



(12)发明专利

(10)授权公告号 CN 106384098 B

(45)授权公告日 2019.11.26

(21)申请号 201610849615.0

(22)申请日 2016.09.23

(65)同一申请的已公布的文献号  
申请公布号 CN 106384098 A

(43)申请公布日 2017.02.08

(73)专利权人 北京小米移动软件有限公司  
地址 100085 北京市海淀区清河中街68号  
华润五彩城购物中心二期9层01房间

(72)发明人 杨松 王百超 张旭华

(74)专利代理机构 北京同立钧成知识产权代理  
有限公司 11205  
代理人 杨贝贝 刘芳

(51)Int.Cl.  
G06K 9/00(2006.01)  
G06N 3/02(2006.01)

(56)对比文件

CN 104778448 A,2015.07.15,  
CN 105718868 A,2016.06.29,  
CN 105426930 A,2016.03.23,  
US 2016104058 A1,2016.04.14,  
US 2003179931 A1,2003.09.25,  
桑高丽 等.一种基于深度卷积网络的鲁棒  
头部姿态估计方法.《四川大学学报(工程科学  
版)》.2016,第48卷

审查员 王桂珍

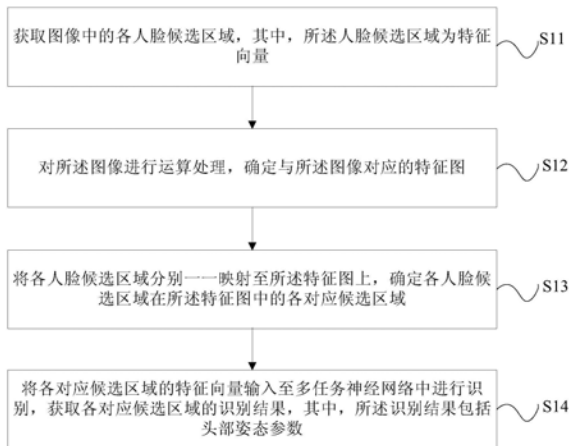
权利要求书3页 说明书10页 附图3页

(54)发明名称

基于图像的头部姿态检测方法、装置以及终端

(57)摘要

本公开是关于一种基于图像的头部姿态检测方法、装置以及终端,其中,该方法包括:获取图像中的各人脸候选区域,其中,人脸候选区域为特征向量;对图像进行运算处理,确定与图像对应的特征图;将各人脸候选区域分别一一映射至特征图上,确定各人脸候选区域在特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,识别结果包括头部姿态参数。通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。



1. 一种基于图像的头部姿态检测方法,其特征在于,包括:
  - 获取图像中的各人脸候选区域,其中,所述人脸候选区域为特征向量;
  - 对所述图像进行运算处理,确定与所述图像对应的特征图;
  - 将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;
  - 将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数;
  - 所述对应候选区域为 $R_f = (x_f, y_f, w_f, h_f)$ ;
  - 其中, $R_f = s * R$ , $R = (x, y, w, h)$ 为人脸候选区域, $x$ 为人脸候选区域的左上角的 $x$ 轴坐标值, $x_f$ 为对应候选区域的左上角的 $x$ 轴坐标值,并且 $x_f = s * x$ ; $y$ 为人脸候选区域的左上角的 $y$ 轴坐标值, $y_f$ 为对应候选区域的左上角的 $y$ 轴坐标值,并且 $y_f = s * y$ ; $w$ 为人脸候选区域的宽度, $w_f$ 为对应候选区域的宽度,并且 $w_f = s * w$ ; $h$ 为人脸候选区域的高度, $h_f$ 为对应候选区域的高度,并且 $h_f = s * h$ ; $s$ 为所述图像的尺寸到所述特征图的尺寸的缩放系数;
  - 所述识别结果还包括人脸概率;
  - 相应的,在所述将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果之后,还包括:
    - 采用交除并参数为0.5的非最大值抑制方法,对根据各对应候选区域的识别结果中的人脸概率进行处理,以保留交除并参数大于0.5中的人脸概率最大的各对应候选区域;
    - 去除交除并参数大于0.5中的人脸概率最大的各对应候选区域中,人脸概率小于0.8的对应候选区域,以确定过滤后的各对应候选区域的识别结果。
2. 根据权利要求1所述的方法,其特征在于,所述获取图像中的各人脸候选区域,包括:
  - 采用显著性检测方法,对图像进行检测,以获取图像中的各人脸候选区域。
3. 根据权利要求1所述的方法,其特征在于,所述对所述图像进行运算处理,确定与所述图像对应的特征图,包括:
  - 对所述图像进行 $N$ 层的卷积神经网络CNN卷积运算,以确定与所述图像对应的特征图,其中, $N$ 为正整数。
4. 根据权利要求1所述的方法,其特征在于,在所述将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域之后,还包括:
  - 采用自适应池化ROI Pooling方法对各对应候选区域进行处理,以将各对应候选区域映射为相同尺寸的各对应候选区域。
5. 一种基于图像的头部姿态检测装置,其特征在于,包括:
  - 获取模块,被配置为获取图像中的各人脸候选区域,其中,所述人脸候选区域为特征向量;
  - 确定模块,被配置为对所述图像进行运算处理,确定与所述图像对应的特征图;
  - 映射模块,被配置为将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;
  - 识别模块,被配置为将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数;
  - 所述对应候选区域为 $R_f = (x_f, y_f, w_f, h_f)$ ;

其中,  $R_f = s * R$ ,  $R = (x, y, w, h)$  为人脸候选区域,  $x$  为人脸候选区域的左上角的  $x$  轴坐标值,  $x_f$  为对应候选区域的左上角的  $x$  轴坐标值, 并且  $x_f = s * x$ ;  $y$  为人脸候选区域的左上角的  $y$  轴坐标值,  $y_f$  为对应候选区域的左上角的  $y$  轴坐标值, 并且  $y_f = s * y$ ;  $w$  为人脸候选区域的宽度,  $w_f$  为对应候选区域的宽度, 并且  $w_f = s * w$ ;  $h$  为人脸候选区域的高度,  $h_f$  为对应候选区域的高度, 并且  $h_f = s * h$ ;  $s$  为所述图像的尺寸到所述特征图的尺寸的缩放系数;

所述识别结果还包括人脸概率;

相应的, 所述装置, 还包括:

筛选模块, 被配置为在所述识别模块将各对应候选区域的特征向量输入至多任务神经网络中进行识别, 获取各对应候选区域的识别结果之后, 采用交除并参数为 0.5 的非最大值抑制方法, 对根据各对应候选区域的识别结果中的人脸概率进行处理, 以保留交除并参数大于 0.5 中的人脸概率最大的各对应候选区域; 去除交除并参数大于 0.5 中的人脸概率最大的各对应候选区域中, 人脸概率小于 0.8 的对应候选区域, 以确定过滤后的各对应候选区域的识别结果。

6. 根据权利要求 5 所述的装置, 其特征在于, 所述获取模块, 被具体配置为:

采用显著性检测方法, 对图像进行检测, 以获取图像中的各人脸候选区域。

7. 根据权利要求 5 所述的装置, 其特征在于, 所述确定模块, 被具体配置为:

对所述图像进行  $N$  层的 CNN 卷积运算, 以确定与所述图像对应的特征图, 其中,  $N$  为正整数。

8. 根据权利要求 5 所述的装置, 其特征在于, 所述装置, 还包括:

处理模块, 被配置为在所述映射模块将各人脸候选区域分别一一映射至所述特征图上, 确定各人脸候选区域在所述特征图中的各对应候选区域之后, 采用 ROI Pooling 方法对各对应候选区域进行处理, 以将各对应候选区域映射为相同尺寸的各对应候选区域。

9. 一种终端, 其特征在于, 包括:

处理器, 以及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: 获取图像中的各人脸候选区域, 其中, 所述人脸候选区域为特征向量; 对所述图像进行运算处理, 确定与所述图像对应的特征图; 将各人脸候选区域分别一一映射至所述特征图上, 确定各人脸候选区域在所述特征图中的各对应候选区域; 将各对应候选区域的特征向量输入至多任务神经网络中进行识别, 获取各对应候选区域的识别结果, 其中, 所述识别结果包括头部姿态参数;

所述对应候选区域为  $R_f = (x_f, y_f, w_f, h_f)$ ;

其中,  $R_f = s * R$ ,  $R = (x, y, w, h)$  为人脸候选区域,  $x$  为人脸候选区域的左上角的  $x$  轴坐标值,  $x_f$  为对应候选区域的左上角的  $x$  轴坐标值, 并且  $x_f = s * x$ ;  $y$  为人脸候选区域的左上角的  $y$  轴坐标值,  $y_f$  为对应候选区域的左上角的  $y$  轴坐标值, 并且  $y_f = s * y$ ;  $w$  为人脸候选区域的宽度,  $w_f$  为对应候选区域的宽度, 并且  $w_f = s * w$ ;  $h$  为人脸候选区域的高度,  $h_f$  为对应候选区域的高度, 并且  $h_f = s * h$ ;  $s$  为所述图像的尺寸到所述特征图的尺寸的缩放系数;

所述识别结果还包括人脸概率;

相应的, 在所述将各对应候选区域的特征向量输入至多任务神经网络中进行识别, 获取各对应候选区域的识别结果之后, 还包括:

采用交除并参数为 0.5 的非最大值抑制方法, 对根据各对应候选区域的识别结果中的

人脸概率进行处理,以保留交除并参数大于0.5中的人脸概率最大的各对应候选区域;

去除交除并参数大于0.5中的人脸概率最大的各对应候选区域中,人脸概率小于0.8的对应候选区域,以确定过滤后的各对应候选区域的识别结果。

## 基于图像的头部姿态检测方法、装置以及终端

### 技术领域

[0001] 本公开涉及图像技术领域,尤其涉及基于图像的头部姿态检测方法、装置以及终端。

### 背景技术

[0002] 随着人脸识别技术的发展,头部姿态检测已经成为动机检测、人脸识别研究中的一项关键技术。头部姿态可以反映出人的思维意识,以及人脸朝向和眼睛的注视方向,因此头部姿态检测也是人际交互、心理意识理解领域中的重要组成部分。

[0003] 相关技术中,头部姿态的检测需要先进行人脸检测,以确定头位置信息,然后采用3D模型去模拟人脸,以进行头部姿态参数的计算,从而完成头部姿态的检测。

[0004] 然而相关技术中,由于先计算出头部位置信息,再进行头部姿态参数的计算,这样的检测方式分成了两个步骤完成头部姿态的检测,检测速度较慢,检测效率低。

### 发明内容

[0005] 为克服相关技术中存在的问题,本公开提供一种基于图像的头部姿态检测方法、装置以及终端,用于解决相关技术中,头部姿态检测的检测速度较慢,检测效率低的问题。

[0006] 根据本公开实施例的第一方面,提供一种基于图像的头部姿态检测方法,包括:

[0007] 获取图像中的各人脸候选区域,其中,所述人脸候选区域为特征向量;

[0008] 对所述图像进行运算处理,确定与所述图像对应的特征图;

[0009] 将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;

[0010] 将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数。

[0011] 进一步地,所述获取图像中的各人脸候选区域,包括:

[0012] 采用显著性检测方法,对图像进行检测,以获取图像中的各人脸候选区域。

[0013] 进一步地,所述对所述图像进行运算处理,确定与所述图像对应的特征图,包括:

[0014] 对所述图像进行N层的卷积神经网络(Convolutional Neural Networks,简称CNN)卷积运算,以确定与所述图像对应的特征图,其中,N为正整数。

[0015] 进一步地,所述对应候选区域为 $R_f = (x_f, y_f, w_f, h_f)$ ;

[0016] 其中, $R_f = s * R$ , $R = (x, y, w, h)$ 为人脸候选区域,x为人脸候选区域的左上角的x轴坐标值,y为人脸候选区域的左上角的y轴坐标值,w为人脸候选区域的宽度,h为人脸候选区域的高度,s为所述图像的尺寸到所述特征图的尺寸的缩放系数。

[0017] 进一步地,在所述将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域之后,还包括:

[0018] 采用自适应池化(Region Of Interest Pooling,简称ROI Pooling)方法对各对应候选区域进行处理,以将各对应候选区域映射为相同尺寸的各对应候选区域。

- [0019] 进一步地,所述识别结果还包括人脸概率;
- [0020] 相应的,在所述将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果之后,还包括:
- [0021] 采用交除并参数为0.5的非最大值抑制方法,对根据各对应候选区域的识别结果中的人脸概率进行处理,以保留交除并参数大于0.5中的人脸概率最大的各对应候选区域;
- [0022] 去除交除并参数大于0.5中的人脸概率最大的各对应候选区域中,人脸概率小于0.8的对应候选区域,以确定过滤后的各对应候选区域的识别结果。
- [0023] 本公开的实施例提供的技术方案可以包括以下有益效果:
- [0024] 本实施例通过获取图像中的各人脸候选区域;对所述图像进行运算处理,确定与所述图像对应的特征图;将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果中的头部姿态参数。从而可以通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。
- [0025] 根据本公开实施例的第二方面,提供一种基于图像的头部姿态检测装置,包括:
- [0026] 获取模块,被配置为获取图像中的各人脸候选区域,其中,所述人脸候选区域为特征向量;
- [0027] 确定模块,被配置为对所述图像进行运算处理,确定与所述图像对应的特征图;
- [0028] 映射模块,被配置为将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;
- [0029] 识别模块,被配置为将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数。
- [0030] 进一步地,所述获取模块,被具体配置为:
- [0031] 采用显著性检测方法,对图像进行检测,以获取图像中的各人脸候选区域。
- [0032] 进一步地,所述确定模块,被具体配置为:
- [0033] 对所述图像进行N层的CNN卷积运算,以确定与所述图像对应的特征图,其中,N为正整数。
- [0034] 进一步地,所述对应候选区域为 $R_f = (x_f, y_f, w_f, h_f)$ ;
- [0035] 其中, $R_f = s * R$ , $R = (x, y, w, h)$ 为人脸候选区域,x为人脸候选区域的左上角的x轴坐标值,y为人脸候选区域的左上角的y轴坐标值,w为人脸候选区域的宽度,h为人脸候选区域的高度,s为所述图像的尺寸到所述特征图的尺寸的缩放系数。
- [0036] 进一步地,所述装置,还包括:
- [0037] 处理模块,被配置为在所述映射模块将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域之后,采用ROI Pooling方法对各对应候选区域进行处理,以将各对应候选区域映射为相同尺寸的各对应候选区域。
- [0038] 进一步地,所述识别结果还包括人脸概率;
- [0039] 相应的,所述装置,还包括:
- [0040] 筛选模块,被配置为在所述识别模块将各对应候选区域的特征向量输入至多任务

神经网络中进行识别,获取各对应候选区域的识别结果之后,采用交除并参数为0.5的非最大值抑制方法,对根据各对应候选区域的识别结果中的人脸概率进行处理,以保留交除并参数大于0.5中的人脸概率最大的各对应候选区域;去除交除并参数大于0.5中的人脸概率最大的各对应候选区域中,人脸概率小于0.8的对应候选区域,以确定过滤后的各对应候选区域的识别结果。

[0041] 本公开的实施例提供的技术方案可以包括以下有益效果:

[0042] 本实施例通过获取图像中的各人脸候选区域;对所述图像进行运算处理,确定与所述图像对应的特征图;将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果中的头部姿态参数。从而可以通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。

[0043] 根据本公开实施例的第三方面,提供一种终端,包括:

[0044] 处理器,以及用于存储处理器可执行指令的存储器;

[0045] 其中,所述处理器被配置为:获取图像中的各人脸候选区域,其中,所述人脸候选区域为特征向量;对所述图像进行运算处理,确定与所述图像对应的特征图;将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数。

[0046] 本公开的实施例提供的技术方案可以包括以下有益效果:

[0047] 本实施例通过获取图像中的各人脸候选区域;对所述图像进行运算处理,确定与所述图像对应的特征图;将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果中的头部姿态参数。从而可以通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。

[0048] 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,并不能限制本公开。

## 附图说明

[0049] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本发明的实施例,并与说明书一起用于解释本发明的原理。

[0050] 图1是根据一示例性实施例示出的一种基于图像的头部姿态检测方法实施例一的流程图;

[0051] 图2是根据一示例性实施例示出的一种基于图像的头部姿态检测方法实施例二的流程图;

[0052] 图3是根据一示例性实施例示出的一种基于图像的头部姿态检测装置实施例三的

框图；

[0053] 图4是根据一示例性实施例示出的一种基于图像的头部姿态检测装置实施例四的框图；

[0054] 图5是根据一示例性实施例示出的一种终端的实体的框图；

[0055] 图6是根据一示例性实施例示出的一种终端设备800的框图。

### 具体实施方式

[0056] 这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本发明相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本发明的一些方面相一致的装置和方法的例子。

[0057] 图1是根据一示例性实施例示出的一种基于图像的头部姿态检测方法实施例一的流程图，如图1所示，该方法主要是用在终端中，该终端包括公共终端、移动终端、个人终端等等，例如台式机、触屏手机、平板电脑、个人电脑等，该基于图像的头部姿态检测方法包括以下步骤。

[0058] 在步骤S11中，获取图像中的各人脸候选区域，其中，所述人脸候选区域为特征向量。

[0059] 其中，步骤S11的具体实现方式为：采用显著性检测方法，对图像进行检测，以获取图像中的各人脸候选区域。

[0060] 在本步骤中，获取图像，然后去获取图像中的各人脸候选区域，一般来说，一幅图像中可以获取到几百到几千个的人脸候选区域。

[0061] 具体来说，可以采用显著性检测方法对图像进行检测，常见的显著性检测方法有 Selective Search, 全卷积神经网络 (Region Proposal Network, 简称RPN) 等方法，在检测之后可以得到各人脸候选区域。这些人脸候选区域中有一些可能包括了人脸，有一些可能不包括人脸。其中，人脸候选区域为特征向量，可以采用  $R = (x, y, w, h)$  表示，其中， $x$  为人脸候选区域的左上角的  $x$  轴坐标值， $y$  为人脸候选区域的左上角的  $y$  轴坐标值， $w$  为人脸候选区域的宽度， $h$  为人脸候选区域的高度。

[0062] 举例来说，获取一幅图像，然后采用 Selective Search 方法对该图像进行检测之后，可以得到多个人脸候选区域。

[0063] 在步骤S12中，对所述图像进行运算处理，确定与所述图像对应的特征图。

[0064] 其中，步骤S11的具体实现方式为：对所述图像进行  $N$  层的 CNN 卷积运算，以确定与所述图像对应的特征图，其中， $N$  为正整数。

[0065] 在本步骤中，对步骤S11中获取到的图像进行运算处理，进而可以计算出与图像对应的特征图，此时，针对一幅图像可以得到一幅特征图。

[0066] 具体来说，对步骤S11中获取到的图像，进行  $N$  层的 CNN 卷积运算，从而可以获取到与图像对应的特征图。其中， $N$  为正整数。此时，CNN 卷积的层数，以及每一层的卷积核大小和数量，可以根据实际情况进行调整，一般来说，根据算法速度和算法精度，调整 CNN 卷积的层数，以及每一层的卷积核大小和数量；如果想提高算法速度，可以减少 CNN 卷积的层数和每一层卷积核的数量。



[0067] 举例来说,可以采用10层的CNN卷积运算,对图像进行处理,进而计算出与图像对应的特征图。

[0068] 在步骤S13中,将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域。

[0069] 其中,所述对应候选区域为 $R_f = (x_f, y_f, w_f, h_f)$ ;  $R_f = s * R$ ,  $R = (x, y, w, h)$  为人脸候选区域, $x$ 为人脸候选区域的左上角的 $x$ 轴坐标值, $y$ 为人脸候选区域的左上角的 $y$ 轴坐标值, $w$ 为人脸候选区域的宽度, $h$ 为人脸候选区域的高度, $s$ 为所述图像的尺寸到所述特征图的尺寸的缩放系数。

[0070] 在本步骤中,将步骤S11得到的各人脸候选区域,分别一一映射到步骤S12得到的特征图中,得到映射后的结果,此时映射后的结果为各人脸候选区域在特征图中的各对应候选区域 $R_f = (x_f, y_f, w_f, h_f)$ 。

[0071] 其中, $R_f = s * R$ ,  $R = (x, y, w, h)$  为人脸候选区域。由于 $x$ 为人脸候选区域的左上角的 $x$ 轴坐标值,则 $x_f$ 为对应候选区域的左上角的 $x$ 轴坐标值,并且 $x_f = s * x$ ;由于 $y$ 为人脸候选区域的左上角的 $y$ 轴坐标值,则 $y_f$ 为对应候选区域的左上角的 $y$ 轴坐标值,并且 $y_f = s * y$ ;由于 $w$ 为人脸候选区域的宽度,则 $w_f$ 为对应候选区域的宽度,并且 $w_f = s * w$ ;由于 $h$ 为人脸候选区域的高度,则 $h_f$ 为对应候选区域的高度,并且 $h_f = s * h$ 。 $s$ 为图像的尺寸到特征图的尺寸的缩放系数。

[0072] 在步骤S14中,将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数。

[0073] 在本步骤中,通过学习训练已经得到了成熟的多任务神经网络,由于人脸候选区域为特征向量,则得到的与人脸候选区域对应的对应候选区域也是特征向量。

[0074] 然后可以将步骤S13中各对应候选区域的特征向量输入至多任务神经网络中进行识别,各对应候选区域的识别结果。其中,一个对应候选区域的识别结果中包括了头部姿态参数、人脸框的位置信息。多任务神经网络由两个全连接层、以及3个任务相关的输出层组成。

[0075] 头部姿态是某一时刻头部的特写,也是人的头部相对某个坐标轴偏转程度的一种状态。头部姿态检测,从姿态估计精度上看,可以把头部姿态检测方式分为两大类:粗估计方式和细估计方式;其中,粗估计方式中,可以粗略地估计出人的头部偏转方向,例如估计出头部向左偏转、或头部向上偏转等等;细估计方式中,可以在三维空间上对头部偏转的精确度量,从而对头部相对于某一坐标轴的偏转角度进行推断。

[0076] 在本实施中,得到的细估计方式检测到的结果,从而本实施例的识别结果中的头部姿态参数包括了参数(yaw, pitch, roll);其中,pitch方向为X轴,围绕于X轴的头部姿态范围为 $-60.4^{\circ} \sim 69.6^{\circ}$ ;yaw方向为Y轴,围绕Y轴的头部姿态范围为 $-40.9^{\circ} \sim 36.3^{\circ}$ ;roll方向为Z轴,围绕于Z轴的头部姿态范围为 $-79.8^{\circ} \sim 75.37^{\circ}$ 。

[0077] 由于头部姿态可以反映一个人眼睛注视方向或注意力方向;具体来说,当人眼被遮挡时,通过头部姿态检测得到的头部姿态参数,可以大致辨别出被观察者的注视方向;当人眼未被遮挡时,通过头部姿态检测得到的头部姿态参数,可以成为精确预测被观察者注视方向的必要条件。

[0078] 本实施例通过获取图像中的各人脸候选区域;对所述图像进行运算处理,确定与

所述图像对应的特征图;将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果中的头部姿态参数。从而可以通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。

[0079] 图2是根据一示例性实施例示出的一种基于图像的头部姿态检测方法实施例二的流程图,如图2所示,在步骤S13之后,还包括以下步骤:

[0080] 在步骤S15中,采用ROI Pooling方法对各对应候选区域进行处理,以将各对应候选区域映射为相同尺寸的各对应候选区域。

[0081] 在本步骤中,在步骤S13得到各对应候选区域之后,需要采用ROI Pooling方法对各对应候选区域进行处理,在处理过程中,会将不同尺寸的各对应候选区域映射为相同尺寸的各对应候选区域。举例来说,可以将不同尺寸的各对应候选区域,映射为特征向量的长度为1024的各对应候选区域。

[0082] 将各对应候选区域映射为相同尺寸的各对应候选区域,是由于后续的多任务神经网络的识别中,需要各对应候选区域的特征向量的长度相同。

[0083] 所述识别结果还包括人脸概率;相应的,在步骤S14之后,还包括以下步骤:

[0084] 在步骤S16中,采用交除并参数为0.5的非最大值抑制方法,对根据各对应候选区域的识别结果中的人脸概率进行处理,以保留交除并参数大于0.5中的人脸概率最大的各对应候选区域;去除交除并参数大于0.5中的人脸概率最大的各对应候选区域中,人脸概率小于0.8的对应候选区域,以确定过滤后的各对应候选区域的识别结果。

[0085] 在本步骤中,在步骤S14中得到的识别结果中还包括人脸概率。首先,需要采用非最大值抑制方法,对各对应候选区域的识别结果进行过滤,具体来说,设置非最大值抑制方法的交除并参数(Intersection-over-union,简称IOU)为0.5,然后采用交除并参数为0.5的非最大值抑制方法,对根据各对应候选区域的识别结果中的人脸概率进行过滤处理,只保留交除并参数大于0.5中的人脸概率最大的各对应候选区域。

[0086] 然后,针对交除并参数大于0.5中的人脸概率最大的各对应候选区域,去人脸概率小于0.8的对应候选区域,从而德奥最终的过滤后的各对应候选区域、以及各对应候选区域的识别结果。进而得到最终的人脸位置坐标和头部姿态参数。

[0087] 本实施例将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域之后,采用ROI Pooling方法对各对应候选区域进行处理,以将各对应候选区域映射为相同尺寸的各对应候选区域;然后将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果中的头部姿态参数;最终采用非最大值抑制方法、人脸概率小于0.8的过滤方式,对各对应候选区域的识别结果的进行过滤,得到最终的结果。从而可以通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。

[0088] 图3是根据一示例性实施例示出的一种基于图像的头部姿态检测装置实施例三的

框图。参照图3,该装置包括获取模块31、确定模块32、映射模块33和识别模块34。

[0089] 该获取模块31,被配置为获取图像中的各人脸候选区域,其中,所述人脸候选区域为特征向量;

[0090] 该确定模块32,被配置为对所述图像进行运算处理,确定与所述图像对应的特征图;

[0091] 该映射模块33,被配置为将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;

[0092] 该识别模块34,被配置为将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数。

[0093] 关于上述实施例中的装置,其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述,此处将不做详细阐述说明。

[0094] 本实施例通过获取图像中的各人脸候选区域;对所述图像进行运算处理,确定与所述图像对应的特征图;将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果中的头部姿态参数。从而可以通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。

[0095] 在上述图3所示的实施例的基础上,图4是根据一示例性实施例示出的一种基于图像的头部姿态检测装置实施例四的框图。参照图4,所述获取模块31,被具体配置为:

[0096] 采用显著性检测方法,对图像进行检测,以获取图像中的各人脸候选区域。

[0097] 所述确定模块32,被具体配置为:

[0098] 对所述图像进行N层的CNN卷积运算,以确定与所述图像对应的特征图,其中,N为正整数。

[0099] 所述对应候选区域为 $R_f = (x_f, y_f, w_f, h_f)$ ;其中, $R_f = s * R$ , $R = (x, y, w, h)$ 为人脸候选区域, $x$ 为人脸候选区域的左上角的 $x$ 轴坐标值, $y$ 为人脸候选区域的左上角的 $y$ 轴坐标值, $w$ 为人脸候选区域的宽度, $h$ 为人脸候选区域的高度, $s$ 为所述图像的尺寸到所述特征图的尺寸的缩放系数。

[0100] 所述装置,还包括:

[0101] 处理模块41,被配置为在所述映射模块33将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域之后,采用ROI Pooling方法对各对应候选区域进行处理,以将各对应候选区域映射为相同尺寸的各对应候选区域。

[0102] 所述识别结果还包括人脸概率;相应的,所述装置,还包括:

[0103] 筛选模块42,被配置为在所述识别模块34将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果之后,采用交除并参数为0.5的非最大值抑制方法,对根据各对应候选区域的识别结果中的人脸概率进行处理,以保留交除并参数大于0.5中的人脸概率最大的各对应候选区域;去除交除并参数大于0.5中的人脸概率最大的各对应候选区域中,人脸概率小于0.8的对应候选区域,以确定过滤后的各对应

候选区域的识别结果。

[0104] 关于上述实施例中的装置,其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述,此处将不做详细阐述说明。

[0105] 本实施例将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域之后,采用ROI Pooling方法对各对应候选区域进行处理,以将各对应候选区域映射为相同尺寸的各对应候选区域;然后将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果中的头部姿态参数;最终采用非最大值抑制方法、人脸概率小于0.8的过滤方式,对各对应候选区域的识别结果的进行过滤,得到最终的结果。从而可以通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。

[0106] 图5是根据一示例性实施例示出的一种终端的实体的框图。参照图5,该终端可以具体实现为:处理器71,以及被配置为存储处理器可执行指令的存储器72;

[0107] 其中,所述处理器71被配置为:获取图像中的各人脸候选区域,其中,所述人脸候选区域为特征向量;对所述图像进行运算处理,确定与所述图像对应的特征图;将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数。

[0108] 在上述实施例中,应理解,该处理器可以是中央处理单元(英文:Central Processing Unit,简称:CPU),还可以是其他通用处理器、数字信号处理器(英文:Digital Signal Processor,简称:DSP)、专用集成电路(英文:Application Specific Integrated Circuit,简称:ASIC)等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,而前述的存储器可以是只读存储器(英文:read-only memory,缩写:ROM)、随机存取存储器(英文:random access memory,简称:RAM)、快闪存储器、硬盘或者固态硬盘。SIM卡也称为用户身份识别卡、智能卡,数字移动电话机必须装上此卡方能使用。即在电脑芯片上存储了数字移动电话客户的信息,加密的密钥以及用户的电话簿等内容。结合本发明实施例所公开的方法的步骤可以直接体现为硬件处理器执行完成,或者用处理器中的硬件及软件模块组合执行完成。

[0109] 关于上述实施例中的终端,其中各个模块执行操作的具体方式已经在有关该方法和装置的实施例中进行了详细描述,此处将不做详细阐述说明。

[0110] 本实施例通过获取图像中的各人脸候选区域;对所述图像进行运算处理,确定与所述图像对应的特征图;将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果中的头部姿态参数。从而可以通过一次识别就可以得到头部姿态参数,同时可以得到人脸的位置,进而提高了头部姿态检测的速度,提高了头部姿态检测的检测效率;同时,通过多任务神经网络的识别,可以消除现有检测方法中任务之间的累积误差,进而提高检测的准确性。

[0111] 图6是根据一示例性实施例示出的一种终端设备800的框图。例如,终端设备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。

[0112] 参照图6,终端设备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/O)的接口812,传感器组件814,以及通信组件816。

[0113] 处理组件802通常控制终端设备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。

[0114] 存储器804被配置为存储各种类型的数据以支持在终端设备800的操作。这些数据的示例包括用于在终端设备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。

[0115] 电源组件806为终端设备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,及其他与为终端设备800生成、管理和分配电力相关联的组件。

[0116] 多媒体组件808包括在所述终端设备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。当终端设备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。

[0117] 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),当终端设备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。

[0118] I/O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。

[0119] 传感器组件814包括一个或多个传感器,用于为终端设备800提供各个方面的状态评估。例如,传感器组件814可以检测到终端设备800的打开/关闭状态,组件的相对定位,例如所述组件为终端设备800的显示器和小键盘,传感器组件814还可以检测终端设备800或终端设备800一个组件的位置改变,用户与终端设备800接触的存在或不存在,终端设备800方位或加速/减速和终端设备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CMOS或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。

[0120] 通信组件816被配置为便于终端设备800和其他设备之间有线或无线方式的通信。终端设备800可以接入基于通信标准的无线网络,如WiFi,2G或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。

[0121] 在示例性实施例中,终端设备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。

[0122] 在示例性实施例中,还提供了一种包括指令的非临时性计算机可读存储介质,例如包括指令的存储器804,上述指令可由终端设备800的处理器820执行以完成上述方法。例如,所述非临时性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。

[0123] 一种非临时性计算机可读存储介质,当所述存储介质中的指令由移动终端的处理器执行时,使得移动终端能够执行一种基于图像的头部姿态检测方法,所述方法包括:

[0124] 获取图像中的各人脸候选区域,其中,所述人脸候选区域为特征向量;

[0125] 对所述图像进行运算处理,确定与所述图像对应的特征图;

[0126] 将各人脸候选区域分别一一映射至所述特征图上,确定各人脸候选区域在所述特征图中的各对应候选区域;

[0127] 将各对应候选区域的特征向量输入至多任务神经网络中进行识别,获取各对应候选区域的识别结果,其中,所述识别结果包括头部姿态参数。

[0128] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本发明的其它实施方案。本申请旨在涵盖本发明的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本发明的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本发明的真正范围和精神由下面的权利要求指出。

[0129] 应当理解的是,本发明并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本发明的范围仅由所附的权利要求来限制。

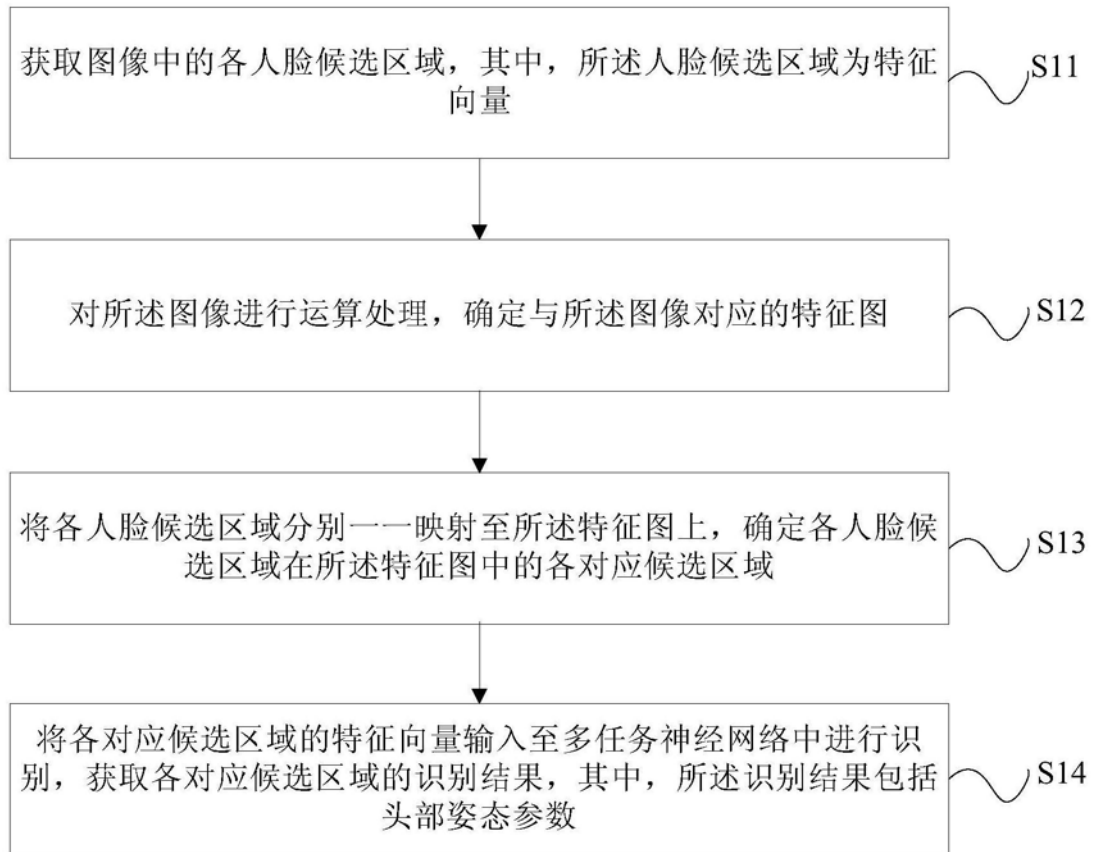


图1

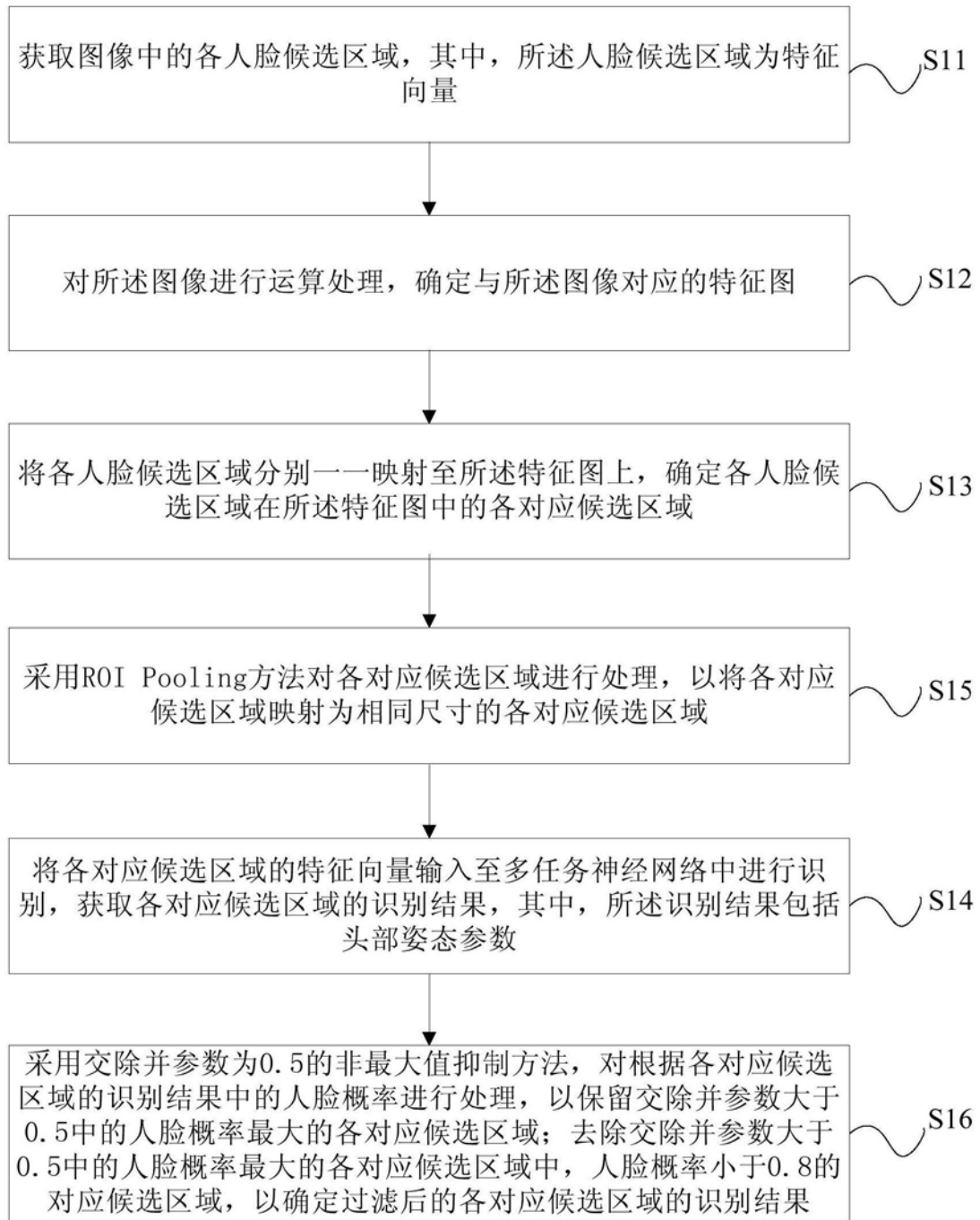


图2



图3



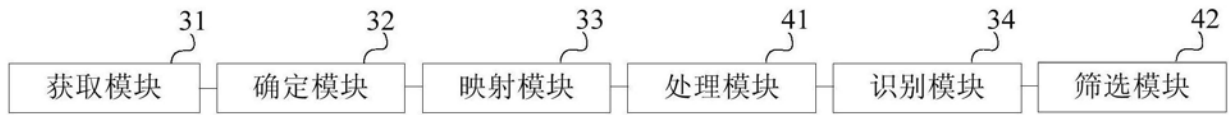


图4

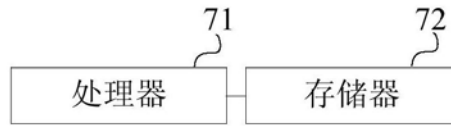


图5

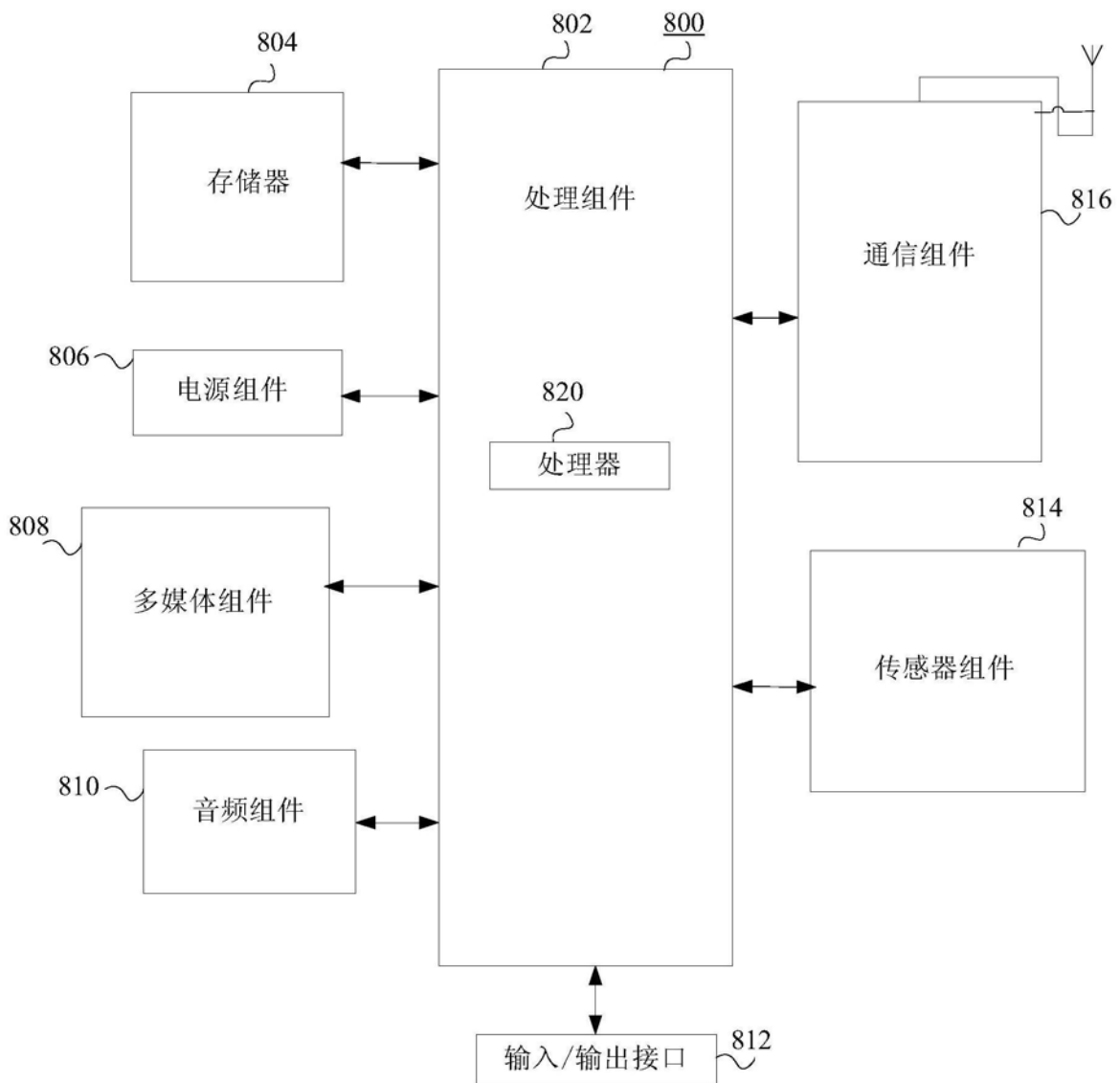


图6