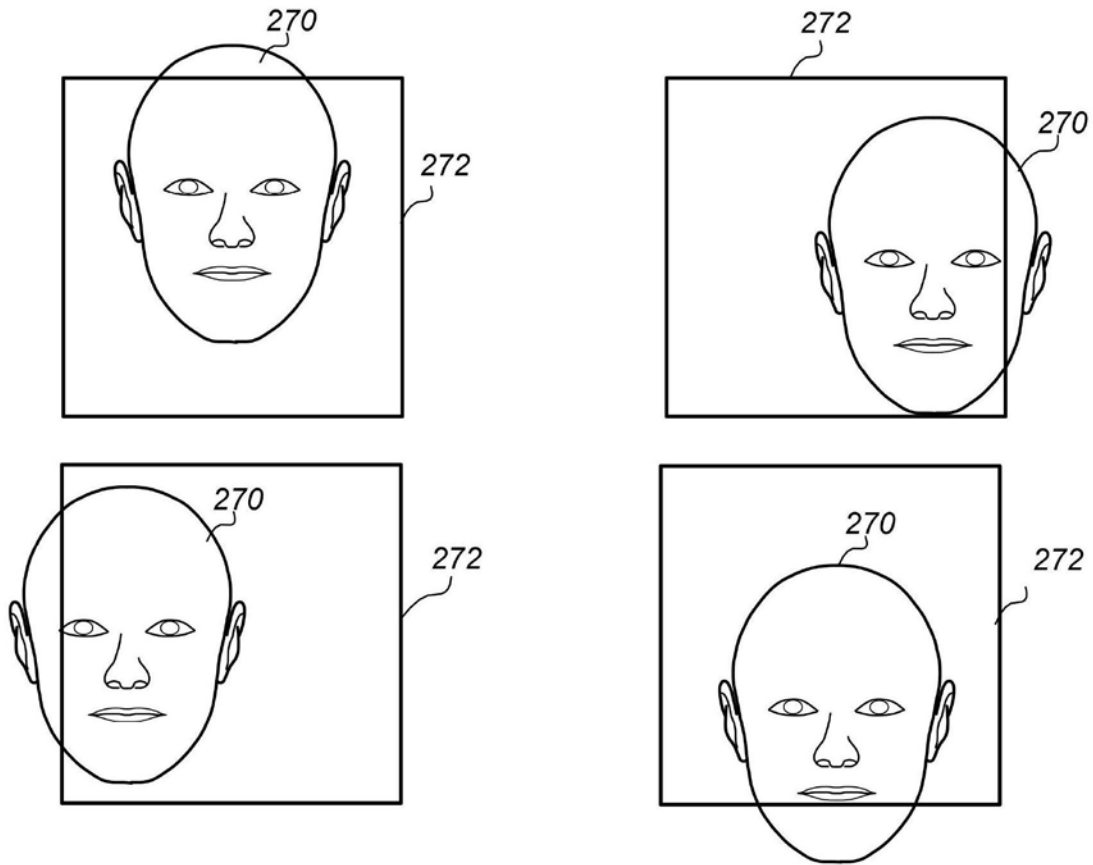


图8

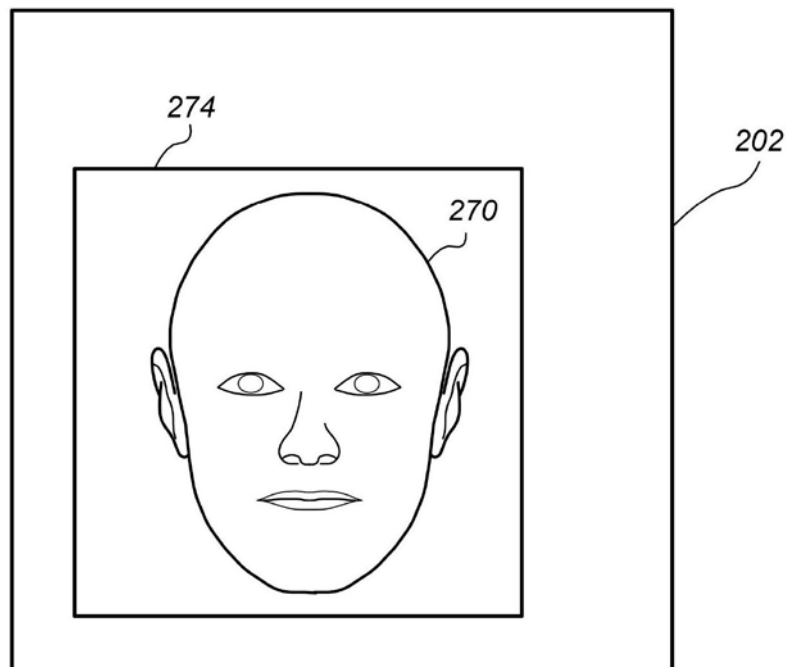


图10

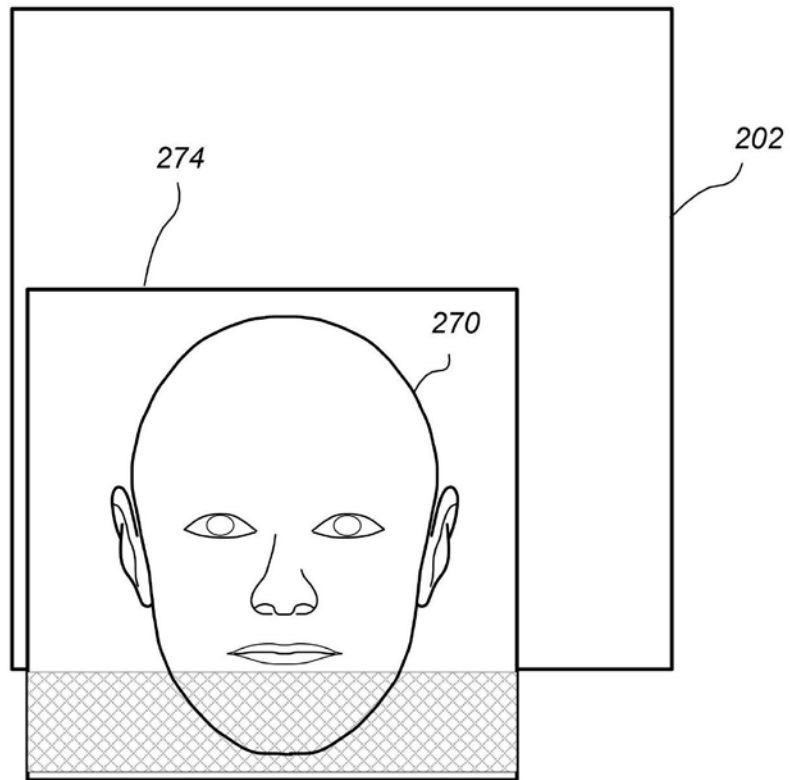


图13

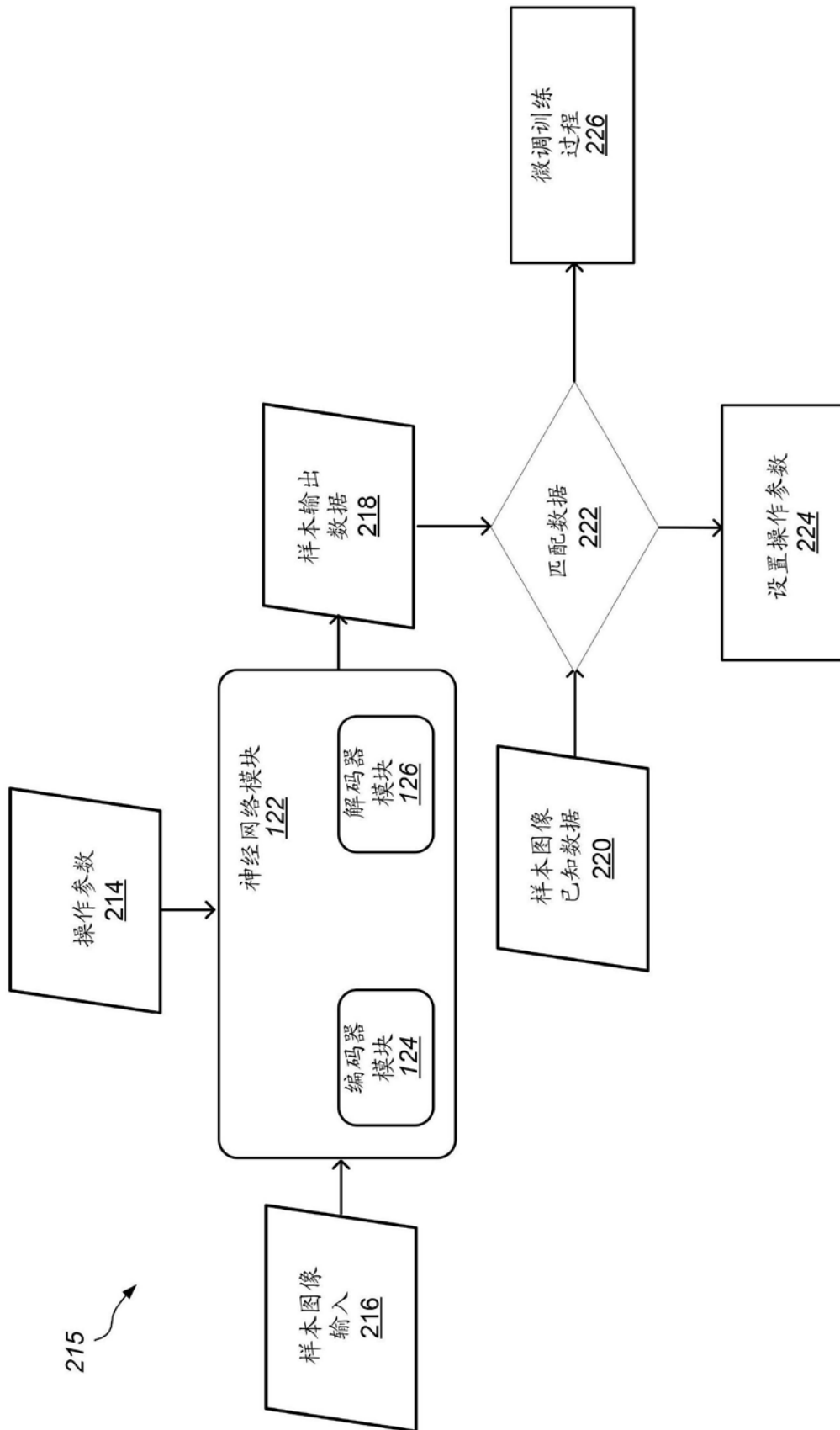


图11

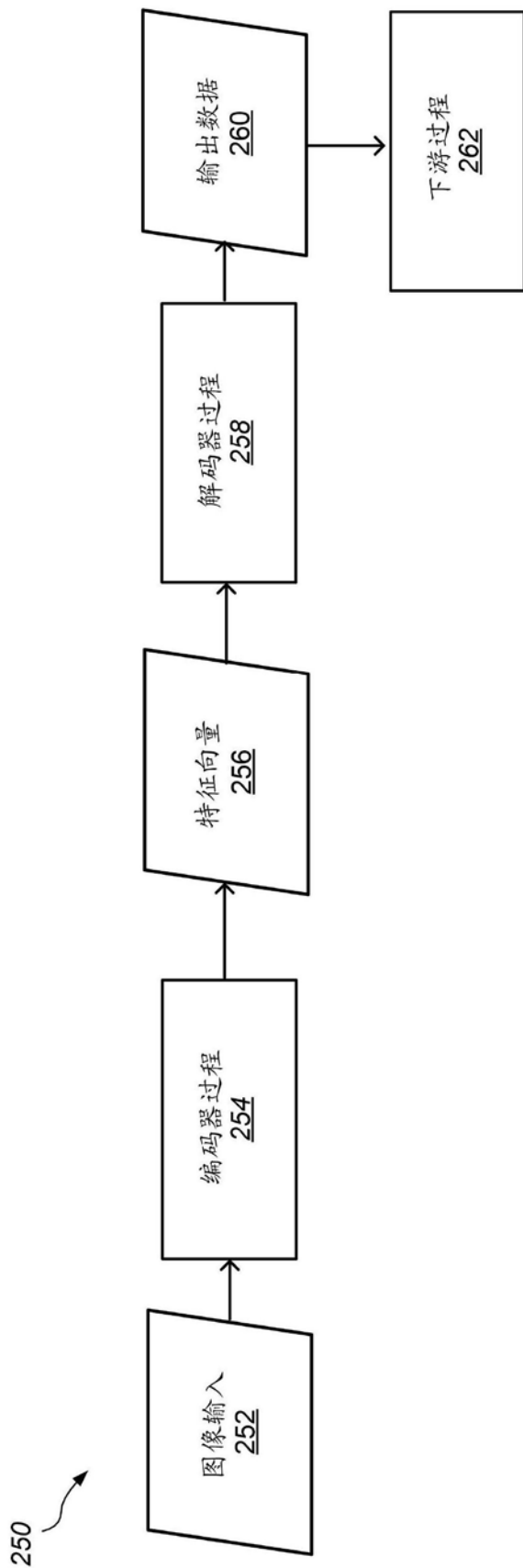


图12

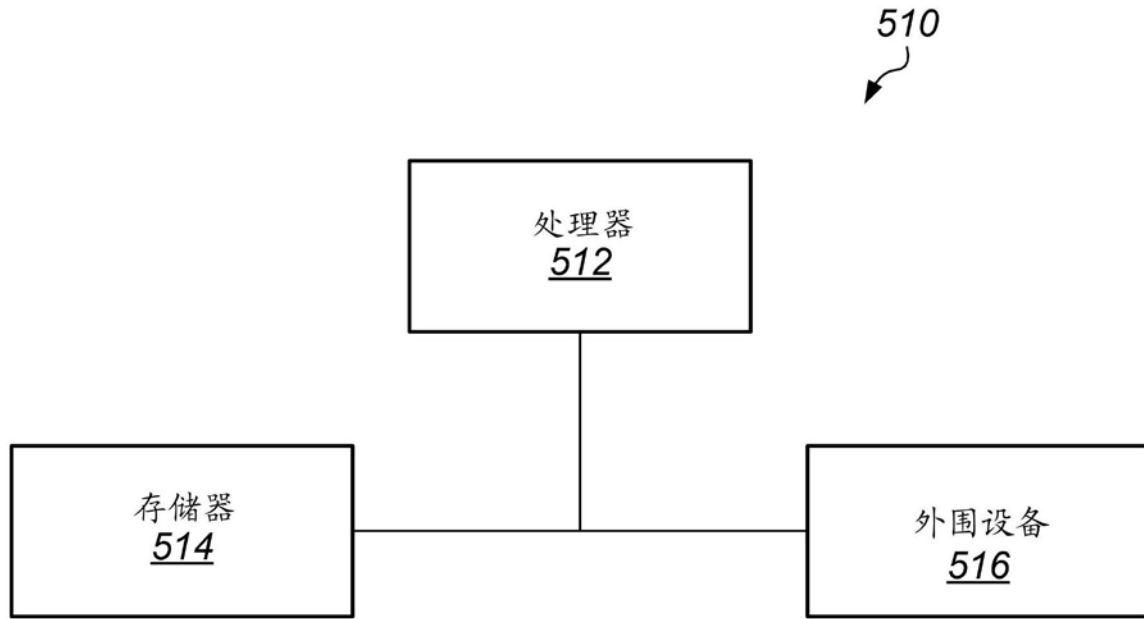


图14

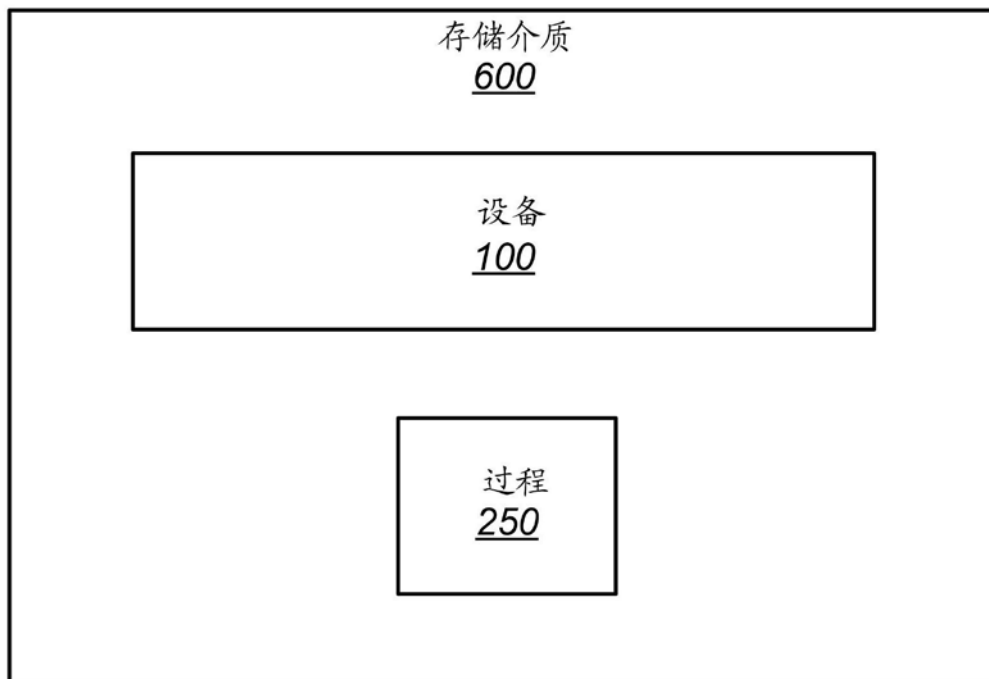


图15