



(12)发明专利申请

(10)申请公布号 CN 109034102 A

(43)申请公布日 2018.12.18

(21)申请号 201810924096.9

(22)申请日 2018.08.14

(71)申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72)发明人 吴双 丁守鸿 李季標 梁小龙
梁麟开 杨静

(74)专利代理机构 北京三高永信知识产权代理
有限责任公司 11138

代理人 张所明

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

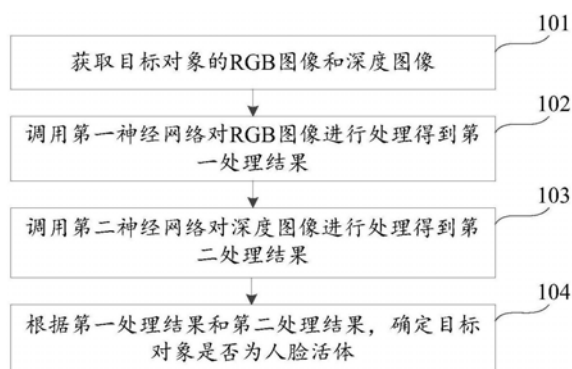
权利要求书3页 说明书18页 附图9页

(54)发明名称

人脸活体检测方法、装置、设备及存储介质

(57)摘要

本申请实施例公开了一种人脸活体检测方法、装置、设备及存储介质。所述方法包括：获取目标对象的RGB图像和深度图像；调用第一神经网络对RGB图像进行处理得到第一处理结果；调用第二神经网络对深度图像进行处理得到第二处理结果；根据第一处理结果和第二处理结果，确定目标对象是否为人脸活体。本申请实施例通过获取目标对象的RGB图像和深度图像，结合上述两个图像对目标对象进行人脸活体检测，融合了人脸纹理线索和3D人脸结构线索进行人脸活体检测，提高了准确性，可以有效拦截人脸纸片、高清屏幕人脸翻拍、合成人脸视频、面具、3D假体模型等攻击。并且，RGB图像和深度图像的拍摄过程中，无需用户执行交互动作，更为简单高效。



1. 一种人脸活体检测方法,其特征在于,所述方法包括:

获取目标对象的RGB图像和深度图像;

调用第一神经网络对所述RGB图像进行处理得到第一处理结果,所述第一处理结果是从所述RGB图像中提取的用于表征所述目标对象是否为人脸活体的分类信息;

调用第二神经网络对所述深度图像进行处理得到第二处理结果,所述第二处理结果是从所述深度图像中提取的用于表征所述目标对象是否为所述人脸活体的分类信息;

根据所述第一处理结果和所述第二处理结果,确定所述目标对象是否为所述人脸活体。

2. 根据权利要求1所述的方法,其特征在于,

所述第一处理结果为第一概率值,所述第一概率值是根据所述RGB图像所确定的所述目标对象为所述人脸活体的概率;

所述第二处理结果为第二概率值,所述第二概率值是根据所述深度图像所确定的所述目标对象为所述人脸活体的概率。

3. 根据权利要求2所述的方法,其特征在于,所述根据所述第一处理结果和所述第二处理结果,确定所述目标对象是否为所述人脸活体,包括:

若所述第一概率值大于第一阈值,且所述第二概率值大于第二阈值,则确定所述目标对象为所述人脸活体。

4. 根据权利要求2所述的方法,其特征在于,所述根据所述第一处理结果和所述第二处理结果,确定所述目标对象是否为所述人脸活体,包括:

调用第三神经网络对所述第一概率值和所述第二概率值进行处理,得到综合概率值;

根据所述综合概率值确定所述目标对象是否为所述人脸活体。

5. 根据权利要求1所述的方法,其特征在于,所述第一处理结果为第一输出特征,所述第二处理结果为第二输出特征;其中,所述第一输出特征是从所述RGB图像中提取的深度特征,所述第二输出特征是从所述深度图像中提取的深度特征;

所述根据所述第一处理结果和所述第二处理结果,确定所述目标对象是否为所述人脸活体,包括:

调用概率输出网络对所述第一输出特征和所述第二输出特征进行处理,得到概率值,所述概率值用于指示所述目标对象为所述人脸活体的概率;

根据所述概率值确定所述目标对象是否为所述人脸活体。

6. 根据权利要求1至5任一项所述的方法,其特征在于,

所述第一神经网络在对所述RGB图像进行处理的过程中,融合所述第二神经网络从所述深度图像中提取的特征,得到所述第一处理结果;

和/或,

所述第二神经网络在对所述深度图像进行处理的过程中,融合所述第一神经网络从所述RGB图像中提取的特征,得到所述第二处理结果。

7. 根据权利要求6所述的方法,其特征在于,所述人脸活体检测模型还包括:特征融合网络;

所述特征融合网络用于:

对所述第一神经网络的第i个隐藏层输出的特征和所述第二神经网络的第j个隐藏层

输出的特征进行特征数组加和处理和/或特征通道叠加处理,得到融合特征;其中,所述*i*为正整数,所述*j*为正整数;

将所述融合特征输入至所述第一神经网络的第*i*+1个隐藏层和/或将所述融合特征输入至所述第二神经网络的第*j*+1个隐藏层。

8. 根据权利要求1至5任一项所述的方法,其特征在于,所述获取目标对象的RGB图像和深度图像之后,还包括:

检测所述RGB图像中的人脸区域;

获取所述RGB图像中所述人脸区域中的RGB图像内容;

根据所述人脸区域在所述RGB图像中的位置,获取所述深度图像中所述位置处的深度图像内容;

其中,所述第一神经网络用于对所述RGB图像中的所述RGB图像内容进行处理得到所述第一处理结果,所述第二神经网络用于对所述深度图像中的所述深度图像内容进行处理得到所述第二处理结果。

9. 根据权利要求1至5任一项所述的方法,其特征在于,所述方法还包括:

获取所述目标对象的红外图像;

调用第四神经网络对所述红外图像进行处理得到第三处理结果,所述第三处理结果是从所述红外图像中提取的用于表征所述目标对象是否为所述人脸活体的分类信息;

所述根据所述第一处理结果和所述第二处理结果,确定所述目标对象是否为所述人脸活体,包括:

根据所述第一处理结果、所述第二处理结果和所述第三处理结果,确定所述目标对象是否为所述人脸活体。

10. 一种人脸活体检测方法,其特征在于,所述方法包括:

终端通过摄像头采集目标对象的RGB图像和深度图像;

所述终端调用第二神经网络对所述深度图像进行处理得到第二概率值,所述第二概率值是根据所述深度图像所确定的所述目标对象为所述人脸活体的概率;

若所述第二概率值大于第二阈值,则所述终端向服务器发送所述RGB图像;

所述服务器调用第一神经网络对所述RGB图像进行处理得到第一概率值,所述第一概率值是根据所述RGB图像所确定的所述目标对象为所述人脸活体的概率;

若所述第一概率值大于第一阈值,则所述服务器确定所述目标对象为所述人脸活体。

11. 一种人脸活体检测模型的训练方法,其特征在于,所述方法包括:

获取训练样本,所述训练样本包括:样本对象的RGB图像和深度图像,以及所述样本对象对应的标签,所述标签用于指示所述样本对象是否为所述人脸活体;

通过所述人脸活体检测模型对所述样本对象的RGB图像和深度图像进行处理,得到第一处理结果和第二处理结果;其中,所述人脸活体检测模型包括第一神经网络和第二神经网络,所述第一神经网络用于对所述样本对象的RGB图像进行处理得到所述第一处理结果,所述第二神经网络用于对所述样本对象的深度图像进行处理得到所述第二处理结果;

根据所述第一处理结果和第二处理结果,预测所述样本对象是否为所述人脸活体;

根据所述样本对象的预测结果和所述标签对所述人脸活体检测模型的参数进行修正,直至所述人脸活体检测模型满足停止训练条件时,结束对所述人脸活体检测模型的训练。

12. 一种人脸活体检测装置,其特征在于,所述装置包括:

图像获取模块,用于获取目标对象的RGB图像和深度图像;

模型调用模块,用于调用第一神经网络对所述RGB图像进行处理得到第一处理结果,所述第一处理结果是从所述RGB图像中提取的用于表征所述目标对象是否为人脸活体的分类信息;以及,调用第二神经网络对所述深度图像进行处理得到第二处理结果,所述第二处理结果是从所述深度图像中提取的用于表征所述目标对象是否为所述人脸活体的分类信息;

活体检测模块,用于根据所述第一处理结果和所述第二处理结果,确定所述目标对象是否为所述人脸活体。

13. 一种人脸活体检测模型的训练装置,其特征在于,所述装置包括:

样本获取模块,用于获取训练样本,所述训练样本包括:样本对象的RGB图像和深度图像,以及所述样本对象对应的标签,所述标签用于指示所述样本对象是否为人脸活体;

概率计算模块,用于通过所述人脸活体检测模型对所述样本对象的RGB图像和深度图像进行处理,得到第一处理结果和第二处理结果;其中,所述人脸活体检测模型包括第一神经网络和第二神经网络,所述第一神经网络用于对所述样本对象的RGB图像进行处理得到所述第一处理结果,所述第二神经网络用于对所述样本对象的深度图像进行处理得到所述第二处理结果;

活体预测模块,用于根据所述第一处理结果和第二处理结果,预测所述样本对象是否为所述人脸活体;

模型训练模块,用于根据所述样本对象的预测结果和所述标签对所述人脸活体检测模型的参数进行修正,直至所述人脸活体检测模型满足停止训练条件时,结束对所述人脸活体检测模型的训练。

14. 一种计算机设备,其特征在于,所述计算机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、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现如权利要求1至9任一项所述的人脸活体检测方法,或者实现如权利要求11所述的人脸活体检测模型的训练方法。

15. 一种计算机可读存储介质,其特征在于,所述存储介质中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由处理器加载并执行以实现如权利要求1至9任一项所述的人脸活体检测方法,或者实现如权利要求11所述的人脸活体检测模型的训练方法。

人脸活体检测方法、装置、设备及存储介质

技术领域

[0001] 本申请实施例涉及图像识别技术领域，特别涉及一种人脸活体检测方法、装置、设备及存储介质。

背景技术

[0002] 人脸识别技术已经广泛应用，然而人脸极易用照片、视频等方式进行复制，因此需要对人脸进行活体检测。

[0003] 在相关技术中，提供了一种基于动作指令的人脸活体检测方法。终端实时给出动作指令，如眨眼、左右摇头、点头、张嘴等指令，用户根据上述动作指令执行相应动作。终端通常依次给出多个动作指令，若执行各个动作指令全部正确则认为是人脸活体；反之，若任一动作指令执行错误则认为不是人脸活体。

[0004] 上述基于动作指令的人脸活体检测方法，需要用户按照动作指令依次执行多个动作，完成人脸活体检测的效率较低，且对于合成视频、高清屏幕人脸翻拍、3D面具等攻击的拦截能力较弱。

发明内容

[0005] 本申请实施例提供了一种人脸活体检测方法、装置、设备及存储介质，可用于解决基于动作指令的人脸活体检测方法，完成人脸活体检测的效率较低，且对于合成视频、高清屏幕人脸翻拍、3D面具等攻击的拦截能力较弱的问题。技术方案如下：

[0006] 一方面，本申请实施例提供一种人脸活体检测方法，所述方法包括：

[0007] 获取目标对象的RGB图像和深度图像；

[0008] 调用第一神经网络对所述RGB图像进行处理得到第一处理结果，所述第一处理结果是从所述RGB图像中提取的用于表征所述目标对象是否为人脸活体的分类信息；

[0009] 调用第二神经网络对所述深度图像进行处理得到第二处理结果，所述第二处理结果是从所述深度图像中提取的用于表征所述目标对象是否为所述人脸活体的分类信息；

[0010] 根据所述第一处理结果和所述第二处理结果，确定所述目标对象是否为所述人脸活体。

[0011] 另一方面，本申请实施例提供一种人脸活体检测方法，所述方法包括：

[0012] 终端通过摄像头采集目标对象的RGB图像和深度图像；

[0013] 所述终端调用第二神经网络对所述深度图像进行处理得到第二概率值，所述第二概率值是根据所述深度图像所确定的所述目标对象为所述人脸活体的概率；

[0014] 若所述第二概率值大于第二阈值，则所述终端向服务器发送所述RGB图像；

[0015] 所述服务器调用第一神经网络对所述RGB图像进行处理得到第一概率值，所述第一概率值是根据所述RGB图像所确定的所述目标对象为所述人脸活体的概率；

[0016] 若所述第一概率值大于第一阈值，则所述服务器确定所述目标对象为所述人脸活体。

- [0017] 再一方面,本申请实施例提供一种人脸活体检测模型的训练方法,所述方法包括:
- [0018] 获取训练样本,所述训练样本包括:样本对象的RGB图像和深度图像,以及所述样本对象对应的标签,所述标签用于指示所述样本对象是否为人脸活体;
- [0019] 通过所述人脸活体检测模型对所述样本对象的RGB图像和深度图像进行处理,得到第一处理结果和第二处理结果;其中,所述人脸活体检测模型包括第一神经网络和第二神经网络,所述第一神经网络用于对所述样本对象的RGB图像进行处理得到所述第一处理结果,所述第二神经网络用于对所述样本对象的深度图像进行处理得到所述第二处理结果;
- [0020] 根据所述第一处理结果和第二处理结果,预测所述样本对象是否为所述人脸活体;
- [0021] 根据所述样本对象的预测结果和所述标签对所述人脸活体检测模型的参数进行修正,直至所述人脸活体检测模型满足停止训练条件时,结束对所述人脸活体检测模型的训练。
- [0022] 再一方面,本申请实施例提供一种人脸活体检测装置,所述装置包括:
- [0023] 图像获取模块,用于获取目标对象的RGB图像和深度图像;
- [0024] 模型调用模块,用于调用第一神经网络对所述RGB图像进行处理得到第一处理结果,所述第一处理结果是从所述RGB图像中提取的用于表征所述目标对象是否为人脸活体的分类信息;以及,调用第二神经网络对所述深度图像进行处理得到第二处理结果,所述第二处理结果是从所述深度图像中提取的用于表征所述目标对象是否为所述人脸活体的分类信息;
- [0025] 活体检测模块,用于根据所述第一处理结果和所述第二处理结果,确定所述目标对象是否为所述人脸活体。
- [0026] 再一方面,本申请实施例提供一种人脸活体检测模型的训练装置,所述装置包括:
- [0027] 样本获取模块,用于获取训练样本,所述训练样本包括:样本对象的RGB图像和深度图像,以及所述样本对象对应的标签,所述标签用于指示所述样本对象是否为人脸活体;
- [0028] 概率计算模块,用于通过所述人脸活体检测模型对所述样本对象的RGB图像和深度图像进行处理,得到第一处理结果和第二处理结果;其中,所述人脸活体检测模型包括第一神经网络和第二神经网络,所述第一神经网络用于对所述样本对象的RGB图像进行处理得到所述第一处理结果,所述第二神经网络用于对所述样本对象的深度图像进行处理得到所述第二处理结果;
- [0029] 活体预测模块,用于根据所述第一处理结果和所述第二处理结果,预测所述样本对象是否为所述人脸活体;
- [0030] 模型训练模块,用于根据所述样本对象的预测结果和所述标签对所述人脸活体检测模型的参数进行修正,直至所述人脸活体检测模型满足停止训练条件时,结束对所述人脸活体检测模型的训练。
- [0031] 又一方面,本申请实施例提供一种计算机设备,所述计算机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、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现上述方面所述的人脸活体检测方法,或者实现上述方面所述的人脸活体检测模型的训练方法。

[0032] 还一方面,本申请实施例提供一种计算机可读存储介质,所述存储介质中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由处理器加载并执行以实现上述方面所述的人脸活体检测方法,或者实现上述方面所述的人脸活体检测模型的训练方法。

[0033] 还一方面,本申请实施例提供一种计算机程序产品,当该计算机程序产品被执行时,其用于执行上述方面所述的人脸活体检测方法,或者实现上述方面所述的人脸活体检测模型的训练方法。

[0034] 本申请实施例提供的技术方案中,通过获取目标对象的RGB图像和深度图像,结合上述两个图像对目标对象进行人脸活体检测,融合了人脸纹理线索和3D人脸结构线索进行人脸活体检测,提高了检测的准确性,可以有效拦截人脸纸片、高清屏幕人脸翻拍、合成人脸视频、面具、3D假体模型等攻击。并且,RGB图像和深度图像的拍摄过程中,无需用户执行交互动作,更为简单高效,且降低交互成本,提高用户体验。

附图说明

[0035] 图1是本申请一个实施例提供的人脸活体检测方法的流程图;

[0036] 图2是本申请另一个实施例提供的人脸活体检测方法的流程图;

[0037] 图3示例性示出了特征融合的示意图;

[0038] 图4是本申请另一个实施例提供的人脸活体检测方法的流程图;

[0039] 图5示例性示出了图4实施例对应的模型架构的示意图;

[0040] 图6是本申请另一个实施例提供的人脸活体检测方法的流程图;

[0041] 图7是图6实施例对应的处理流程图;

[0042] 图8是本申请另一个实施例提供的人脸活体检测方法的流程图;

[0043] 图9是本申请另一个实施例提供的人脸活体检测方法的流程图;

[0044] 图10是本申请一个实施例提供的人脸活体检测装置的框图;

[0045] 图11是本申请一个实施例提供的人脸活体检测模型的训练装置的框图;

[0046] 图12是本申请一个实施例提供的计算机设备的结构框图。

具体实施方式

[0047] 下面将结合附图对本申请实施方式作进一步地详细描述。

[0048] 请参考图1,其示出了本申请一个实施例提供的人脸活体检测方法的流程图。该方法可应用于计算机设备中,所述计算机设备可以是诸如手机、平板电脑、可穿戴设备、PC(Personal Computer,个人计算机)等终端,也可以是用于为终端提供后台服务的服务器,或者是其它具备计算和处理能力的电子设备。

[0049] 该方法可以包括如下几个步骤:

[0050] 步骤101,获取目标对象的RGB图像和深度图像。

[0051] 目标对象的RGB图像是指通过摄像头对目标对象进行拍摄得到的彩色图像。RGB图像的每个像素值采用R(红)、G(绿)、B(蓝)三个通道的颜色分量来表示。

[0052] 目标对象的深度图像是指将从摄像头到目标对象所处场景中各点的距离值作为像素值的图像。深度图像也称为距离影像。深度图像直接反映了物体可见表面的几何形状。

深度图像类似于灰度图像,只是深度图像的每个像素值是摄像头距离物体的实际距离。目标对象的RGB图像和深度图像可以是配准的,两者的像素点之间具有一对一的对应关系。

[0053] 目标对象的RGB图像和深度图像可以是采用3D摄像头拍摄得到的两张图像,相较于普通2D摄像头仅能够拍摄RGB图像,3D摄像头除了能够拍摄RGB图像之外,还能够拍摄深度图像。其中,3D摄像头可以包括彩色摄像头和深度传感器,彩色摄像头用于拍摄RGB图像,深度传感器用于采集深度信息以生成深度图像。

[0054] 步骤102,调用第一神经网络对RGB图像进行处理得到第一处理结果。

[0055] 步骤103,调用第二神经网络对深度图像进行处理得到第二处理结果。

[0056] 在本申请实施例中,人脸活体检测模型是采用机器学习算法对神经网络进行训练得到的模型,人脸活体检测模型包括第一神经网络和第二神经网络。第一处理结果是由第一神经网络从RGB图像中提取的,用于表征目标对象是否为人脸活体的分类信息。第二处理结果是由第二神经网络从深度图像中提取的,用于表征目标对象是否为人脸活体的分类信息。

[0057] 第一神经网络可以是深度神经网络(Deep Neural Network,DNN),如卷积神经网络(Convolutional Neural Network,CNN)。同样地,第二神经网络也可以是DNN,如CNN。

[0058] 通过采用训练样本对人脸活体检测模型中的上述两个神经网络进行训练,即可得到用于实现上述功能的两个神经网络。

[0059] 有关第一神经网络和第二神经网络的具体结构,以及人脸活体检测模型的训练过程,可参见下文实施例中的介绍说明。

[0060] 步骤104,根据第一处理结果和第二处理结果,确定目标对象是否为人脸活体。

[0061] 例如,计算机设备可以根据第一处理结果和第二处理结果得到概率值,该概率值用于指示目标对象为人脸活体的概率。可选地,概率值的取值范围是 $[0,1]$ 。概率值越大,表明目标对象为人脸活体的可能性越高;反之,概率值越小,表明目标对象是人脸活体的可能性越低。计算机设备可以将概率值和预设阈值进行比对,当概率值大于预设阈值时,确定目标对象是人脸活体,当概率值小于预设阈值时,确定目标对象不是人脸活体。

[0062] 针对不同的处理结果,根据处理结果确定目标对象是否为人脸活体的方式也有所不同,具体可参见下文实施例中的介绍说明。

[0063] 需要说明的一点是,在本申请实施例中,对步骤102和步骤103的执行先后顺序不作限定,步骤103可以在步骤102之后执行,也可以在步骤102之前执行,或者与步骤102同时执行。

[0064] 综上所述,本申请实施例提供的技术方案中,通过获取目标对象的RGB图像和深度图像,结合上述两个图像对目标对象进行人脸活体检测,融合了人脸纹理线索和3D人脸结构线索进行人脸活体检测,提高了检测的准确性,可以有效拦截人脸纸片、高清屏幕人脸翻拍、合成人脸视频、面具、3D假体模型等攻击。并且,RGB图像和深度图像的拍摄过程中,无需用户执行交互动作,更为简单高效,且降低交互成本,提高用户体验。

[0065] 另外,在本申请实施例中,采用神经网络从图像中提取特征进行人脸活体检测,整个过程无需人工制定所要提取的图像特征,因此方案更具鲁棒性,且有助于提升人脸活体检测的准确度。

[0066] 请参考图2,其示出了本申请另一个实施例提供的人脸活体检测方法的流程图。该

方法可应用于上文介绍的计算机设备中。该方法可以包括如下几个步骤：

[0067] 步骤201,获取目标对象的RGB图像和深度图像。

[0068] 本步骤和图1实施例中的步骤101相同,可参见图1实施例中的介绍说明,本实施例对此不再赘述。

[0069] 步骤202,调用第一神经网络对RGB图像进行处理,得到第一概率值。

[0070] 第一概率值是根据RGB图像所确定的目标对象为人脸活体的概率。可选地,第一概率值的取值范围是 $[0,1]$ 。第一概率值越大,表明目标对象为人脸活体的可能性越高;反之,第一概率值越小,表明目标对象是人脸活体的可能性越低。

[0071] 第一神经网络可以包括输入层、隐藏层和输出层。输入层用于输入RGB图像,输出层用于输出第一概率值,隐藏层用于对RGB图像进行处理得到上述第一概率值。

[0072] 可选地,第一神经网络为深度神经网络(Deep Neural Network,DNN)。例如,第一神经网络为卷积神经网络(Convolutional Neural Network,CNN)。卷积神经网络可以包括输入层、卷积层(convolutional layer)、池化层(pooling layer)、全连接层(fully connected layer)和输出层。其中,输入层用于输入RGB图像,卷积层用于将RGB图像转化为特征图(feature map),池化层用于对特征图进行池化处理,以减少特征图中的特征数量,全连接层用于将经过卷积层和池化层处理后的特征图映射为一维特征向量,输出层用于根据全连接层输出的一维特征向量,输出第一概率值。卷积神经网络的主干网络可以采用AlexNet、VGG16、GoogleNet、ResNet、DenseNet、MobileNets等网络模型。卷积神经网络的输出层可以由分类器实现,用于实现二分类任务,输出是人脸活体以及不是人脸活体的概率。例如,上述分类器可以是Softmax分类器。可选地,卷积神经网络还可以包括批规范化层(batch normalization,BN)和ReLU(Rectified Linear Unit,修正线性单元)激活层。

[0073] 步骤203,调用第二神经网络对深度图像进行处理,得到第二概率值。

[0074] 第二概率值是根据深度图像所确定的目标对象为人脸活体的概率。可选地,第二概率值的取值范围是 $[0,1]$ 。第二概率值越大,表明目标对象为人脸活体的可能性越高;反之,第二概率值越小,表明目标对象是人脸活体的可能性越低。

[0075] 第二神经网络可以包括输入层、隐藏层和输出层。输入层用于输入深度图像,输出层用于输出第二概率值,隐藏层用于对深度图像进行处理得到上述第二概率值。与第一神经网络相类似,第二神经网络也可以是深度神经网络,如卷积神经网络。关于卷积神经网络的网络结构可参见上文介绍说明,此处不再赘述。

[0076] 第一神经网络的结构和第二神经网络的结构可以相同,也可以不同,本申请实施例对此不作限定。例如,第一神经网络和第二神经网络所包含的隐藏层数量、隐藏层中的神经元数量可以有所不同。在模型训练过程中,采用训练样本的RGB图像对神经网络进行训练,即可得到适用于处理RGB图像的第一神经网络;类似地,采用训练样本的深度图像对神经网络进行训练,即可得到适用于处理深度图像的第二神经网络。

[0077] 可选地,第一神经网络在对RGB图像进行处理的过程中,融合第二神经网络从深度图像中提取的特征,得到第一概率值;和/或,第二神经网络在对深度图像进行处理的过程中,融合第一神经网络从RGB图像中提取的特征,得到第二概率值。

[0078] 在预测第一概率值时,融合第一神经网络从RGB图像中提取的低维深度特征和第二神经网络从深度图像中提取的低维深度特征,使得第一神经网络具备融合RGB线索和深

度线索进行第一概率值预测的能力,有助于提升鲁棒性。同样地,在预测第二概率值时,融合第二神经网络从深度图像中提取的低维深度特征和第一神经网络从RGB图像中提取的低维深度特征,使得第二神经网络具备融合深度线索和RGB线索进行第二概率值预测的能力,也有助于提升鲁棒性。

[0079] 在本申请实施例中,对于特征融合(feature fusion)所采用的方式不作限定。例如,可以把第一神经网络和第二神经网络的隐藏层输出的特征数组进行加和,也可以在特征通道上进行叠加等。

[0080] 在一种可能的实施方式中,本实施例提供的人脸活体检测模型除了包括第一神经网络和第二神经网络之外,还包括特征融合网络。该特征融合网络用于:对第一神经网络的第*i*个隐藏层输出的特征和第二神经网络的第*j*个隐藏层输出的特征进行特征数组加和处理和/或特征通道叠加处理,得到融合特征;将融合特征输入至第一神经网络的第*i*+1个隐藏层和/或将融合特征输入至第二神经网络的第*j*+1个隐藏层;其中,*i*为正整数,*j*为正整数。

[0081] 其中,特征数组加和处理是指将两个神经网络输出的两个特征数组的相同位置处的数组元素相加,得到一个加和后的特征数组。特征通道叠加处理是指将两个神经网络输出的两个特征数组进行拼接,得到一个拼接后的特征数组。特征融合网络在进行特征数组加和处理和/或特征通道叠加处理之前,可以先对两个神经网络输出的两个特征数组进行规范化处理,使得两个特征数组具有相同的维度,以提示特征融合的效果。

[0082] 在本申请实施例中,对于特征融合网络所插入的位置不作限定,其可以将第一神经网络的任意一个隐藏层输出的特征和第二神经网络的任意一个隐藏层输出的特征进行特征融合处理。另外,也可以在多个不同的位置插入多个特征融合网络。在实际应用中,可以根据经验和模型训练效果,设定和调整特征融合网络所插入的位置和数量。

[0083] 请参考图3,其示例性示出了特征融合的示意图。以第一神经网络和第二神经网络均为卷积神经网络为例,可以在卷积层、池化层或者全连接层中的任意一个或多个层进行特征融合。

[0084] 步骤204,根据第一概率值和第二概率值,确定目标对象是否为人脸活体。

[0085] 在得到第一概率值和第二概率值之后,可以综合这两个概率值,最终确定目标对象是否为人脸活体。

[0086] 在一种可能的实施方式中,若第一概率值大于第一阈值,且第二概率值大于第二阈值,则确定目标对象为人脸活体;若第一概率值小于第一阈值和/或第二概率值小于第二阈值,则确定目标对象不是人脸活体。其中,第一阈值和第二阈值可以根据实际需求预先设定,本申请实施例对此不作限定,例如第一阈值为0.5且第二阈值也为0.5。

[0087] 在另一种可能的实施方式中,调用第三神经网络对第一概率值和第二概率值进行处理,得到综合概率值;根据综合概率值确定目标对象是否为人脸活体。第三神经网络可以包括输入层、隐藏层和输出层。输入层用于输入第一概率值和第二概率值,输出层用于输出综合概率值,隐藏层用于对第一概率值和第二概率值进行处理得到上述综合概率值。综合概率值是综合第一概率值和第二概率值之后,所确定的目标对象为人脸活体的概率。可选地,综合概率值的取值范围是[0,1]。综合概率值越大,表明目标对象为人脸活体的可能性越高;反之,综合概率值越小,表明目标对象是人脸活体的可能性越低。例如,当综合概率值

大于预设门限值时,确定目标对象为人脸活体;当综合概率值小于预设门限值时,确定目标对象不是人脸活体。其中,预设门限值可以根据实际需求预先设定,本申请实施例对此不作限定,例如预设门限值为0.5或其它数值。

[0088] 需要说明的一点是,在本申请实施例中,对步骤202和步骤203的执行先后顺序不作限定,步骤203可以在步骤202之后执行,也可以在步骤202之前执行,或者与步骤202同时执行。

[0089] 另外,除了以单帧RGB图像和深度图像作为神经网络输入之外,还可以采集多帧RGB图像和深度图像作为神经网络输入,使用RNN (Recurrent Neural Network,递归神经网络)等具备上下文感知的神经网络对多帧图像进行特征提取。

[0090] 综上所述,本申请实施例提供的技术方案中,通过两个神经网络分别对RGB图像和深度图像进行处理,得到第一概率值和第二概率值,使得人脸活体检测模型的模型结构更为灵活,可以将两个神经网络单独训练和部署。

[0091] 另外,还通过特征融合网络对第一神经网络和第二神经网络的隐藏层特征进行融合,使得第一神经网络具备融合RGB线索和深度线索进行第一概率值预测的能力,第二神经网络具备融合深度线索和RGB线索进行第二概率值预测的能力,有助于提升鲁棒性。

[0092] 请参考图4,其示出了本申请另一个实施例提供的人脸活体检测方法的流程图。该方法可应用于上文介绍的计算机设备中。该方法可以包括如下几个步骤:

[0093] 步骤401,获取目标对象的RGB图像和深度图像。

[0094] 本步骤与图1实施例中的步骤101相同,参见图1实施例中的介绍说明,本实施例对此不再赘述。

[0095] 与图2实施例不同的是,在本实施例中,人脸活体检测模型包括第一神经网络、第二神经网络和概率输出网络,该模型架构可参见图5所示。

[0096] 步骤402,调用第一神经网络对RGB图像进行处理得到第一输出特征。

[0097] 在本实施例中,第一神经网络可以是图2实施例介绍的第一神经网络中去除输出层后的网络架构。例如,当第一神经网络为CNN时,第一神经网络可以包括输入层、卷积层、池化层和全连接层。有关各层的功能可参见图2实施例中的介绍说明。

[0098] 步骤403,调用第二神经网络对深度图像进行处理得到第二输出特征。

[0099] 在本实施例中,第二神经网络可以是图2实施例介绍的第二神经网络中去除输出层后的网络架构。例如,当第二神经网络为CNN时,第二神经网络可以包括输入层、卷积层、池化层和全连接层。有关各层的功能可参见图2实施例中的介绍说明。

[0100] 可选地,第一神经网络在对RGB图像进行处理的过程中,也可以融合第二神经网络从深度图像中提取的特征,得到第一输出特征;和/或,第二神经网络在对深度图像进行处理的过程中,也可以融合第一神经网络从RGB图像中提取的特征,得到第二输出特征。有关特征融合的具体方式,可参见上文介绍说明。

[0101] 步骤404,调用概率输出网络对第一输出特征和第二输出特征进行处理,得到概率值。

[0102] 概率值用于指示目标对象为人脸活体的概率。可选地,概率值的取值范围是[0, 1]。概率值越大,表明目标对象为人脸活体的可能性越高;反之,概率值越小,表明目标对象是人脸活体的可能性越低。

[0103] 概率输出网络可以包括若干个隐藏层和输出层。隐藏层用于对上述第一输出特征和第二输出特征进行处理,得到上述概率值。输出层用于输出上述概率值。概率输出网络也可以是CNN,包括卷积层、池化层、全连接层和输出层。概率输出网络的输出层可以由分类器实现,用于实现二分类任务,输出是人脸活体以及不是人脸活体的概率。例如,上述分类器可以是Softmax分类器。

[0104] 步骤405,根据概率值,确定目标对象是否为人脸活体。

[0105] 例如,将概率值和预设阈值进行比对,当概率值大于预设阈值时,确定目标对象是人脸活体,当概率值小于预设阈值时,确定目标对象不是人脸活体。

[0106] 需要说明的一点是,在本申请实施例中,对步骤402和步骤403的执行先后顺序不作限定,步骤403可以在步骤402之后执行,也可以在步骤402之前执行,或者与步骤402同时执行。

[0107] 综上所述,本申请实施例提供的技术方案中,通过两个神经网络分别对RGB图像和深度图像进行处理,提取相应特征,并通过概率输出网络对上述特征进行融合处理,最终输出用于表示目标对象是否为人脸活体的概率值,有助于提升模型的鲁棒性。

[0108] 请参考图6,其示出了本申请另一个实施例提供的人脸活体检测方法的流程图。该方法可应用于上文介绍的计算机设备中。该方法可以包括如下几个步骤:

[0109] 步骤601,获取目标对象的RGB图像和深度图像。

[0110] 本步骤与图1实施例中的步骤101相同,参见图1实施例中的介绍说明,本实施例对此不再赘述。

[0111] 步骤602,检测RGB图像中的人脸区域。

[0112] 人脸区域是指包含有人脸的图像区域。人脸区域可以是一个矩形区域,人脸区域的位置可以采用该人脸区域在RGB图像中的左上角顶点坐标、人脸区域的宽度以及人脸区域的高度来表示。

[0113] 对于从RGB图像中检测人脸区域所采用的人脸检测算法,本申请实施例并不做具体限定。人脸检测的一般过程可以包括如下步骤:选择RGB图像中的某个矩形区域作为一个观察窗口,在选定的观察窗口中提取一些特征对其包含的图像内容进行描述,根据上述特征来判断该观察窗口是不是正好框住了一个人脸,通过不断重复上面步骤,直到遍历完成所有的观察窗口后停止。

[0114] 另外,检测结果为RGB图像中不包含人脸区域,则结束流程,可以判定目标对象不是人脸活体。

[0115] 步骤603,获取RGB图像中人脸区域中的RGB图像内容。

[0116] 如果从RGB图像中检测到人脸区域,则从RGB图像中提取该人脸区域中的RGB图像内容,作为人脸活体检测模型的输入图像。

[0117] 步骤604,根据人脸区域在RGB图像中的位置,获取深度图像中上述位置处的深度图像内容。

[0118] 在确定人脸区域在RGB图像中的位置之后,从深度图像中提取相同位置处的深度图像内容,作为人脸活体检测模型的输入图像。

[0119] 步骤605,调用第一神经网络对RGB图像的人脸区域中的RGB图像内容进行处理得到第一处理结果。

[0120] 步骤606,调用第二神经网络对深度图像的人脸区域中的深度图像内容进行处理得到第二处理结果。

[0121] 步骤607,根据第一处理结果和第二处理结果,确定目标对象是否为人脸活体。

[0122] 上述步骤605-607与图1实施例中的步骤102-104相同或类似,参见图1实施例中的介绍说明,本实施例对此不再赘述。

[0123] 另外,在本申请实施例中,对步骤605和步骤606的执行先后顺序不作限定,步骤606可以在步骤605之后执行,也可以在步骤605之前执行,或者与步骤605同时执行。

[0124] 结合参考图7,在获取目标对象的RGB图像和深度图像之后,对RGB图像进行人脸检测和人脸配准,确定RGB图像中的人脸区域,然后根据人脸区域的位置确定深度图像中的人脸区域;分别从RGB图像和深度图像中进行抠图处理,得到人脸区域中的RGB图像内容和深度图像内容,然后将人脸区域中的RGB图像内容和深度图像内容输入至人脸活体检测模型(包括上文介绍的第一神经网络和第二神经网络),得到第一概率值和第二概率值,最后结合第一概率值和第二概率值进行融合决策,确定目标对象是否为人脸活体。

[0125] 综上所述,本申请实施例提供的技术方案中,还通过检测RGB图像中的人脸区域,将人脸区域中的RGB图像内容和深度图像内容作为神经网络的输入,有助于减少神经网络的数据处理量,提高人脸活体检测的效率。

[0126] 在基于上述任一方法实施例提供的另一个可选实施例中,除了获取目标对象的RGB图像和深度图像之外,还可以获取目标对象的红外图像,而后综合RGB图像、深度图像和红外图像这3种图像信息,确定目标对象是否为人脸。

[0127] 目标对象的红外图像是指采用红外摄像头对目标对象进行拍摄得到的图像。红外图像的每个像素值采用红外特征来表示。目标对象的RGB图像、深度图像和红外图像可以是配准的,任意两个图像的像素点之间具有一对一的对应关系。

[0128] 在一个示例中,人脸活体检测模型包括第一神经网络、第二神经网络和第四神经网络。计算机设备调用第四神经网络对红外图像进行处理,得到第三处理结果;根据第一处理结果、第二处理结果和第三处理结果,确定目标对象是否为人脸活体。

[0129] 第三处理结果是从红外图像中提取的用于表征目标对象是否为人脸活体的分类信息。在一个示例中,第三处理结果为第三概率值,该第三概率值是根据红外图像所确定的目标对象为人脸活体的概率。可选地,第三概率值的取值范围是 $[0,1]$ 。第三概率值越大,表明目标对象为人脸活体的可能性越高;反之,第三概率值越小,表明目标对象是人脸活体的可能性越低。

[0130] 相应地,第四神经网络可以包括输入层、隐藏层和输出层。输入层用于输入红外图像,输出层用于输出第三概率值,隐藏层用于对红外图像进行处理得到上述第三概率值。与第一神经网络相类似,第四神经网络也可以是深度神经网络,如卷积神经网络。关于卷积神经网络的网络结构可参见上文介绍说明,此处不再赘述。

[0131] 可选地,第一神经网络在对RGB图像进行处理的过程中,融合第二神经网络从深度图像中提取的特征和/或第四神经网络从红外图像中提取的特征,得到第一概率值;和/或,第二神经网络在对深度图像进行处理的过程中,融合第一神经网络从RGB图像中提取的特征和/或第四神经网络从红外图像中提取的特征,得到第二概率值;和/或,第四神经网络在对红外图像进行处理的过程中,融合第一神经网络从RGB图像中提取的特征和/或第二神经

网络从深度图像中提取的特征,得到第三概率值。有关特征融合的方式可参见上文介绍说明,此处不再赘述。

[0132] 另外,在得到第一概率值、第二概率值和第三概率值之后,可以综合这三个概率值,最终确定目标对象是否为人脸活体。

[0133] 在一种可能的实施方式中,若第一概率值大于第一阈值、第二概率值大于第二阈值、且第三概率值大于第三阈值,则确定目标对象为人脸活体;若第一概率值小于第一阈值和/或第二概率值小于第二阈值和/或第三概率值小于第三阈值,则确定目标对象不是人脸活体。

[0134] 在另一种可能的实施方式中,调用第三神经网络对第一概率值、第二概率值和第三概率值进行处理,得到综合概率值;根据综合概率值确定目标对象是否为人脸活体。

[0135] 在另一个示例中,人脸活体检测模型包括第一神经网络、第二神经网络、第四神经网络和概率输出网络。计算机设备调用第四神经网络对红外图像进行处理得到第三输出特征;然后调用概率输出网络对第一输出特征、第二输出特征和第三输出特征进行处理,得到概率值。

[0136] 综上所述,本申请实施例提供的技术方案中,利用红外图像对太阳光敏感、对灯光不敏感,抗干扰性强的特点进行人脸活体检测,在RGB线索、深度线索的基础上进一步结合红外线索,形成更强大的活体攻击拦截模型。

[0137] 请参考图8,其示出了本申请另一个实施例提供的人脸活体检测方法的流程图。在本实施例中,由终端和服务端各进行一次人脸活体检测,通过双重校验的方式进一步提升安全性和鲁棒性。

[0138] 终端通过摄像头采集目标对象的RGB图像和深度图像。终端可以是诸如手机、平板电脑、可穿戴设备、PC等电子设备。终端配备有摄像头,具备采集RGB图像和深度图像的功能。

[0139] 终端根据目标对象的RGB图像和所述深度图像,对目标对象进行人脸活体初步校验。在本申请实施例中,对终端进行人脸活体初步校验的方式不作限定,其可以采用上文介绍的通过神经网络对目标对象进行人脸活体初步校验,也可以采用其它方式。终端对目标对象进行人脸活体初步校验,其目的是在前端拦截一些低级别的攻击(如人脸纸片、屏幕翻拍等),减少后台服务器的访问量,提高前端交互的响应速度,避免因网络传输延迟导致前端响应速度过慢。

[0140] 若目标对象通过人脸活体初步校验,则终端向服务器发送目标对象的RGB图像和深度图像。终端和服务端之间可以通过网络建立通信连接,如有线网络或者无线网络。服务器可以是一台服务器,也可以是由多台服务器组成的服务器集群,或者是一个云计算服务中心。

[0141] 之后,服务器调用上文介绍的人脸活体检测模型,对目标对象的RGB图像和深度图像进行处理,得到概率值,并根据该概率值确定目标对象是否为人脸活体。例如,如图8所示,服务器调用第一神经网络对目标对象的RGB图像进行处理得到第一概率值,以及调用第二神经网络对目标对象的深度图像进行处理得到第二概率值,服务器根据第一概率值和所述第二概率值,确定目标对象是否为人脸活体。有关服务器侧执行的上述步骤的介绍说明,可参见上文,本实施例对此不再赘述。

[0142] 综上所述,在本实施例中,由终端和服务端各进行一次人脸活体检测,通过双重校验的方式进一步提升安全性和鲁棒性。

[0143] 请参考图9,其示出了本申请另一个实施例提供的人脸活体检测方法的流程图。在本实施例中,人脸活体检测模型包括上文介绍的第一神经网络和第二神经网络。由终端和服务端配合完成人脸活体检测。

[0144] 终端通过摄像头采集目标对象的RGB图像和深度图像。终端调用第二神经网络对深度图像进行处理得到第二概率值;若第二概率值大于第二阈值,则终端向服务器发送RGB图像。

[0145] 服务器调用第一神经网络对RGB图像进行处理得到第一概率值;若第一概率值大于第一阈值,则服务器确定目标对象为人脸活体。

[0146] 综上所述,在本实施例中,分别在终端和服务端中部署第二神经网络和第一神经网络,由终端执行计算量相对较小的深度图像识别,由服务器执行计算量相对较大的RGB图像识别,终端和服务端配合完成整个人脸活体检测的流程。并且,第二神经网络可以选用轻量化的适合在诸如手机之类的终端中部署的神经网络,第一神经网络可以选用复杂且精度高的神经网络。

[0147] 本申请实施例提供的人脸活体检测的技术方案,可应用于任何有人脸活体检测需求的应用场景。下面,示例性介绍几个可能的应用场景:

[0148] 1、基于人脸识别的门禁系统

[0149] 相比于基于刷卡方式的门禁系统,基于人脸识别的门禁系统可以减少用户交互成本,提升用户体验,且刷脸核身的准确性远高于刷卡。在基于人脸识别的门禁系统中,对人脸进行活体检测,可以有效抵御人脸纸片、高清屏幕翻拍等方式的身份伪造,降低人工审核成本。

[0150] 门禁系统可以包括部署在门口的3D摄像头,以及用于对图像进行处理的后台处理设备(如PC)。当目标用户需要通过门口时,3D摄像头采集目标用户的RGB图像和深度图像,然后将目标用户的RGB图像和深度图像发送给后台处理设备。后台处理设备调用人脸活体检测模型对上述目标用户的RGB图像和深度图像进行处理,得到第一概率值和第二概率值,当第一概率值大于第一阈值且第二概率值大于第二阈值时,后台处理设备确定目标用户通过人脸活体检测,执行后续的人脸验证流程,并在人脸验证通过后开启门禁允许目标用户通行。

[0151] 2、基于人脸识别的支付验证场景

[0152] 相比于基于指纹识别的支付验证场景,基于人脸识别的支付验证场景只需用户站在摄像头前方,交互更为简单快捷。但是,人脸的造假成本比指纹更低,收集用户的人脸照片是很容易的事情。因此,需要在基于人脸识别的支付验证场景中,对人脸进行活体检测,以拦截仿冒他人人脸的恶意支付,保证用户的财产安全。

[0153] 在支付验证场景中,终端(如手机)的3D摄像头可以采集目标用户的RGB图像和深度图像,然后调用部署在终端的第二神经网络对上述目标用户的深度图像进行处理,得到第二概率值;当第二概率值大于第二阈值时,终端将目标用户的RGB图像发送给服务器;当第二概率值小于第二阈值时,终端提示用户人脸验证失败。服务器接收到终端发送的目标用户的RGB图像之后,调用第一神经网络对RGB图像进行处理得到第一概率值;若第一概率

值大于第一阈值,则服务器确定目标用户是人脸活体。在目标用户是人脸活体的情况下,服务器可以根据目标用户的RGB图像对目标用户作进一步的人脸验证流程,并在人脸验证通过后通知终端,终端在接收到人脸验证通过的响应信息之后,执行支付流程。

[0154] 上文仅介绍了两种可能的应用场景,本申请实施例提供的技术方案还可应用于终端解锁场景、应用登录场景、签到场景等任何具有人脸活体检测需求的应用场景中,本申请实施例对此不作限定。

[0155] 本申请一示例性实施例还提供了一种人脸活体检测模型的训练方法。该方法可应用于诸如PC之类的计算机设备中。该方法可以包括如下步骤:

[0156] 1、获取训练样本;

[0157] 训练样本包括:样本对象的RGB图像和深度图像,以及样本对象对应的标签,该标签用于指示样本对象是否为人脸活体。

[0158] 可选地,训练样本还包括样本对象的红外图像。

[0159] 在通常情况下,计算机设备获取的训练样本的数量为多个。训练样本越多,最终训练得到的模型的准确性通常来说越好。

[0160] 2、通过人脸活体检测模型对样本对象的RGB图像和深度图像进行处理,得到第一处理结果和第二处理结果;

[0161] 其中,人脸活体检测模型包括第一神经网络和第二神经网络,第一神经网络用于对样本对象的RGB图像进行处理得到第一处理结果,第二神经网络用于对样本对象的深度图像进行处理得到第二处理结果。

[0162] 可选地,人脸活体检测模型还包括第四神经网络。第四神经网络用于对样本对象的红外图像进行处理得到第三处理结果。

[0163] 可选地,第一神经网络在对样本对象的RGB图像进行处理的过程中,融合第二神经网络从样本对象的深度图像中提取的特征,得到第一处理结果;和/或,第二神经网络在对样本对象的深度图像进行处理的过程中,融合第一神经网络从样本对象的RGB图像中提取的特征,得到第二处理结果。

[0164] 可选地,当人脸活体检测模型还包括第四神经网络时,第一神经网络在对RGB图像进行处理的过程中,融合第二神经网络从深度图像中提取的特征和/或第四神经网络从红外图像中提取的特征,得到第一处理结果;和/或,第二神经网络在对深度图像进行处理的过程中,融合第一神经网络从RGB图像中提取的特征和/或第四神经网络从红外图像中提取的特征,得到第二处理结果;和/或,第四神经网络在对红外图像进行处理的过程中,融合第一神经网络从RGB图像中提取的特征和/或第二神经网络从深度图像中提取的特征,得到第三处理结果。

[0165] 3、根据第一处理结果和第二处理结果,预测样本对象是否为人脸活体;

[0166] 在一个示例中,人脸活体检测模型还包括第三神经网络。通过第三神经网络对第一概率值和第二概率值进行处理得到综合概率值,根据该综合概率值确定样本对象是否为人脸活体。

[0167] 在另一个示例中,若第一概率值大于第一阈值,且第二概率值大于第二阈值,则确定样本对象为人脸活体。

[0168] 在又一个示例中,人脸活体检测模型还包括概率输出网络。通过概率输出网络对

第一输出特征和第二输出特征进行处理,得到概率值。然后,根据概率值确定样本对象是否为人脸活体。

[0169] 4、根据样本对象的预测结果和标签对人脸活体检测模型的参数进行修正,直至人脸活体检测模型满足停止训练条件时,结束对人脸活体检测模型的训练。

[0170] 对于每一个样本对象,计算机设备可以将人脸活体检测模型对该样本对象的预测结果以及该样本对象对应的标签进行比较,以确定模型对该样本对象是否预测准确。

[0171] 另外,人脸活体检测模型的停止训练条件可以预先设定,例如停止训练条件包括人脸活体检测模型的预测准确度达到预设阈值,如95%。

[0172] 当人脸活体检测模型不满足停止训练条件时,计算机设备采用新的训练样本继续对该模型进行训练,以优化该模型的参数,直至人脸活体检测模型满足停止训练条件时停止,最终得到满足实际应用需求的人脸活体检测模型。

[0173] 完成训练的人脸活体检测模型可用于实现对目标对象执行上文实施例介绍的人脸活体检测。

[0174] 下述为本申请装置实施例,可以用于执行本申请方法实施例。对于本申请装置实施例中未披露的细节,请参照本申请方法实施例。

[0175] 请参考图10,其示出了本申请一个实施例提供的人脸活体检测装置的框图。该装置具有实现上述人脸活体检测方法的功能,所述功能可以由硬件实现,也可以由硬件执行相应的软件实现。该装置1000可以包括:图像获取模块1010、模型调用模块1020和活体检测模块1030。

[0176] 图像获取模块1010,用于获取目标对象的RGB图像和深度图像。

[0177] 模型调用模块1020,用于调用第一神经网络对所述RGB图像进行处理得到第一处理结果,所述第一处理结果是从所述RGB图像中提取的用于表征所述目标对象是否为人脸活体的分类信息;以及,调用第二神经网络对所述深度图像进行处理得到第二处理结果,所述第二处理结果是从所述深度图像中提取的用于表征所述目标对象是否为所述人脸活体的分类信息。

[0178] 活体检测模块1030,用于根据所述第一处理结果和所述第二处理结果,确定所述目标对象是否为所述人脸活体。

[0179] 综上所述,本申请实施例提供的技术方案中,通过获取目标对象的RGB图像和深度图像,结合上述两个图像对目标对象进行人脸活体检测,融合了人脸纹理线索和3D人脸结构线索进行人脸活体检测,提高了检测的准确性,可以有效拦截人脸纸片、高清屏幕人脸翻拍、合成人脸视频、面具、3D假体模型等攻击。并且,RGB图像和深度图像的拍摄过程中,无需用户执行交互动作,更为简单高效,且降低交互成本,提高用户体验。

[0180] 在基于图10实施例提供的一个可选实施例中,所述第一处理结果为第一概率值,所述第一概率值是根据所述RGB图像所确定的所述目标对象为所述人脸活体的概率;所述第二处理结果为第二概率值,所述第二概率值是根据所述深度图像所确定的所述目标对象为所述人脸活体的概率。

[0181] 相应地,在一个示例中,所述活体检测模块1030,用于当所述第一概率值大于第一阈值,且所述第二概率值大于第二阈值时,确定所述目标对象为所述人脸活体。在另一个示例中,所述活体检测模块1030,用于调用第三神经网络对所述第一概率值和所述第二概率

值进行处理,得到综合概率值;根据所述综合概率值确定所述目标对象是否为所述人脸活体。

[0182] 在基于图10实施例提供的另一个可选实施例中,所述第一处理结果为第一输出特征,所述第二处理结果为第二输出特征;其中,所述第一输出特征是从所述RGB图像中提取的深度特征,所述第二输出特征是从所述深度图像中提取的深度特征。

[0183] 相应地,所述活体检测模块1030,用于调用概率输出网络对所述第一输出特征和所述第二输出特征进行处理,得到概率值,所述概率值用于指示所述目标对象为所述人脸活体的概率;根据所述概率值确定所述目标对象是否为所述人脸活体。

[0184] 在基于图10实施例或者上述任一可选实施例提供的另一个可选实施例中,所述第一神经网络在对所述RGB图像进行处理的过程中,融合所述第二神经网络从所述深度图像中提取的特征,得到所述第一处理结果;和/或,所述第二神经网络在对所述深度图像进行处理的过程中,融合所述第一神经网络从所述RGB图像中提取的特征,得到所述第二处理结果。

[0185] 可选地,所述人脸活体检测模型还包括:特征融合网路;

[0186] 所述特征融合网路用于:

[0187] 对所述第一神经网络的第i个隐藏层输出的特征和所述第二神经网络的第j个隐藏层输出的特征进行特征数组加和处理和/或特征通道叠加处理,得到融合特征;其中,所述i为正整数,所述j为正整数;

[0188] 将所述融合特征输入至所述第一神经网络的第i+1个隐藏层和/或将所述融合特征输入至所述第二神经网络的第j+1个隐藏层。

[0189] 在基于图10实施例或者上述任一可选实施例提供的另一个可选实施例中,装置1000还包括:人脸检测模块、第一提取模块和第二提取模块。

[0190] 人脸检测模块,用于检测所述RGB图像中的人脸区域。

[0191] 第一提取模块,用于获取所述RGB图像中所述人脸区域中的RGB图像内容。

[0192] 第二提取模块,用于根据所述人脸区域在所述RGB图像中的位置,获取所述深度图像中所述位置处的深度图像内容。

[0193] 其中,所述第一神经网络用于对所述RGB图像中的所述RGB图像内容进行处理得到所述第一处理结果,所述第二神经网络用于对所述深度图像中的所述深度图像内容进行处理得到所述第二处理结果。

[0194] 在基于图10实施例或者上述任一可选实施例提供的另一个可选实施例中,

[0195] 所述图像获取模块1010,还用于获取所述目标对象的红外图像;

[0196] 所述模型调用模块1020,还用于调用第四神经网络对所述红外图像进行处理得到第三处理结果,所述第三处理结果是从所述红外图像中提取的用于表征所述目标对象是否为所述人脸活体的分类信息。

[0197] 所述活体检测模块1030,还用于根据所述第一处理结果、所述第二处理结果和所述第三处理结果,确定所述目标对象是否为所述人脸活体。

[0198] 请参考图11,其示出了本申请一个实施例提供的人脸活体检测模型的训练装置的框图。该装置具有实现上述人脸活体检测模型的训练方法的功能,所述功能可以由硬件实现,也可以由硬件执行相应的软件实现。该装置1100可以包括:样本获取模块1110、概率计

算模块1120、活体预测模块1130和模型训练模块1140。

[0199] 样本获取模块1110,用于获取训练样本,所述训练样本包括:样本对象的RGB图像和深度图像,以及所述样本对象对应的标签,所述标签用于指示所述样本对象是否为人脸活体。

[0200] 概率计算模块1120,用于通过所述人脸活体检测模型对所述样本对象的RGB图像和深度图像进行处理,得到第一处理结果和第二处理结果;其中,所述人脸活体检测模型包括第一神经网络和第二神经网络,所述第一神经网络用于对所述样本对象的RGB图像进行处理得到所述第一处理结果,所述第二神经网络用于对所述样本对象的深度图像进行处理得到所述第二处理结果。

[0201] 活体预测模块1130,用于根据所述第一处理结果和第二处理结果,预测所述样本对象是否为所述人脸活体。

[0202] 模型训练模块1140,用于根据所述样本对象的预测结果和所述标签对所述人脸活体检测模型的参数进行修正,直至所述人脸活体检测模型满足停止训练条件时,结束对所述人脸活体检测模型的训练。

[0203] 需要说明的是,上述实施例提供的装置,在实现其功能时,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将设备的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。另外,上述实施例提供的装置与方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0204] 请参考图12其示出了本申请一个实施例提供的计算机设备1200的结构框图。该计算机设备1200可以是手机、平板电脑、游戏主机、电子书阅读器、多媒体播放设备、可穿戴设备、PC等。

[0205] 通常,计算机设备1200包括有:处理器1201和存储器1202。

[0206] 处理器1201可以包括一个或多个处理核心,比如4核心处理器、8核心处理器等。处理器1201可以采用DSP(Digital Signal Processing,数字信号处理)、FPGA(Field Programmable Gate Array,现场可编程门阵列)、PLA(Programmable Logic Array,可编程逻辑阵列)中的至少一种硬件形式来实现。处理器1201也可以包括主处理器和协处理器,主处理器是用于对在唤醒状态下的数据进行处理的处理器,也称CPU(Central Processing Unit,中央处理器);协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中,处理器1201可以在集成有GPU(Graphics Processing Unit,图像处理器),GPU用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中,处理器1201还可以包括AI(Artificial Intelligence,人工智能)处理器,该AI处理器用于处理有关机器学习的计算操作。

[0207] 存储器1202可以包括一个或多个计算机可读存储介质,该计算机可读存储介质可以是非暂态的。存储器1202还可包括高速随机存取存储器,以及非易失性存储器,比如一个或多个磁盘存储设备、闪存存储设备。在一些实施例中,存储器1202中的非暂态的计算机可读存储介质用于存储至少一个指令,该至少一个指令用于被处理器1201所执行以实现本申请中方法实施例提供的人脸活体检测方法,或者实现本申请中方法实施例提供的人脸活体检测模型的训练方法。

[0208] 在一些实施例中, 计算机设备1200还可选包括有: 外围设备接口1203和至少一个外围设备。处理器1201、存储器1202和外围设备接口1203之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外围设备接口1203相连。具体地, 外围设备包括: 射频电路1204、触摸显示屏1205、摄像头1206、音频电路1207、定位组件1208和电源1209中的至少一种。

[0209] 外围设备接口1203可被用于将I/O (Input/Output, 输入/输出) 相关的至少一个外围设备连接到处理器1201和存储器1202。在一些实施例中, 处理器1201、存储器1202和外围设备接口1203被集成在同一芯片或电路板上; 在一些其他实施例中, 处理器1201、存储器1202和外围设备接口1203中的任意一个或两个可以在单独的芯片或电路板上实现, 本实施例对此不加以限定。

[0210] 射频电路1204用于接收和发射RF (Radio Frequency, 射频) 信号, 也称电磁信号。射频电路1204通过电磁信号与通信网络以及其他通信设备进行通信。射频电路1204将电信号转换为电磁信号进行发送, 或者, 将接收到的电磁信号转换为电信号。可选地, 射频电路1204包括: 天线系统、RF收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频电路1204可以通过至少一种无线通信协议来与其它设备进行通信。该无线通信协议包括但不限于: 城域网、各代移动通信网络 (2G、3G、4G及5G)、无线局域网和/或Wi-Fi (Wireless Fidelity, 无线保真) 网络。在一些实施例中, 射频电路1204还可以包括NFC (Near Field Communication, 近距离无线通信) 有关的电路, 本申请对此不加以限定。

[0211] 显示屏1205用于显示UI (User Interface, 用户界面)。该UI可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏1205是触摸显示屏时, 显示屏1205还具有采集在显示屏1205的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器1201进行处理。此时, 显示屏1205还可以用于提供虚拟按钮和/或虚拟键盘, 也称软按钮和/或软键盘。在一些实施例中, 显示屏1205可以为一个, 设置计算机设备1200的前面板; 在另一些实施例中, 显示屏1205可以为至少两个, 分别设置在计算机设备1200的不同表面或呈折叠设计; 在再一些实施例中, 显示屏1205可以是柔性显示屏, 设置在计算机设备1200的弯曲表面上或折叠面上。甚至, 显示屏1205还可以设置成非矩形的不规则图形, 也即异形屏。显示屏1205可以采用LCD (Liquid Crystal Display, 液晶显示屏)、OLED (Organic Light-Emitting Diode, 有机发光二极管) 等材质制备。

[0212] 摄像头组件1206用于采集图像或视频。可选地, 摄像头组件1206包括前置摄像头和后置摄像头。通常, 前置摄像头设置在计算机设备的前面板, 后置摄像头设置在计算机设备的背面。在一些实施例中, 后置摄像头为至少两个, 分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种, 以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及VR (Virtual Reality, 虚拟现实) 拍摄功能或者其它融合拍摄功能。在一些实施例中, 摄像头组件1206还可以包括闪光灯。闪光灯可以是单色温闪光灯, 也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合, 可以用于不同色温下的光线补偿。

[0213] 音频电路1207可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波, 并将声波转换为电信号输入至处理器1201进行处理, 或者输入至射频电路1204以实现语音通

信。出于立体声采集或降噪的目的,麦克风可以为多个,分别设置在计算机设备1200的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器1201或射频电路1204的电信号转换为声波。扬声器可以是传统的薄膜扬声器,也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时,不仅可以将电信号转换为人类可听见的声波,也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中,音频电路1207还可以包括耳机插孔。

[0214] 定位组件1208用于定位计算机设备1200的当前地理位置,以实现导航或LBS (Location Based Service,基于位置的服务)。定位组件1208可以是基于美国的GPS (Global Positioning System,全球定位系统)、中国的北斗系统、俄罗斯的格雷纳斯系统或欧盟的伽利略系统的定位组件。

[0215] 电源1209用于为计算机设备1200中的各个组件进行供电。电源1209可以是交流电、直流电、一次性电池或可充电电池。当电源1209包括可充电电池时,该可充电电池可以是有线充电电池或无线充电电池。该可充电电池还可以用于支持快充技术。

[0216] 在一些实施例中,计算机设备1200还包括有一个或多个传感器1210。该一个或多个传感器1210包括但不限于:加速度传感器1211、陀螺仪传感器1212、压力传感器1213、指纹传感器1214、光学传感器1215以及接近传感器1216。

[0217] 加速度传感器1211可以检测以计算机设备1200建立的坐标系的三个坐标轴上的加速度大小。比如,加速度传感器1211可以用于检测重力加速度在三个坐标轴上的分量。处理器1201可以根据加速度传感器1211采集的重力加速度信号,控制触摸显示屏1205以横向视图或纵向视图进行用户界面的显示。加速度传感器1211还可以用于游戏或者用户的运动数据的采集。

[0218] 陀螺仪传感器1212可以检测计算机设备1200的机体方向及转动角度,陀螺仪传感器1212可以与加速度传感器1211协同采集用户对计算机设备1200的3D动作。处理器1201根据陀螺仪传感器1212采集的数据,可以实现如下功能:动作感应(比如根据用户的倾斜操作来改变UI)、拍摄时的图像稳定、游戏控制以及惯性导航。

[0219] 压力传感器1213可以设置在计算机设备1200的侧边框和/或触摸显示屏1205的下层。当压力传感器1213设置在计算机设备1200的侧边框时,可以检测用户对计算机设备1200的握持信号,由处理器1201根据压力传感器1213采集的握持信号进行左右手识别或快捷操作。当压力传感器1213设置在触摸显示屏1205的下层时,由处理器1201根据用户对触摸显示屏1205的压力操作,实现对UI界面上的可操作性控件进行控制。可操作性控件包括按钮控件、滚动条控件、图标控件、菜单控件中的至少一种。

[0220] 指纹传感器1214用于采集用户的指纹,由处理器1201根据指纹传感器1214采集到的指纹识别用户的身份,或者,由指纹传感器1214根据采集到的指纹识别用户的身份。在识别出用户的身份为可信身份时,由处理器1201授权该用户执行相关的敏感操作,该敏感操作包括解锁屏幕、查看加密信息、下载软件、支付及更改设置等。指纹传感器1214可以被设置计算机设备1200的正面、背面或侧面。当计算机设备1200上设置有物理按键或厂商Logo时,指纹传感器1214可以与物理按键或厂商Logo集成在一起。

[0221] 光学传感器1215用于采集环境光强度。在一个实施例中,处理器1201可以根据光学传感器1215采集的环境光强度,控制触摸显示屏1205的显示亮度。具体地,当环境光强度

较高时,调高触摸显示屏1205的显示亮度;当环境光强度较低时,调低触摸显示屏1205的显示亮度。在另一个实施例中,处理器1201还可以根据光学传感器1215采集的环境光强度,动态调整摄像头组件1206的拍摄参数。

[0222] 接近传感器1216,也称距离传感器,通常设置在计算机设备1200的前面板。接近传感器1216用于采集用户与计算机设备1200的正面之间的距离。在一个实施例中,当接近传感器1216检测到用户与计算机设备1200的正面之间的距离逐渐变小时,由处理器1201控制触摸显示屏1205从亮屏状态切换为息屏状态;当接近传感器1216检测到用户与计算机设备1200的正面之间的距离逐渐变大时,由处理器1201控制触摸显示屏1205从息屏状态切换为亮屏状态。

[0223] 本领域技术人员可以理解,图12中示出的结构并不构成对计算机设备1200的限定,可以包括比图示更多或更少的组件,或者组合某些组件,或者采用不同的组件布置。

[0224] 在示例中实施例中,还提供了一种计算机设备,所述计算机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、至少一段程序、代码集或指令集。所述至少一条指令、至少一段程序、代码集或指令集经配置以由一个或者一个以上处理器执行,以实现上述人脸活体检测方法,或者实现上述人脸活体检测模型的训练方法。

[0225] 在示例性实施例中,还提供了一种计算机可读存储介质,所述存储介质中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或所述指令集在被计算机设备的处理器执行时实现上述人脸活体检测方法,或者实现上述人脸活体检测模型的训练方法。

[0226] 可选地,上述计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0227] 在示例性实施例中,还提供了一种计算机程序产品,当该计算机程序产品被执行时,其用于实现上述人脸活体检测方法,或者实现上述人脸活体检测模型的训练方法。

[0228] 应当理解的是,在本文中提及的“多个”是指两个或两个以上。“和/或”,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。字符“/”一般表示前后关联对象是一种“或”的关系。

[0229] 以上所述仅为本申请的示例性实施例,并不用以限制本申请,凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

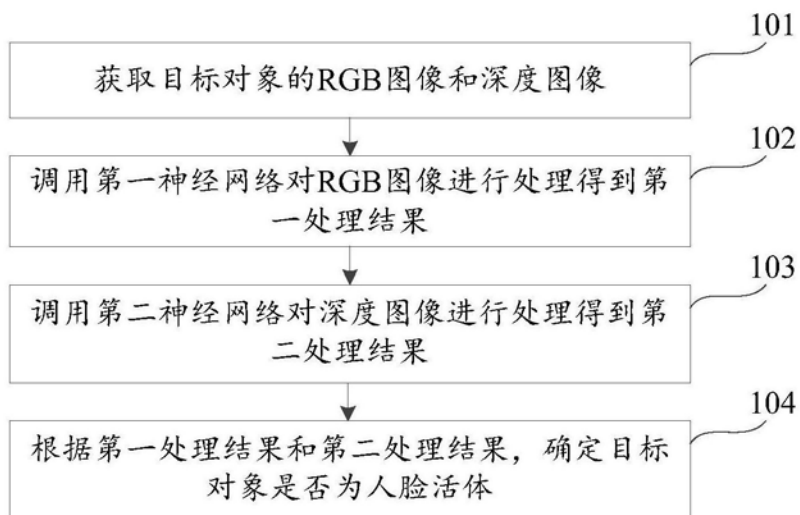


图1

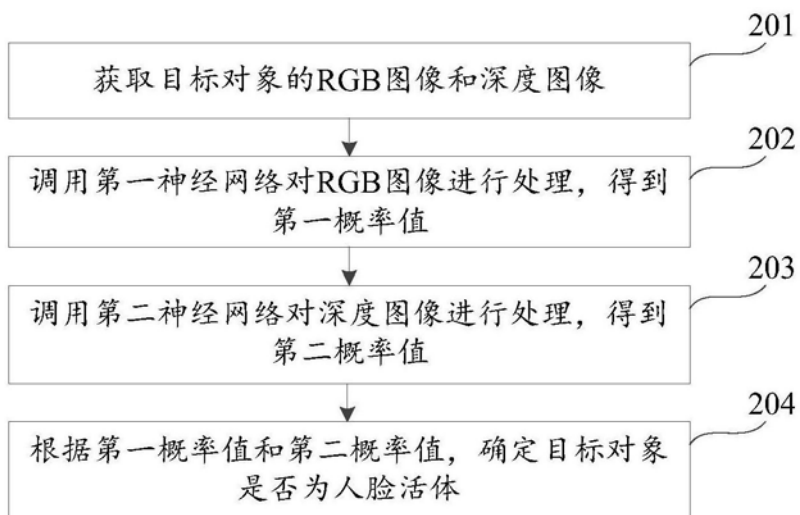


图2

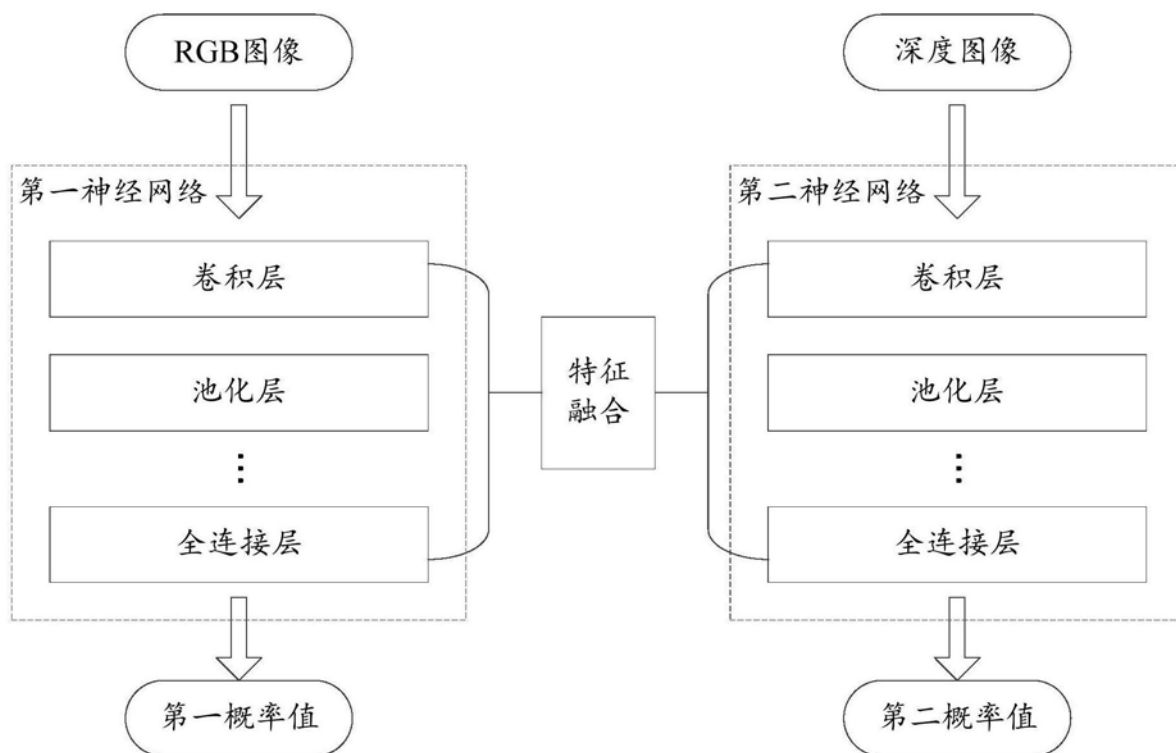


图3

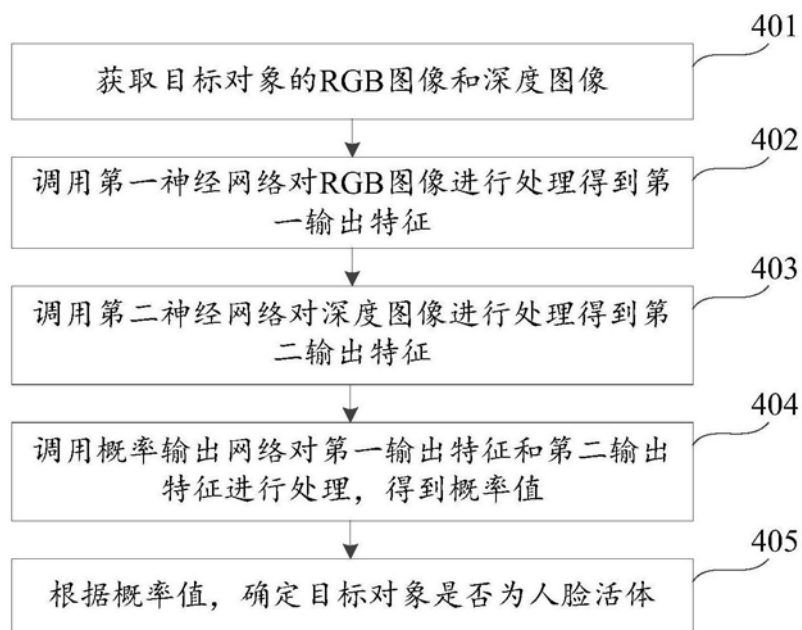


图4

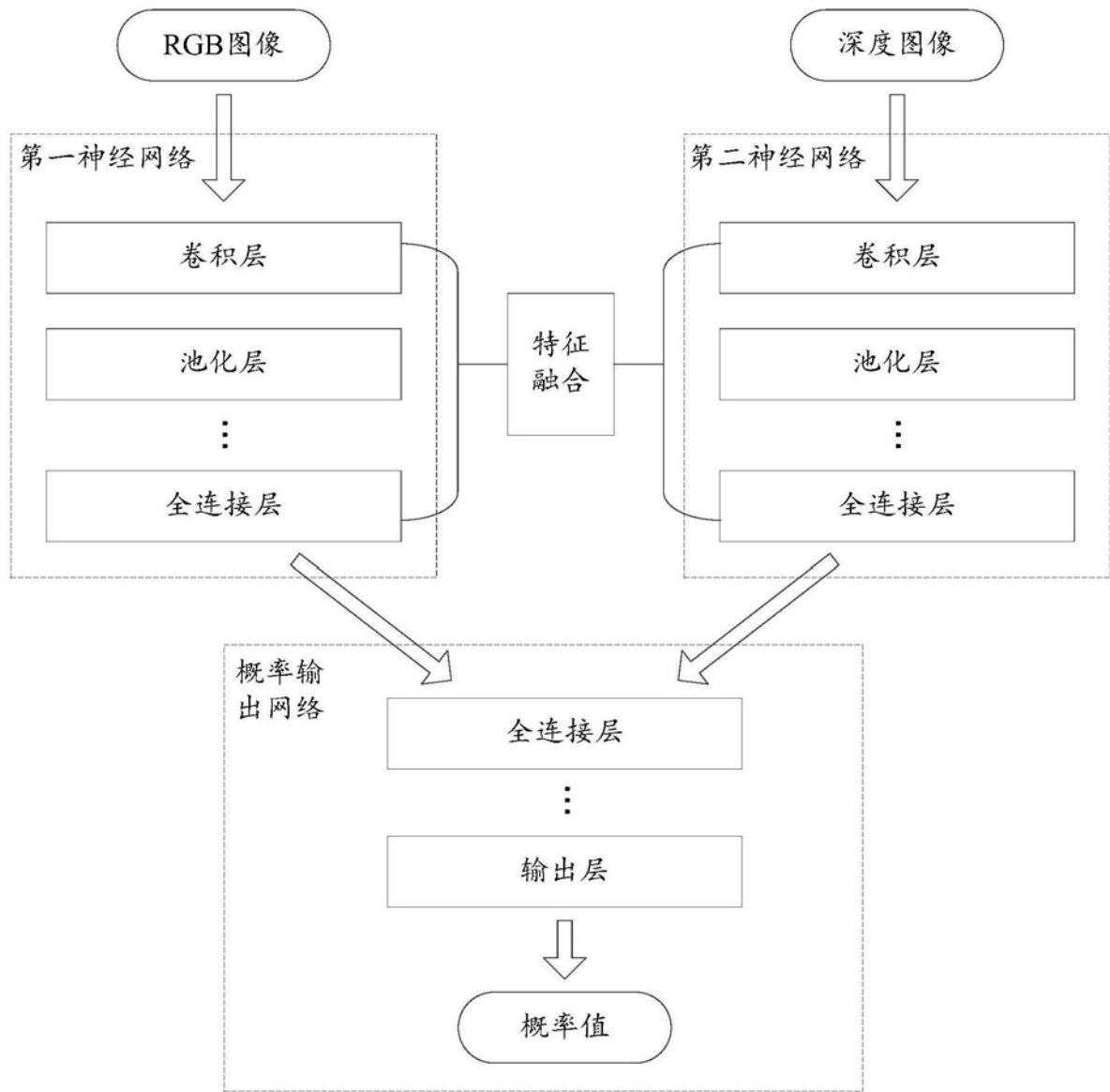


图5

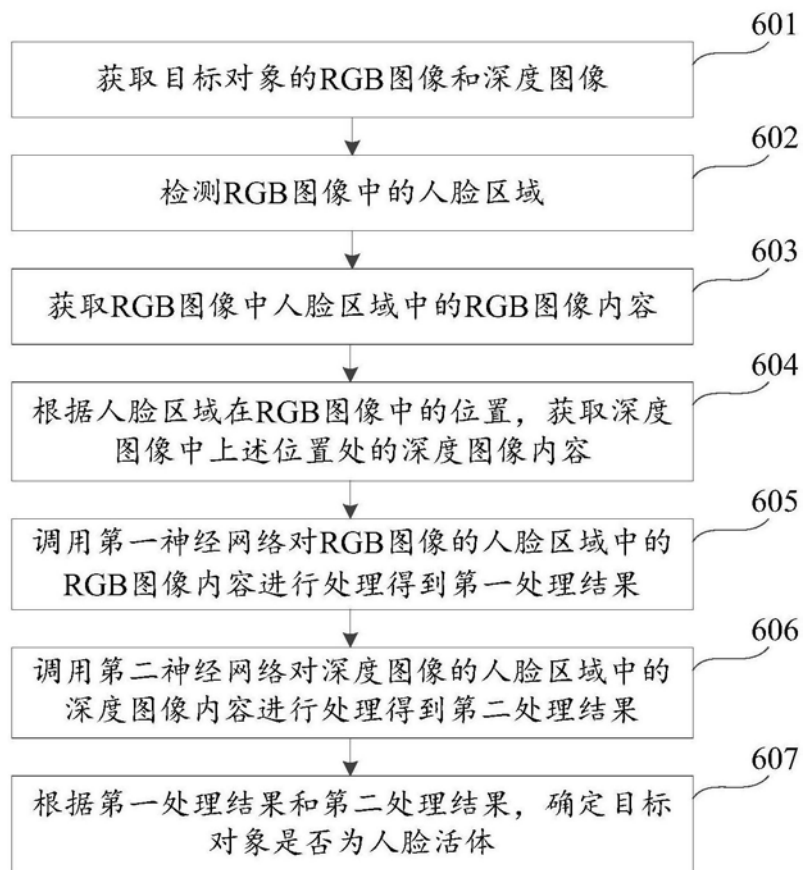


图6

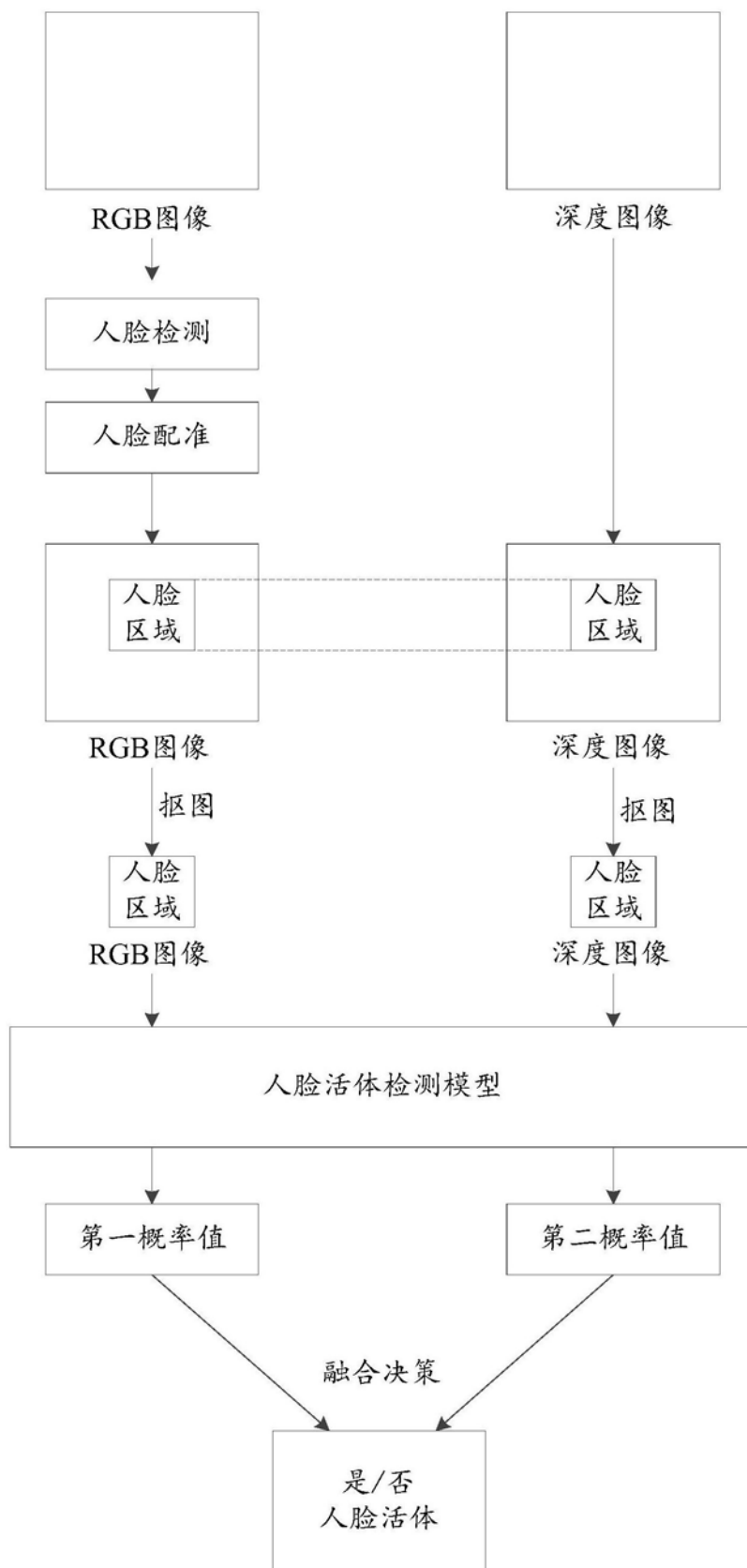


图7

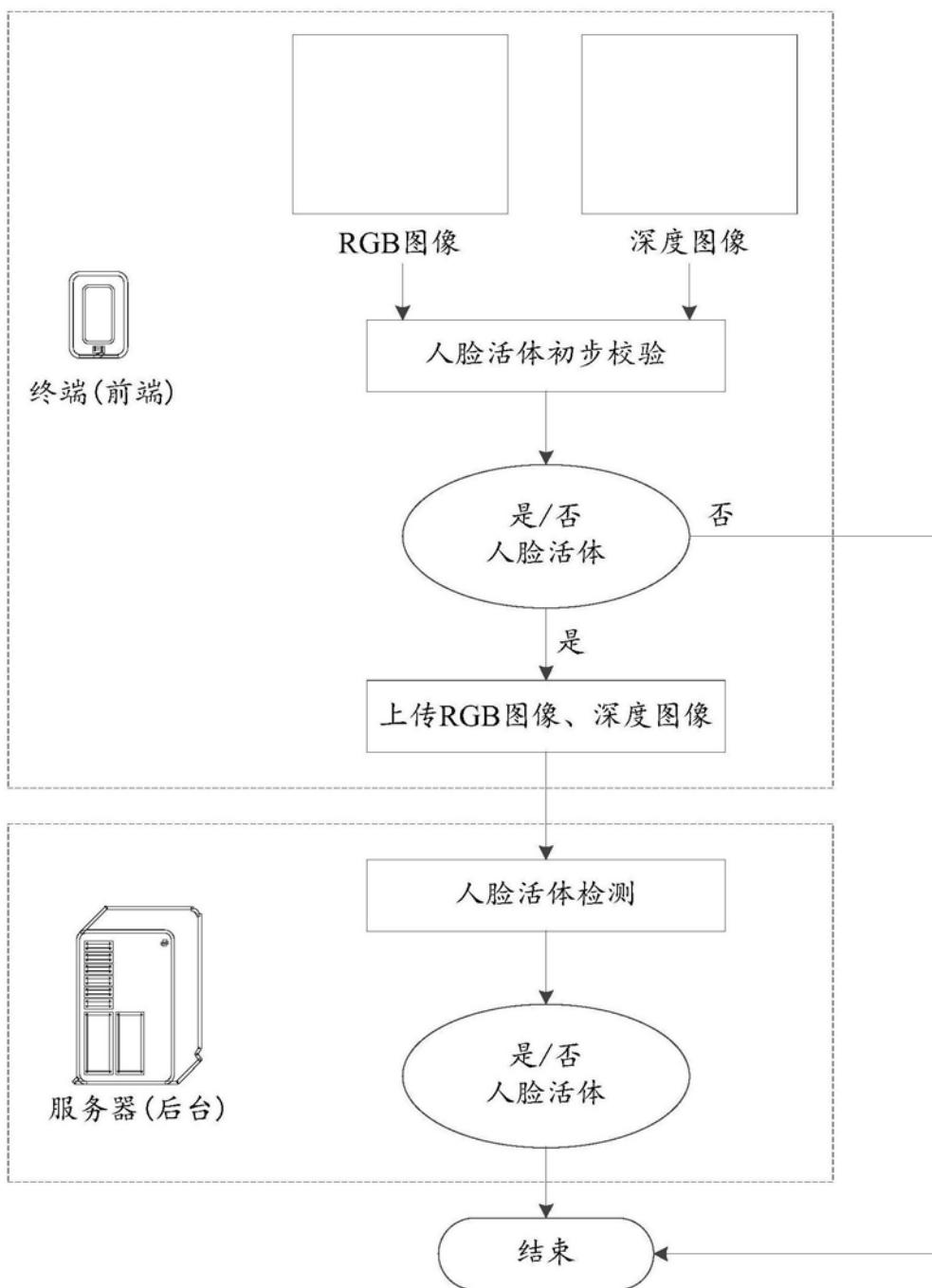


图8

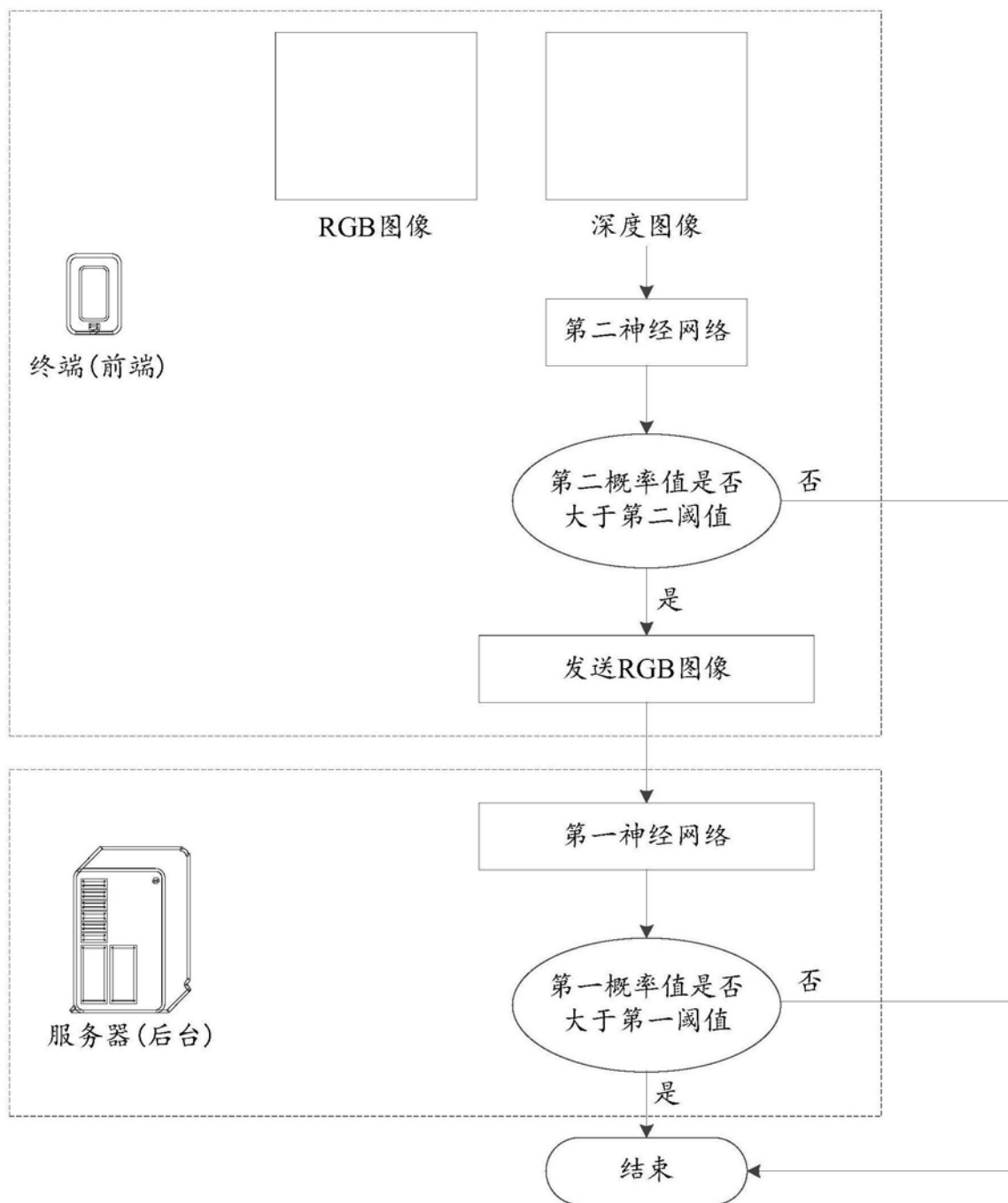


图9

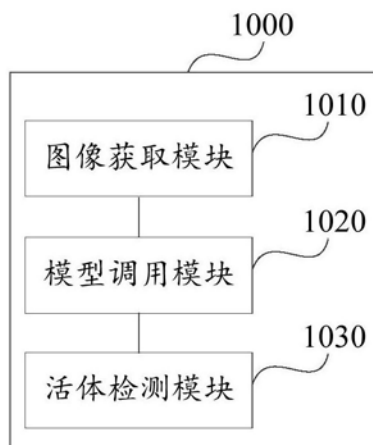


图10

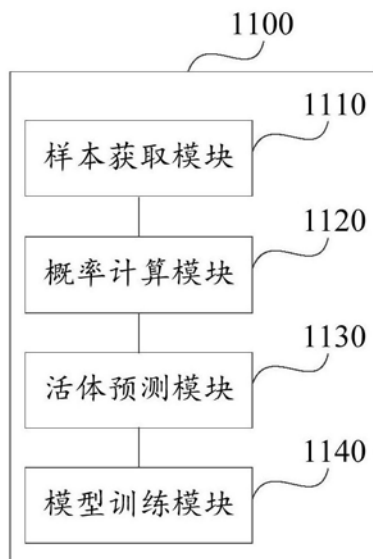


图11

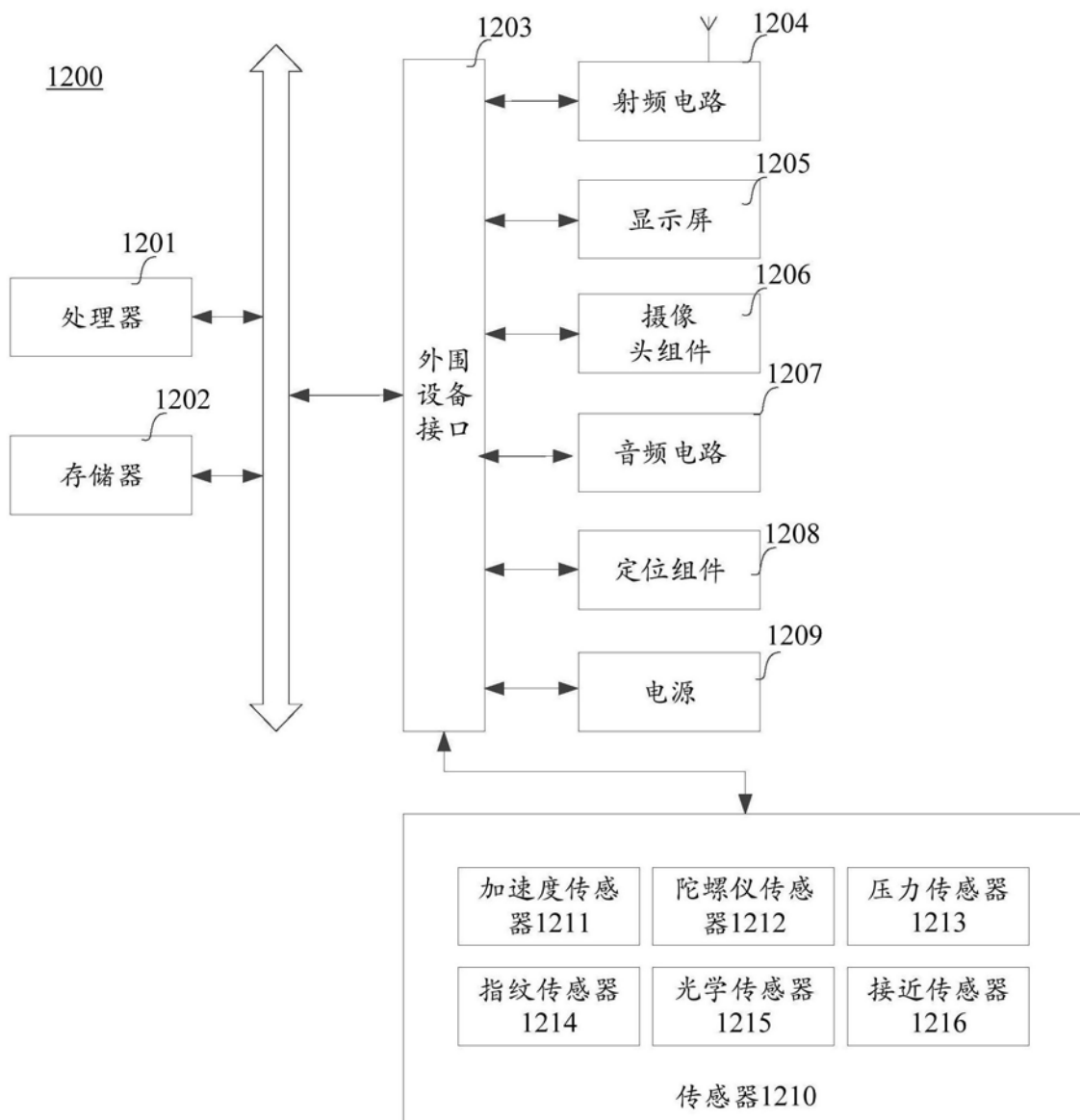


图12