



(12) 发明专利申请

(10) 申请公布号 CN 115661872 A

(43) 申请公布日 2023. 01. 31

(21) 申请号 202211292529.6

G06V 10/764 (2022.01)

(22) 申请日 2022.10.21

G06V 10/766 (2022.01)

G06V 10/774 (2022.01)

(71) 申请人 深圳市人工智能与机器人研究院
地址 518129 广东省深圳市龙岗区坂田街
道雅宝路1号星河WORLD G2-14层
申请人 香港中文大学(深圳)

G06V 10/80 (2022.01)

G06V 10/82 (2022.01)

(72) 发明人 梁旭 陈骏安 张大鹏

(74) 专利代理机构 深圳市君胜知识产权代理事
务所(普通合伙) 44268
专利代理师 朱阳波

(51) Int. Cl.

G06V 40/12 (2022.01)

G06V 10/25 (2022.01)

G06V 10/26 (2022.01)

G06V 10/24 (2022.01)

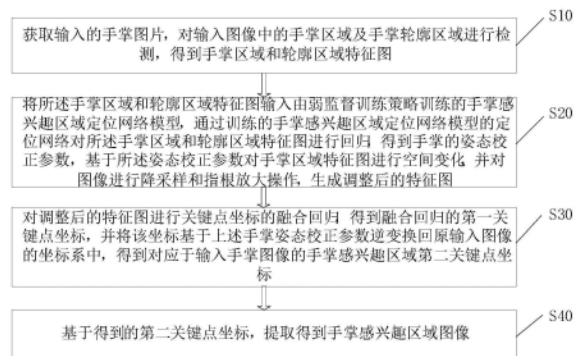
权利要求书3页 说明书20页 附图8页

(54) 发明名称

一种自然场景下鲁棒的手掌感兴趣区域定位方法

(57) 摘要

本发明公开了自然场景下鲁棒的手掌感兴趣区域定位方法,包括:获取输入图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;将特征图输入手掌感兴趣区域定位网络模型,通过其中的定位网络对特征图进行回归,得到姿态校正参数后对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大,生成调整后的特征图;对调整后的特征图进行关键点坐标融合回归,得到第一关键点坐标,将该坐标逆变换回原输入图像的坐标系中,得到第二关键点坐标;基于第二关键点坐标,提取得到手掌感兴趣区域图像。本发明无需红外成像去除背景,能够处理大范围内平移、旋转等的问题,实现自然场景下的非接触式手掌感兴趣区域定位需求。



1. 一种自然场景下鲁棒的手掌感兴趣区域定位方法,其特征在于,所述方法包括:

获取输入的手掌图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;

将所述手掌区域和轮廓区域特征图输入由弱监督训练策略训练的手掌感兴趣区域定位网络模型,通过训练的手掌感兴趣区域定位网络模型的定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大操作,生成调整后的特征图;

对调整后的特征图进行关键点坐标的融合回归,得到融合回归的第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标;

基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像。

2. 根据权利要求1所述自然场景下鲁棒的手掌感兴趣区域定位方法,其特征在于,所述弱监督训练策略训练手掌感兴趣区域定位网络模型的训练步骤包括:

通过手掌区域弱分割器,采用由单色背景手掌图像集训练得到的二分类器对输入的第一手掌图像中的像素进行分类,提取手掌区域,对手掌区域执行边缘检测得到手掌轮廓;

通过迭代手掌感兴趣区域定位算法处理手掌轮廓,得到输入的第一手掌图像的第一标注信息;

基于第一手掌图像的手掌区域提取手掌像素,结合手掌像素和自然图库中的自然背景图像进行背景合成,生成自然环境下的第二手掌图像;

对上述得到的第二手掌图像和第一标注信息通过同时执行相同的随机空间变化操作,得到第三手掌图像和第二标注信息,再对第三手掌图像执行图像质量随机扰动得到第四手掌图像,得到数据增广后包含有第四手掌图像和第二标注信息的训练样本;

通过得到的训练样本对所述手掌感兴趣区域定位网络模型进行训练。

3. 根据权利要求2所述自然场景下鲁棒的手掌感兴趣区域定位方法,其特征在于,所述迭代手掌感兴趣区域定位算法,具体包括:

对提取到的手掌区域进行关键点的检测,并校验关键点的个数;

基于校验后的关键点个数,当检测到完整的5个指尖点和4个指谷点则执行基于距离极值点的手掌感兴趣区域定位方法;

当校验未得到完整的5个指尖点和4个指谷点时则进行基于线扫描的手指边缘检测,当检测到4个指尖点及3个指谷点时则执行基于线扫描的手掌感兴趣区域定位方法;

当未检测到4个指尖点及3个指谷点时则移动到下一个扫描点执行基于线扫描的手指边缘检测,通过检测再次判断是否扫描到4个指尖点及3个指谷点,当检测到4个指尖点及3个指谷点时则执行基于线扫描的手掌感兴趣区域定位方法,当未检测到4个指尖点及3个指谷点时移动到下一个扫描位置再次执行基于线扫描的手指边缘检测;

当扫描触发停止条件时退出检测。

4. 根据权利要求3所述自然场景下鲁棒的手掌感兴趣区域定位方法,其特征在于,所述迭代手掌感兴趣区域定位算法和所述基于线扫描的手指边缘检测,具体还包括:

所述迭代手掌感兴趣区域定位算法通过定位方法得到手掌感兴趣区域后对所述手掌感兴趣区域进行异常检测;

所述基于线扫描的手指边缘检测在输入的手掌区域中由上到下竖直检测手掌区域图的亮度值,当完整出现了变化模式时即检测到了正常的四指手指边缘,当没有检测到完整的变化模式,则本次竖直扫描失败,向右移动一个步长,继续进行下一次竖直扫描;

当扫描超过了预先设定的扫描区域时,则触发终止扫描条件,退出扫描检测。

5. 根据权利要求1所述自然场景下鲁棒的手掌感兴趣区域定位方法,其特征在于,所述获取输入图片,对所述输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图具体包括:

通过具有多尺度金字塔结构的主干网络提取图片中的初级特征;

基于所获取的初级特征,采用语义分割网络对初级特征进行处理生成包含有背景区域图、手掌区域图、手掌轮廓区域图三个通道的区域特征图,且所述三通道输出通过焦点损失对其语义分割效果进行监督;

生成的三通道区域特征图经过预处理后生成含有六个通道的特征图。

6. 根据权利要求1所述自然场景下鲁棒的手掌感兴趣区域定位方法,其特征在于,所述对调整后的特征图进行关键点坐标的融合回归,具体包括:

通过多头注意力网络对关键点特征进行注意力感知及提取,再通过多层感知机对关键点的坐标进行全局回归,得到全局回归所预测的关键点坐标;

通过Vision Transformer网络对调整后的特征图进行注意力感知及特征提取,再通过CNN卷积网络对局部关键点进行检测,最后通过可微空间到数值变换将关键点响应映射为局部回归的关键点坐标;

通过由当前样本特征预测得到的融合权重对全局回归得到的关键点坐标和局部回归得到的关键点坐标进行融合,得到融合回归的第一关键点坐标;

对指根边缘距离损失进行计算,所述指根边缘距离损失定义为: $L_{edge} = 1/2 \sum_{i=1}^2 \min(d_i)$, d_i 是对应第*i*条指根边缘与第*i*个指根关键点的欧式距离序列, L_{edge} 为指根边缘距离损失。

7. 根据权利要求1所述自然场景下鲁棒的手掌感兴趣区域定位方法,其特征在于,所述基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像,具体包括:

根据预先设置的阈值对训练样本进行判断;

当样本规模达到阈值时选择使用第二关键点坐标输出手掌感兴趣区域角点的坐标;

当样本规模未达到阈值时选择使用第二关键点坐标输出手掌感兴趣区域中心点与指根关键点坐标来建立坐标系并定位正方形手掌感兴趣区域;

对获取到的手掌感兴趣区域角点坐标执行射影变换提取手掌感兴趣区域图像。

8. 一种自然场景下鲁棒的手掌感兴趣区域定位装置,其特征在于,所述装置包括:

手掌区域及轮廓提取模块,用于获取输入的手掌图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;

手掌姿态调整模块,用于将所述手掌区域和轮廓区域特征图输入由弱监督训练策略训练的手掌感兴趣区域定位网络模型,通过训练的手掌感兴趣区域定位网络模型的定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大操作,生成调整后的特征图;

手掌关键点坐标融合回归模块,用于对调整后的特征图进行关键点坐标的融合回归,

得到融合回归的第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标;

手掌感兴趣区域图像提取模块,用于基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像。

9. 一种特征提取识别器,其特征在于,所述特征提取器对权利要求1-7任一项所述的自然场景下鲁棒的手掌感兴趣区域定位方法获取的手掌感兴趣区域图像进行手掌生物信息检测,基于检测到的手掌生物信息实现用户身份识别及核验。

10. 一种终端设备,其特征在于,所述终端设备包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的自然场景下鲁棒的手掌感兴趣区域定位程序,所述处理器执行所述自然场景下鲁棒的手掌感兴趣区域定位程序时,实现如权利要求1-7任一项所述的自然场景下鲁棒的手掌感兴趣区域定位方法的步骤。

一种自然场景下鲁棒的手掌感兴趣区域定位方法

技术领域

[0001] 本发明涉及图像处理领域,尤其涉及一种自然场景下鲁棒的手掌感兴趣区域定位方法、装置、特征提取识别器及设备。

背景技术

[0002] 掌纹识别技术能够根据人体手掌表面的生物特征自动地对用户的身份进行鉴别,掌纹识别包括如下几个部分:手掌图像采集、掌纹感兴趣区域定位、掌纹特征提取、特征匹配及识别。对于掌纹感兴趣区域定位而言,其背景较为复杂、手掌姿态变化较多,因此,手掌感兴趣区域的定位成为了识别系统的瓶颈所在。

[0003] 然而,当前掌纹感兴趣区域定位方案较多依赖红外成像去除背景,且无法处理大范围内的手掌旋转、平移、缩放的问题,即无法实现自然场景下的非接触式手掌感兴趣区域定位需求,对用户而言较为不便,这是目前亟需解决的问题。

[0004] 因此,现有技术还有待于改进和发展。

发明内容

[0005] 本申请要解决的技术问题在于,针对现有技术的不足,提供一种自然场景下鲁棒的手掌感兴趣区域定位方法装置、设备及介质,本发明能够通过深度网络得到手掌相关区域特征图,区分手掌背景图和手掌相关区域特征图,同时对得到的特征图进行姿态校正并通过坐标回归得到手掌区域关键点坐标,从而由坐标实现对手掌高清感兴趣区域ROI图像的提取,使得本发明对复杂的背景和自然的手掌姿态具有鲁棒性,能够从不同设备在自然场景中所拍摄的手掌图像中成功地定位ROI。

[0006] 为了解决上述现有技术问题的不足,本申请实施例第一方面提供了一种自然场景下鲁棒的手掌感兴趣区域定位方法,所述方法包括:

[0007] 获取输入的手掌图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;

[0008] 将所述手掌区域和轮廓区域特征图输入由弱监督训练策略训练的手掌感兴趣区域定位网络模型,通过训练的手掌感兴趣区域定位网络模型的定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大操作,生成调整后的特征图;

[0009] 对调整后的特征图进行关键点坐标的融合回归,得到融合回归的第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标;

[0010] 基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像。

[0011] 所述弱监督训练策略训练手掌感兴趣区域定位网络模型的训练步骤包括:通过手掌区域弱分割器,采用由单色背景手掌图像集训练得到的二分类器对输入的第一手掌图像中的像素进行分类,提取手掌区域,并对手掌区域执行边缘检测得到手掌轮廓;

- [0012] 通过迭代手掌感兴趣区域定位算法处理手掌轮廓,得到输入的第一手掌图像的第一标注信息;
- [0013] 基于第一手掌图像的手掌区域提取手掌像素,结合手掌像素和自然图库中的自然背景图像进行背景合成,生成自然环境下的第二手掌图像;
- [0014] 对上述得到的第二手掌图像和第一标注信息通过同时执行相同的随机空间变化操作,得到第三手掌图像和第二标注信息,再对第三手掌图像执行图像质量随机扰动得到第四手掌图像,得到数据增广后包含有第四手掌图像和第二标注信息的训练样本;
- [0015] 通过得到的训练样本对所述手掌感兴趣区域定位网络模型进行训练。
- [0016] 所述迭代手掌感兴趣区域定位算法,具体包括:
- [0017] 对提取到的手掌区域进行关键点的检测,并校验关键点的个数;
- [0018] 基于校验后的关键点个数,当检测到完整的5个指尖点和4个指谷点则执行基于距离极值点的手掌感兴趣区域定位方法;
- [0019] 当校验未得到完整的5个指尖点和4个指谷点时则进行基于线扫描的手指边缘检测,当检测到4个指尖点及3个指谷点时则执行基于线扫描的手掌感兴趣区域定位方法;
- [0020] 当未检测到4个指尖点及3个指谷点时则移动到下一个扫描点执行基于线扫描的手指边缘检测,通过检测再次判断是否扫描到4个指尖点及3个指谷点,当检测到4个指尖点及3个指谷点时则执行基于线扫描的手掌感兴趣区域定位方法,当未检测到4个指尖点及3个指谷点时移动到下一个扫描位置再次执行基于线扫描的手指边缘检测;
- [0021] 当扫描触发停止条件时退出检测。
- [0022] 所述获取输入图片,对所述输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图具体包括:
- [0023] 通过具有多尺度金字塔结构的主干网络提取图片中的初级特征;
- [0024] 基于所获取的初级特征,采用语义分割网络对初级特征进行处理生成包含有背景区域图、手掌区域图、手掌轮廓区域图三个通道的区域特征图,且所述三通道输出通过焦点损失对其语义分割效果进行监督;
- [0025] 生成的三通道区域特征图经过预处理后生成含有六个通道的特征图。
- [0026] 所述迭代手掌感兴趣区域定位算法和所述基于线扫描的手指边缘检测,具体还包括:
- [0027] 所述迭代手掌感兴趣区域定位算法通过定位方法得到手掌感兴趣区域后对所述手掌感兴趣区域进行异常检测;
- [0028] 所述基于线扫描的手指边缘检测在输入的手掌区域中由上到下竖直检测手掌区域图的亮度值,当完整出现了变化模式时即检测到了正常的四指手指边缘,当没有检测到完整的变化模式,则本次竖直扫描失败,向右移动一个步长,继续进行下一次竖直扫描;
- [0029] 当扫描超过了预先设定的扫描区域时,则触发终止扫描条件,退出扫描检测。
- [0030] 所述对调整后的特征图进行关键点坐标的融合回归,具体包括:
- [0031] 通过多头注意力网络对关键点特征进行注意力感知及提取,再通过多层感知机对关键点的坐标进行全局回归,得到全局回归所预测的关键点坐标;
- [0032] 通过Vision Transformer网络对调整后的特征图进行注意力感知及特征提取,再通过CNN卷积网络对局部关键点进行检测,最后通过可微空间到数值变换将关键点响应映

射为局部回归的关键点坐标；

[0033] 通过由当前样本特征预测得到的融合权重对全局回归得到的关键点坐标和局部回归得到的关键点坐标进行融合,得到融合回归的第一关键点坐标；

[0034] 对指根边缘距离损失进行计算,所述指根边缘距离损失定义为:

$L_{edge} = 1/2 \sum_{i=1}^2 \min(d_i)$, d_i 是对应第*i*条指根边缘与第*i*个指根关键点的欧式距离序列,
 L_{edge} 为指根边缘距离损失。

[0035] 所述基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像,具体包括:

[0036] 根据预先设置的阈值对训练样本进行判断;

[0037] 当样本规模达到阈值时选择使用第二关键点坐标输出手掌感兴趣区域角点的坐标;

[0038] 当样本规模未达到阈值时选择使用第二关键点坐标输出手掌感兴趣区域中心点与指根关键点坐标来建立坐标系并定位正方形手掌感兴趣区域;

[0039] 对获取到的手掌感兴趣区域角点坐标执行射影变换提取手掌感兴趣区域图像。

[0040] 本申请实施例第二方面提供了一种自然场景下鲁棒的手掌感兴趣区域定位装置,所述自然场景下鲁棒的手掌感兴趣区域定位装置包括:

[0041] 手掌区域及轮廓提取模块,用于获取输入的手掌图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;

[0042] 手掌姿态调整模块,用于将所述手掌区域和轮廓区域特征图输入由弱监督训练策略训练的手掌感兴趣区域定位网络模型,通过由弱监督训练策略训练的手掌感兴趣区域定位网络模型的定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大操作,生成调整后的特征图;

[0043] 手掌关键点坐标融合回归模块,用于对调整后的特征图进行关键点坐标的融合回归,得到融合回归的第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标;

[0044] 手掌感兴趣区域图像提取模块,用于基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像。

[0045] 本申请实施例第三方面提供了一种特征提取识别器,其特征在于,所述特征提取器对如上任所述的自然场景下鲁棒的手掌感兴趣区域定位方法获取的手掌感兴趣区域图像进行手掌生物信息检测,基于检测到的手掌生物信息实现用户身份识别及核验。

[0046] 本申请实施例第四方面提供了一种终端设备,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的自然场景下鲁棒的手掌感兴趣区域定位程序,以实现如上任一所述的自然场景下鲁棒的手掌感兴趣区域定位方法中的步骤。

[0047] 有益效果:与现有技术相比,本申请提供了一种自然场景下鲁棒的手掌感兴趣区域定位方法、装置、特征提取识别器及设备,所述方法包括获取输入的手掌图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;将所述手掌区域和轮廓区域特征图输入训练的手掌感兴趣区域定位网络模型,通过训练的手掌感兴趣区域定位网络模型的定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采

样和指根放大操作,生成调整后的特征图;对调整后的特征图进行关键点坐标的融合回归,得到融合回归的第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标;基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像。这样本发明能够通过语义分割网络实现对手掌区域及手掌轮廓区域特征图的区分,再对得到的图像特征图进行姿态调整校正使得不同手掌姿态校正归一化,为后续网络模块提供高质量且姿态一致的输入样本,之后对得到的姿态进行校正过后的图像特征图进行基于全局信息和局部信息的关键点的坐标加权回归,使得全局和局部两种信息得以融合,实现坐标回归的鲁棒性,最后通过坐标执行射影变换得到手掌感兴趣区域图像,达到不需要红外相机的辅助的情况下对包含有复杂背景的手掌图像的处理;同时在本发明实现的过程中,网络在特征提取时,关注指根边缘区域,得出关键点回归的指根边缘距离损失,为后续坐标回归提供高质量的特征输入,提升回归精度及鲁棒性;此外本发明在训练手掌感兴趣区域定位网络模型的过程中,采用迭代手掌感兴趣区域定位算法,其中结合了异常检测功能模块,且使用基于距离极值点检测的手掌感兴趣区域定位方法以及基于线扫描的手掌感兴趣区域定位方法,使得其既能够剔除错误的手掌感兴趣区域定位,又可以迭代式搜索来提升整体的手掌感兴趣区域定位成功率,而在采用迭代手掌感兴趣区域定位算法得到标定信息后,本发明一种自然场景手掌数据集扩增方法,对原图像进行背景合成后,再对背景合成后得到的图像进行数据增广,从而实现为定位网络提供更加逼真和丰富的训练样本、进一步增强网络的泛化能力的效果;同时本发明对由自然场景下鲁棒的手掌感兴趣区域定位方法得到的手掌感兴趣区域图像进行手掌生物信息检测,可以实现高精度的用户身份识别及核验。

附图说明

[0048] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域普通技术人员而言,在不符创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0049] 图1为本发明提供的自然场景下鲁棒的手掌感兴趣区域定位方法的流程图;

[0050] 图2为本发明实施例提供的手掌感兴趣区域定位网络结构配置参考图;

[0051] 图3为本发明实施例提供的手掌关键点坐标回归网络逻辑框图;

[0052] 图4为本发明实施例提供的训练样本区域标签及关键点标注示意图;

[0053] 图5为本发明实施例提供的手掌关键点标记及手掌感兴趣区域定位示意图;

[0054] 图6为本发明实施例提供的边缘距离损失示意图;

[0055] 图7为本发明实施例提供的弱监督训练策略框图;

[0056] 图8为本发明实施例提供的迭代手掌感兴趣区域定位算法流程图;

[0057] 图9为本发明实施例提供的两种传统手掌感兴趣区域定位方法示意图;

[0058] 图10为本发明实施例提供的基于得到的融合回归的关键点坐标,提取得到手掌感兴趣区域图像效果图;

[0059] 图11为本发明实施例提供的迭代手掌感兴趣区域定位算法在不同数据集上自动生成的标注效果图;

- [0060] 图12为本发明实施例提供的数据增强的过程示意图；
- [0061] 图13为本发明实施例提供的空间变化及边缘感知效果图；
- [0062] 图14为本发明实施例提供的网络模型在自然手掌图像上的定位效果图；
- [0063] 图15为本发明实施例提供的自然场景下鲁棒的手掌感兴趣区域定位装置的原理框图。

具体实施方式

[0064] 本申请提供一种自然场景下鲁棒的手掌感兴趣区域定位方法装置、设备及介质，为使本申请的目的、技术方案及效果更加清楚、明确，以下参照附图并举实施例对本申请进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。

[0065] 本技术领域技术人员可以理解，除非特意声明，这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是，本申请的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和/或组件，但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。应该理解，当我们称元件被“连接”或“耦接”到另一元件时，它可以直接连接或耦接到其他元件，或者也可以存在中间元件。此外，这里使用的“连接”或“耦接”可以包括无线连接或无线耦接。这里使用的措辞“和/或”包括一个或多个相关联的列出项的全部或任一单元和全部组合。

[0066] 本技术领域技术人员可以理解，除非另外定义，这里使用的所有术语(包括技术术语和科学术语)，具有与本申请所属领域中的普通技术人员的一般理解相同的意义。还应该理解的是，诸如通用字典中定义的那些术语，应该被理解为具有与现有技术的上下文中的意义一致的意义，并且除非像这里一样被特定定义，否则不会用理想化或过于正式的含义来解释。

[0067] 另外，若本发明实施例中有涉及“第一”、“第二”等的描述，则该“第一”、“第二”等的描述仅用于描述目的，而不能理解为指示或暗示其相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。另外，各个实施例之间的技术方案可以相互结合，但是必须是以本领域普通技术人员能够实现为基础，当技术方案的结合出现相互矛盾或无法实现时应当认为这种技术方案的结合不存在，也不在本发明要求的保护范围之内。

[0068] 掌纹识别技术能够根据人体手掌表面的生物特征(如表皮纹线、浅层纹理及皮下静脉分布等)自动地对用户的身份进行鉴别，具有身份核验速度快、客观准确的优点。而目前，基于生物特征识别技术的电子哨兵系统拥有较好的发展机遇期。对于当前已有的系统而言，人脸识别存在隐私安全风险，指纹识别接触式的采集方式存在病菌传播的风险。与之相比，非接触式掌纹识别技术具有交互方式自然、隐私敏感性低、特征丰富、识别率高、不易引起病菌传播的优点，因而逐渐成为当前生物特征识别研究及产业化的热门方向。

[0069] 掌纹识别过程包括如下几个部分：手掌图像采集、掌纹感兴趣区域(ROI，在下文中所涉及的ROI其含义均为手掌感兴趣区域)定位、掌纹特征提取、特征匹配及识别。对于非接触式的图像采集而言，其背景较为复杂、手掌姿态变化较多。因此，手掌感兴趣区域的定位成为了识别系统的瓶颈所在。为了增加掌纹识别对环境的鲁棒性，当前的商用系统往往采

用红外摄像头结合红外光源进行成像,以达到去除背景的目的。在得到手掌红外图像之后,即可使用一系列的数字图像处理方法(如自适应二值分割、手掌边缘跟踪、关键点提取等)对手掌的ROI进行定位和提取,该类方案是需要特殊定制的硬件设备的。但是上述方法也有较多问题,首先硬件成本较高;其次,各厂商之间的系统相互不通用,限制了掌纹识别应用的普及和推广;再者,对于各类基于互联网的应用而言(如互联网金融、网上政务等),其登录页面的身份核验功能只能依赖于移动端的手机摄像头;故而,使用手机拍摄自然场景中的手掌图像并进行身份识别的需求正变得越来越迫切。然而,已有的ROI定位方法无法很好地满足当前的需求。目前现有技术的ROI定位方法有两类,第一类为传统方法,具体为基于单色背景手掌图像(如红外辅助成像、或人为布置单色背景)进行手掌区域分割,获得手掌轮廓之后,可以有以下不同的方法来定位ROI:

[0070] 方法一:基于线扫描的手掌感兴趣区域定位方法,检测手指边缘,通过检测指根边缘的切线来确定指根关键点的位置,进而建立手掌坐标系,定位ROI;而该方法对手掌姿态要求严格,适用场景受限,该类方法针对的是约束环境下的手掌ROI定位问题。线扫描位置方向的确定、掌根参照点的选择都是基于固定的手掌姿态和位置而进行的,在系统设计时就被固定下来,并作为一种规则来约束用户放置手掌的过程,因此,此类方法对手掌的旋转、平移、缩放、以及背景复杂度较为敏感,稍有偏差即无法处理。

[0071] 方法二:基于距离极值点的手掌感兴趣区域定位方法,首先确定手掌根部的参照点;求取此参照点与手掌边缘逆时针顺序各点的距离;通过距离变化的局部极大值和局部极小值来检测指尖与指根关键点;通过一定的后处理操作对结果进行去噪和细化,得到更为精准的关键点坐标;获得指根关键点之后,可以仿照方法一建立手掌局部坐标系、定位ROI;该方法效果不稳定,易受噪声影响,传统方法基于局部极值信息(如局部方向变化的拐点、局部距离变化的拐点)来检测关键点的位置,由于缺失全局信息,该类方法易受二值化分割噪声的干扰(如背景干扰物、衣袖、饰品、不均匀或者过暗的光照),因此,在真实的应用场景下此类方法的效果不鲁棒。

[0072] 方法三:与以上两种方法类似,也是基于指根关键点建立矩形ROI区域,区别是ROI区域与手掌轮廓相交,即搜索手掌区域内穿过指根关键点的最大的矩形ROI。该方法基于距离阈值对4个指根关键点(凹点)进行处理,提取食指-中指、无名指-小指对应的关键点,并且对左右手采用相同的算法进行处理。然而,该基于规则的方法所假定的手掌形状过于理想,实际处理中会出现有些用户的手掌大拇指指根关键点与其他四指的指根关键点过近的情况。此时,基于规则的方法就无法较好的区分指根关键点的顺序,有可能定位到错误的ROI。

[0073] 第二类为基于深度卷积网络的方法,该方法往往直接对ROI的关键点或者手掌的轮廓点进行检测和回归,然后基于关键点提取ROI图像。然而该类方法容易过拟合,手掌轮廓点,尤其是左右两侧及手掌掌根处的轮廓点,非常容易受到手指姿态、衣袖、饰物等的影响,因此,此类方法的定位精度和鲁棒性严重依赖于训练样本的容量,在训练数据量有限的情况下难以学到真正的模型参数,很容易产生过拟合,收敛到个例姿态或者干扰物上,在真实场景中表现欠佳,此外,当前的方案,往往需要五指分开,限制了掌纹识别的应用场景;同时该方法视野具有局限,鲁棒性和精度无法兼顾,基于全局视野的坐标回归,学习到的回归函数不具有空间不变性,在样本发生平移、旋转、缩放后往往无法正确地回归坐标,基于局

部响应的数值坐标映射网络模型具有空间不变性,但是缺乏全局视野,易受噪声干扰,定位精度易受影响;此外,该方法中人工标注数据工作量大,同时人工操作缺乏客观标准,在疲劳时亦容易引入随机误差。

[0074] 由上可见,现有技术的多种技术都要依赖红外成像去背景,且无法处理大范围内的手掌旋转、平移、缩放及手指闭合的问题,同时也无法满足自然场景下的非接触式手掌ROI定位需求。

[0075] 为了解决现有技术中的问题,本发明提供的一种自然场景下鲁棒的手掌感兴趣区域定位方法,可以由一种自然场景下鲁棒的手掌感兴趣区域定位装置来执行,所述装置可以由软件或硬件来实现,并可以应用于诸如平板电脑、计算机等之类安装有操作系统的智能终端设备上。在发明实施例中,获取输入的手掌图片后,通过深度网络进行手掌区域及手掌边缘区域的检测生成手掌区域及手掌轮廓区域分割特征图,在无需红外成像的辅助下实现手掌区域分割,之后针对得到的这些手掌相关区域特征图,利用定位网络检测手掌姿态,对手掌区域特征图进行标准化调整,以削弱旋转、平移的影响,同时对指间区及指根区进行放大,以保证后续坐标回归的精度,随后对调整后的特征图进行关键点坐标的融合回归,通过全局回归和局部回归得到的融合回归坐标兼具鲁棒性(空间不变性)和回归精度,最后基于得到的关键点坐标,提取得到手掌感兴趣区域图像;此外,对所述自然场景下鲁棒的手掌感兴趣区域定位方法进行训练,通过弱分类器来提取标准手掌图像的手掌区域、依此生成手掌边缘区域;然后,通过传统手掌感兴趣区域定位方法来检测指根及手掌感兴趣区域关键点,获得训练网络所需的标注信息;最后,结合图像合成技术生成具有复杂背景的仿真手掌图像,结合随机的空间几何变换及图像质量扰动生成丰富的训练样本,使得所述手掌感兴趣区域定位网络模型利用现有的公开掌纹数据库结合数据增强技术充分地进行弱监督网络训练,避免人工标注的繁重工作,增强网络的泛化能力。

[0076] 实例性方法

[0077] 本发明实施例提供一种自然场景下鲁棒的手掌感兴趣区域定位方法,所述自然场景下鲁棒的手掌感兴趣区域定位方法可以应用于智能终端设备。

[0078] 本发明所述方法具体实施时,可以基于一训练的手掌感兴趣区域定位网络模型实现,其中,如图2所示,所述手掌感兴趣区域定位网络模型包括主干网络、手掌区域及轮廓提取模块、手掌姿态调整模块、手掌关键点坐标融合回归模块、手掌感兴趣区域图像提取模块。具体的手掌感兴趣区域定位网络结构配置流程为获取由网络输入的手掌图片,通过主干网络(①)提取手掌图像的初级特征,对提取到的初级特征输入到手掌区域及轮廓区域提取模块,其中手掌区域及轮廓区域提取模块具体包括有语义分割模块与通道间softmax操作,由语义分割模块(②)进行手掌区域及轮廓区域提取,对提取到的特征图(BRE map)做基于像素的通道间的softmax操作(③)得到六通道特征图;将得到的特征图输入到手掌姿态调整模块中,所述手掌姿态调整模块包含有定位网络进行手掌姿态调整以及变形与下采样进行自适应缩放过程,通过定位网络(④)对图像姿态进行空间变换调整,并进行下采样和图像变形(指根区域放大),从而实现图像中手指姿态调整(⑤)得到手掌姿态调整后的特征图即BRE注意力特主图;将所述BRE注意力特征图输入到手掌关键点坐标融合回归模块,关键点坐标回归模块中进行基于局部信息的关键点坐标回归(⑥)和基于全局信息的关键点坐标回归(⑦),对得到的坐标进行融合回归(⑧)后通过坐标变换(⑨)得到手掌感兴趣区域

坐标;而最终通过手指感兴趣区域图像获取模块获取目标图像。此外在整个网络流程中对手掌区域分割后的特征图进行损失计算,对得到的坐标回归以及边缘最小距离的均方误差(MSE)损失等多种损失进行计算从而得到损失项,对其网络进行优化。

[0079] 如图1所示,本发明实施例提供一种自然场景下鲁棒的手掌感兴趣区域定位方法,包括如下步骤:

[0080] 步骤S10、获取输入的手掌图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;

[0081] 本发明实施例中,当该方法实施时,获取输入的图片后,通过一系列网络对输入图像中的手掌区域及手掌轮廓区域进行检测,从而得到手掌区域和轮廓区域特征图;

[0082] 进一步,所述获取输入图片,对所述输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图具体包括:

[0083] 通过具有多尺度金字塔结构的主干网络提取图片中的初级特征;

[0084] 基于所获取的初级特征,采用语义分割网络对初级特征进行处理生成包含有背景区域图、手掌区域图、手掌轮廓区域图三个通道的区域特征图,且所述三通道输出通过焦点损失对其语义分割效果进行监督;

[0085] 生成的三通道区域特征图经过预处理后生成含有六个通道的特征图。

[0086] 具体的,获取到输入的图片后,使用具有多尺度金字塔结构的各类常用主干网络模型,如HRFormer等,来提取初级特征,从而实现尽可能的保留手掌轮廓细节信息、提升关键点坐标的精度效果,并为后续的语义分割和关键点坐标回归做准备,其中通过主干网络模型提取得到的初级特征共包含有224个通道;此外,为了提升网络速度时,也可以选用轻量主干网络模型,如ResNet等来提取初级特征。

[0087] 主干模型获得初级特征之后,采用语义分割网络对初级特征进一步处理,将得到的具有224个通道的初级特征通过语义分割输出区域特征图。此处,区域特征图包含3个通道,分别对应于背景区域图、手掌区域图、手掌轮廓区域图。

[0088] 生成的3通道区域特征图经过预处理后成为6个通道的特征图,具体的预处理方法为:对原始的输入3通道区域特征图做基于像素的通道间的softmax操作,生成新的3通道特征图,从而通过该操作对区域类别进行拉伸(类似于二值化的拉伸),明确当前位置是属于背景、还是手掌内部、还是手掌轮廓,之后将原始的3通道区域特征图与新生成的3通道区域特征图串接起来,构成预处理后的区域特征图,共6个通道。

[0089] 在整个自然场景下鲁棒的手掌感兴趣区域定位方法运行过程,使用分类损失对区域特征图中的每个像素位置进行监督。一个像素位置可对应3个类别:背景像素、手掌区域像素、手掌边缘像素;类别标签分别为0,1,2;特征图的三个通道输出分别代表了当前像素位置隶属于三个类别的概率,像素的三通道输出通过焦点损失(焦点损失英文释义为focal loss,下述出现的focal loss均表示为焦点损失),其主要关注训练样本中困难样本的误差,其对像素的语义分割效果进行监督,该分割损失记为 L_{seg} 。分割损失由区域特征图中每个像素的分类损失构成;对于每个像素,其分类损失为:

$$L_{pix}(t) = -\frac{1}{c_{WH}} \sum_{c=1}^C \sum_{i=1}^{WH} y(c, i) (1 - p_t(c, i))^y \log(p_t(c, i)),$$

其中, $L_{pix}(t)$ 为对应的区域特征图分类损失,WH分别为手掌区域特征图的宽和高;C为总的区域特征图通道个

数,此处 $C=3$;我们定义了三个区域类别:背景区域(记为类别0)、手掌区域(记为类别1)、手掌边缘区域(记为类别2),分别对应特征图的3个特征通道; c 为特征图通道的索引; i 为特征图中像素的索引; γ 为focal loss中定义的超参数,其值取2.0; y 代表特征通道 c 中像素 i 的真实类标,以onehot向量的形式进行表示; t 代表区域特征图; $p_t(c, i)$ 代表特征图通道 c 中像素 i 的类别预测概率值,其具体计算方法为:
$$p_t(c, i) = \frac{e^{t(c, i)}}{\sum_{c=1}^C e^{t(c, i)}}$$
此变换为特征通道间的softmax运算,即, $p_t(c, i)$ 代表了区域特征图通道 c 中像素 i 属于类别 c 的概率, $c \in \{0, 1, 2\}$;分割损失 L_{seg} 由区域特征图的像素分类损失的加权平均构成,针对不同的网络结构有不同的具体表现形式,当只使用一组手掌区域特征图时(特征图记为BRE),则 $L_{seg} = L_{pix}(BRE)$,当有多组手掌区域特征图时,如:当首先预测得到手掌区域粗分割特征图后(记为S),然后对S进一步精细化处理后得到新的手掌区域特征图(记为BRE),如果将两个区域特征图都纳入loss进行监督,则 $L_{seg} = \frac{1}{2} \sum_t L_{pix}(t)$,其中 $t \in \{S, BRE\}$ 。

[0090] 步骤S20、将所述手掌区域和轮廓区域特征图输入训练的手掌感兴趣区域定位网络模型,通过由弱监督训练策略训练的手掌感兴趣区域定位网络模型的定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大操作,生成调整后的特征图;

[0091] 本发明实施例中,通过定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,之后基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大操作,生成调整后的特征图。

[0092] 具体的,在得到手掌区域和轮廓区域特征图之后我们可以对指根和ROI关键点的坐标进行回归。由于非接触式掌纹采集的过程中手掌的姿态较为灵活,因此,手掌倾斜、旋转、平移、过近、过远的情况时有发生。为了保证输出精度,在坐标回归之前我们先对手掌的姿态进行调整。因此采用定位网络来回归手掌的姿态校正参数 θ ,该定位网络结构为3层卷积模块,后接2层全连接层,输出为 θ ;定位网络同时感知手掌的几何姿态及指根区域的几何尺度进行回归得到手掌姿态校正参数 θ ,在得到手掌姿态校正参数 θ 之后,空间变换网络模块对输入的6通道区域特征图基于 θ 执行特征图的空间变换(如:射影变换、仿射变换),即可达到上述手掌姿态校正的目的,生成调整后的指根区特征图,而其中特征图的空间变换通过PyTorch网络的基础操作进行实现。而在对特征图进行调整之后,对特征图进行降采样,使得姿态变换后特征图的尺寸降低为输入特征图的1/4(即横向纵向各降低一半),从而实现降低后端基于全局信息的坐标回归过程中的计算量,而在调整手掌姿态的同时还需要将指根区域进行放大以保证最终关键点坐标回归的精度和输入到后端的指根区域分辨率。

[0093] 不同手掌姿态校正归一化,为后续网络模块提供高质量且姿态一致的输入样本,减小后续关键点坐标回归网络模块的学习难度、降低坐标回归模块的学习空间,降低其过拟合风险,保证模型整体泛化能力。

[0094] 步骤S30、对调整后的特征图进行关键点坐标的融合回归,得到融合回归的第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到

对应于输入手掌图像的手掌感兴趣区域第二关键点坐标；在本发明实施例中，对经过步骤S20调整得到的特征图进行关键点坐标的融合回归，从而得到融合回归后的第一关键点坐标，并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中，得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标；

[0095] 进一步的，所述自然场景下鲁棒的手掌感兴趣区域定位方法，其特征在于，所述对调整后的特征图进行关键点坐标的融合回归，具体包括：

[0096] 通过多头注意力网络对关键点特征进行注意力感知及提取，再通过多层感知机对关键点的坐标进行全局回归，得到全局回归所预测的关键点坐标；

[0097] 通过VisionTransformer网络对调整后的特征图进行注意力感知及特征提取，再通过CNN卷积网络对局部关键点进行检测，最后通过可微空间到数值变换将关键点响应映射为局部回归的关键点坐标；

[0098] 通过由当前样本特征预测得到的融合权重对全局回归得到的关键点坐标和局部回归得到的关键点坐标进行融合，得到融合回归的第一关键点坐标；

[0099] 对指根边缘距离损失进行计算，所述指根边缘距离损失定义为：

$L_{edge} = 1/2 \sum_{i=1}^2 \min(d_i)$, d_i 是对应第 i 条指根边缘与第 i 个指根关键点的欧式距离序列, L_{edge} 为指根边缘距离损失。

[0100] 具体的，可微空间到数值变换即Differentiable spatial to numerical transform, 其缩写为DSNT, DSNT网络能够将局部关键点响应转换为其对应的数值坐标。由于卷积响应所编码的是局部像素区域的分布特征，并且整幅图像中卷积权值共享，这使得DSNT回归器具有空间不变性。即，当手掌发生偏移和旋转时，训练好的网络模型依然能够正确的回归关键点的坐标。但是，局部视野无法感知全局信息，容易受到区域分割噪声的干扰。当网络前端预测的手掌区域不理想时，坐标容易回归到噪声点上。此外，根据指根关键点确定ROI位置时，需要首先明确左右手标记，进而能够确定指根关键点的顺序(大拇指是哪一个)，从而选取无名指-小指及食指-中指两个指根关键点进行手掌局部坐标系的建立。因此，手掌分布的全局信息有助于提升关键点定位的正确性。基于此，本方案的在进行坐标回归时包括局部回归和全局回归两部分坐标预测，并对其进行融合得到关键点坐标，全局与局部两种信息的融合使得坐标回归网络模块能够兼具鲁棒性(空间不变性)和回归精度。

[0101] 所述全局回归具体包括：首先通过多头注意力网络模块对关键点特征进行提取，然后通过一个全连接网络对关键点的坐标进行回归，具体的，全局回归模块为VisionTransformer结构，其中Vision Transformer缩写ViT, 是一种常用的网络模型，共包含一个多头注意力模块(Multi-head self-attention, MHSA) 和一个多层感知机模块(Multilayer perceptron, MLP)。其中，MHSA多头注意力模块的输入特征图通道为6, 区域划分window size为7, 输出通道为6, head个数为3; 经过MHSA处理之后，其生成的二维特征图被拉伸成一维向量，输入后续的MLP模块。MLP模块包含输入层、激活层、dropout层、以及输出层，输出层的输出为所预测的关键点坐标。全局回归网络模块所得坐标标记为 X_g 。 $X_g \in \mathbb{R}^{14}$ 包含了关键点 t_1 、 t_2 及ROI中心点和角点的 x 、 y 坐标，即

[0102] $X_g = [x_{t1}, y_{t1}, x_{t2}, y_{t2}, x_c, y_c, x_{r1}, y_{r1}, x_{r2}, y_{r2}, x_{r3}, y_{r3}, x_{r4}, y_{r4}]$ 。

[0103] 所述局部回归具体包括：首先通过Vision Transformer网络模块对步骤S20生成的姿态调整特征图进行注意力感知及特征提取，然后通过一个CNN卷积模块对局部关键点

进行检测,最后通过DSNT将关键点响应映射为关键点坐标。具体的,Vision Transformer模块输入通道为6,输出通道为6,head个数为3>window size为7;CNN卷积模块结构为:BN+Conv2D+ReLU+Conv2D。所得坐标记为 X_1 。 $X_1 \in \mathbb{R}^{14}$ 包含了关键点t1、t2及ROI中心点和角点的xy坐标,即

$$[0104] \quad X_1 = [x_{t1}, y_{t1}, x_{t2}, y_{t2}, x_c, y_c, x_{r1}, y_{r1}, x_{r2}, y_{r2}, x_{r3}, y_{r3}, x_{r4}, y_{r4}]。$$

[0105] 进一步的,通过由当前样本特征预测得到的融合权重对全局回归得到的关键点坐标和局部回归得到的关键点坐标进行融合,得到融合回归的关键点坐标具体包括:

[0106] 通过权重 w 对两类坐标进行融合。最终的坐标为: $X = w \cdot X_g + (1-w) \cdot X_1$,且 $w \in [0, 1]$ 。此处的融合权重 w 可以通过网络根据当前样本的特征进行预测,也可以固定设置为一个值,如0.5;融合权重 w 的预测方法为在基于全局信息进行关键点坐标回归的同时,也基于全局信息对融合权重 w 进行预测,通过修改全局回归的输出为 $X_g \in \mathbb{R}^{15}$,其包含了关键点t1、t2及ROI中心点和角点的x、y坐标及融合权重预测初值 w_- ,即

$$[0107] \quad X_g = [x_{t1}, y_{t1}, x_{t2}, y_{t2}, x_c, y_c, x_{r1}, y_{r1}, x_{r2}, y_{r2}, x_{r3}, y_{r3}, x_{r4}, y_{r4}, w_-]。$$

[0108] 为了保证 w 位于0到1之间,我们对 w_- 做如下处理:

$$[0109] \quad w = \frac{1}{1 + e^{-s \cdot w_-}}$$

[0110] 其中 s 为函数曲线形状调整参数,通过网络进行学习,初值为0。

[0111] 更进一步的,本发明对训练过程中由于区域分割、关键点定位、指根边缘距离区域感知以及DSNT网络计算造成的损失进行计算,具体包括:

[0112] 整个网络最终预测的关键点包括,指根关键点t1、t2,ROI中心点c及四个角点r1~r4,对于输出的关键点坐标值,我们使用MSE均方损失函数进行监督。对于指根关键点t1、t2的定位损失记为 L_t ,对于ROI关键点的定位损失记为 L_{roi} ,其定义分别为:

$$[0113] \quad L_t = \frac{1}{4} \sum_{i=1}^4 \|k_t(i) - y_t(i)\|_2^2$$

[0114] 其中, $k_t = [x_{t1}, y_{t1}, x_{t2}, y_{t2}]$ 为预测的指根关键点坐标, y_t 为对应的坐标真值, i 为坐标索引;

$$[0115] \quad L_{roi} = \frac{1}{10} \sum_{i=1}^{10} \|k_{roi}(i) - y_{roi}(i)\|_2^2$$

[0116] 其中, $k_{roi} = [x_c, y_c, x_{r1}, y_{r1}, x_{r2}, y_{r2}, x_{r3}, y_{r3}, x_{r4}, y_{r4}]$,其为预测的ROI关键点坐标(ROI中心点c及四个角点r1~r4), y_{roi} 为对应的坐标真值。

[0117] 本发明还提出指根边缘距离损失来促使网络在训练的过程中对指根边缘区域进行感知。指根边缘距离损失定义为: $L = 1/2 \sum_{i=1}^2 \min(d_i)$ 。其中, d_i 是对应第 i 条指根边缘与第 i 个指根关键点的欧式距离序列。具体如图6所示,序列中的每个距离由所回归的指根关键点(记为 t_i)到对应的指根边缘点(记为 e_j)构成,即: $d_i = \{\text{distance}(t_i, e_j)\}$, $h = \{1, 2\}$, $j = 1, \dots, n$,其中 n 为指根边缘点序列中所包含的点的总个数,其中所涉及到的 i 均代指根边缘,且其取值一致,指根边缘点序列中所包含的点的总个数能够覆盖指根边缘即可,图中 i 也为对应的边缘,边缘距离损失记为 L_{edge} 。关键点回归的指根边缘距离损失促使网络模

型对指根边缘进行感知,促使网络在特征提取时,关注指根边缘区域,为后续坐标回归提供高质量的特征输入,提升回归精度及鲁棒性。通过指根边缘感知,使得网络在手指闭合时也能定位到关键点。较佳的,如图13所示,图中上下两列分别为两个输入图像处理按例,图13中(a)为输入图,(b)中间特征图,(c)为检测到的手掌区域图,(d)为姿态调整后的区域特征图,(e)为指根边缘感知特征图,(f)为预测到的关键点;通过边缘距离损失项能够促使网络对指根边缘进行感知,姿态调整时根据需要能够对指根区进行放大、旋转及平移,保护和强化指根边缘特征,使得最终输出的ROI关键点图像更加清晰准确。

[0118] 如前所属,对于手掌区域图及手掌轮廓区域图,我们针对每个像素,执行focal loss从而对分割模块的参数进行监督和训练,该分割损失为 L_{seg} 。

[0119] 而DSNT中响应热图的损失函数沿用官方函数不做展开,记为 L_{hm} 。

[0120] 综上,整体的损失函数为:

$$[0121] \quad L = \alpha L_{seg} + \beta L_t + \gamma L_{roi} + \delta L_{edge} + \mu L_{hm}$$

[0122] 其中, $\alpha, \beta, \gamma, \delta, \mu$ 为非负权重,用来平衡每种损失的重要性。按照不同模块重要性的优先级我们设置如下: $\alpha=1.0, \beta=0.5, \gamma=0.1, \delta=0.1, \mu=0.1$,即首先确保网络生成理想的手掌区域特征图,然后确保指根关键点 t_1, t_2 的坐标回归精度,最后,ROI关键点定位损失 L_{roi} 、指根边缘最小距离损失 L_{edge} 及DSNT中的关键点热图检测损失 L_{hm} 处于最低优先级来对模型的学习进行约束。实践中,每个训练批次在损失反向传播之前,我们需要对损失函数

中的每个损失项进行归一化,即: $L_x = \frac{L_x}{|L_{x1}|}$,其中, $x \in \{seg, t, roi, edge, hm\}$ 代表各个损失项, $|L_{x1}|$ 代表损失项在当前训练批次所获得的具体损失值。此处, L_x 是变量, $|L_{x1}|$ 是常量。这样使得本发明能够归一化不同类型损失的尺度,均衡不同损失项对网络的作用程度,并根据网络损失对网络进行优化。

[0123] 步骤S40、基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像。

[0124] 本发明实施例中,基于步骤S30得到的第二关键点坐标,提取得到手掌感兴趣区域图像。

[0125] 进一步的,所述基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像,具体包括:

[0126] 根据预先设置的阈值对训练样本进行判断;

[0127] 当样本规模达到阈值时选择使用第二关键点坐标输出手掌感兴趣区域角点的坐标;

[0128] 当样本规模未达到阈值时选择使用第二关键点坐标输出手掌感兴趣区域中心点与指根关键点坐标来建立坐标系并定位正方形区域;

[0129] 对获取到的手掌感兴趣区域角点坐标执行射影变换提取手掌感兴趣区域图像。

[0130] 具体的,对于ROI图像的提取,可以直接使用步骤S30得到的融合回归关键点坐标的4个ROI角点执行射影变换提取ROI图像;也可以先基于指根关键点 t_1, t_2 及ROI中心点 c 对矩形ROI的四个角点坐标进行求解,再执行射影变换提取ROI图像;第一种方法即直接使用步骤S30得到的第二关键点坐标的4个ROI角点执行射影变换提取ROI图像,所预测的四个角点可以根据手掌的轮廓进行适应,不是严格的矩形;因此能够适应一定程度上的手掌变形,但是在训练数据量不充分的情况下存在过拟合的风险,在训练集较小的情况下,预测的关

键点很可能受到手指姿态、衣袖、饰品等因素的干扰,预测结果是错误的;第二种方法即基于第二关键点中的指根关键点 t_1 、 t_2 及ROI中心点 c 对正方形ROI的四个角点坐标进行求解进一步的,连接指根关键点 t_1 、 t_2 ,过连线 l 的中点做垂线向掌心延伸建立局部坐标系,根据预定义的ROI偏移量 d 和ROI边长 s 在图像坐标系中计算正方形ROI的四个角点 $r_1 \sim r_4$ 的坐标,再执行射影变换提取ROI图像,该方法生成的ROI是严格的正方形,重点检测指根关键点这一具有明显视觉显著性的区域,泛化能力强,但是对异常的手掌姿态没有适应性,在手掌倾斜时,预测的ROI还是正方形的,此时,其内部的掌纹是有所变形的。

[0131] 实践中,根据训练样本的容量进行选择,并设置相应的样本阈值。此处训练样本指虚拟背景合成及数据增强之前的手掌图像样本,即最开始收集到的手掌图像,用来训练手掌感兴趣区域定位网络模型。较佳的,当原始训练样本规模够大时,达到百万级时,选择使用融合回归关键点坐标输出ROI角点的坐标;当原始训练样本规模未达到百万级时,选择使用融合回归关键点坐标输出ROI中心点与指根关键点坐标来建立坐标系并定位正方形ROI区域。具体的,如图10所示,训练样本的容量达到阈值后,如图10-a采用上述第一种方法,角点能够根随手掌边缘做一定的适应调整,而图10-b采用上述第二种方法,角点构建的仍然为一个正方形,不能够根随手掌边缘做一定的适应调整,而在图10-c中训练样本的容量未达到阈值,采用第一种方法得到的ROI坐标构建出的ROI是错误的。

[0132] 所述弱监督训练策略训练手掌感兴趣区域定位网络模型的训练步骤包括:

[0133] 通过手掌区域弱分割器,采用由单色背景手掌图像集训练得到的二分类器对输入的第一手掌图像中的像素进行分类,提取手掌区域,并对手掌区域执行边缘检测得到手掌轮廓;

[0134] 通过迭代手掌感兴趣区域定位算法处理手掌轮廓,得到输入的第一手掌图像的第一标注信息;

[0135] 基于第一手掌图像的手掌区域提取手掌像素,结合手掌像素和自然图库中的自然背景图像进行背景合成,生成自然环境下的第二手掌图像;

[0136] 对上述得到的第二手掌图像和第一标注信息通过同时执行相同的随机空间变化操作,得到第三手掌图像和第二标注信息,再对第三手掌图像执行图像质量随机扰动得到第四手掌图像,得到数据增广后包含有第四手掌图像和第二标注信息的训练样本;

[0137] 通过上述训练样本对所述手掌感兴趣区域定位网络模型进行训练。

[0138] 具体的,弱监督训练策略首先采用简单的定位算法自动的在受约束的理想手掌图像上进行手掌区域及关键点标注,然后通过背景合成和数据增强技术生成丰富的具有自然背景的虚拟手掌图像样本对上述自然场景下鲁棒的手掌感兴趣区域定位方法进行训练。当前的公开掌纹数据库中的图像都是在严格的限制条件下所采集的,往往采用单色背景,手掌朝向固定,五指分开并且光照适宜。因此,使用传统的ROI定位算法在这些受限数据库上能够较好地ROI进行定位,生成所需标注信息。所述弱监督训练策略主要包括自动化标注、数据增强两个功能模块,其中自动化标注模块主要包括如下三个部分:手掌区域弱分割器、迭代手掌感兴趣区域定位算法、标签生成方法。

[0139] 所述手掌区域弱分割器的实现方式具体为:给定单色背景手掌图像集,通过截图的方式分别随机截取背景区域图像块和手掌区域肤色图像块;将所有背景像素值作为一类,而所有肤色像素值作为另一类,训练二分类器(如支持向量机SVM、逻辑回归LR二分类器

等) ;使用二分类器对新输入的手掌图像中的像素进行分类,提取手掌区域。

[0140] 所述迭代手掌感兴趣区域定位算法,具体包括:

[0141] 对提取到的手掌区域进行关键点的检测,并校验关键点的个数;

[0142] 基于校验后的关键点个数,当检测到完整的5个指尖点和4个指谷点则执行基于距离极值点的手掌感兴趣区域定位方法;

[0143] 当校验未得到完整的5个指尖点和4个指谷点时则进行基于线扫描的手指边缘检测,当检测到4个指尖点及3个指谷点时则执行基于线扫描的手掌感兴趣区域定位方法;

[0144] 当未检测到4个指尖点及3个指谷点时则移动到下一个扫描点执行基于线扫描的手指边缘检测,通过检测再次判断是否扫描到4个指尖点及3个指谷点,当检测到4个指尖点及3个指谷点时则执行基于线扫描的手掌感兴趣区域定位方法,当未检测到4个指尖点及3个指谷点时移动到下一个扫描位置再次执行基于线扫描的手指边缘检测;

[0145] 当扫描触发停止条件时退出检测。

[0146] 所述迭代手掌感兴趣区域定位算法和所述基于线扫描的手指边缘检测,具体还包括:

[0147] 所述迭代手掌感兴趣区域定位算法通过定位方法得到手掌感兴趣区域后对所述ROI进行异常检测;

[0148] 所述基于线扫描的手指边缘检测在输入的手掌区域中由上到下竖直检测手掌区域图的亮度值,当完整出现了变化模式时即检测到了正常的四指手指边缘,当没有检测到完整的变化模式,则本次竖直扫描失败,向右移动一个步长,继续进行下一次竖直扫描;

[0149] 当扫描超过了预先设定的扫描区域时,则触发终止扫描条件,退出扫描检测。

[0150] 更具体的,本方案采用程序自动标注的方法来达到弱监督训练的目的。在迭代手掌ROI定位算法中,实现对传统ROI定位算法进行改进,引入异常检测模块,通过位置、尺寸、角度等几何信息及图像有效面积、清晰度等质量信息来判断当前检测到的ROI是否为正常的ROI区域。启动ROI定位,得到当前ROI后通过异常检测模块进行判断,如果判定为正常ROI则成功返回退出,否则再迭代地进行下一次ROI检测,直至触发终止条件,失败退出。此处,终止条件为预先设置的用来控制线扫描的区域范围,如预先设置为图像的3/4区域,即当扫描位置超过图像宽度的3/4时即触发终止条件,直接退出。迭代手掌ROI定位算法将两种ROI定位方法进行结合,基于距离极值点的方法能够获得五指指尖及指根关键点的信息,对手掌旋转(正负30度以内)具有更好的鲁棒性,但是该方法必须检测到完整的五指且识别到大拇指才能正常地建立手掌坐标系定位ROI,对于非接触式手掌采集,该前提条件有时候无法满足,因此,此方法鲁棒性不足;而基于线扫描的ROI定位方法只需要检测到四指即可,因此可以作为距离极值法的有效补充;

[0151] 如图9所示,本发明对手掌采用基于距离极值点的手掌感兴趣区域定位方法得到图9-a中的图像,具体为,首先对手掌区域进行分割,进一步获得手掌边缘点序列以及手掌最大内接圆(半径为 r 、圆心为 c);再设置一个起始参考点记为 sp (使用图像最右边界上距离手掌内接圆圆心 c 最近的那个点,图a中的start point),求取参考点 sp 与手掌轮廓上每个点之间的距离(由手掌轮廓上距离 sp 最近的点开始,沿着手掌轮廓逆时针顺序求取),构成距离变化曲线(如图b所示);曲线中局部极大值点(参考图b中的peaks)对应的是指尖点,局部极小值点(参考图b中valley points)对应的是指根关键点(此时,由于可能存在手掌旋

转,因此按照距离极值所确定的指尖和指根点与其真实位置相比会有一定的微小误差);为了进一步减小误差,连接中指指尖点与内接圆圆心 c ,并继续向前延伸 $1.2*r$ 的距离,得到新的参照点 rp (图a中的reference point)。连线 $rp \rightarrow c$ 的方向作为当前的手掌方向;以 rp 为新的参考点,逆时针求取 rp 与手掌轮廓上点的距离序列,得到新的距离变化曲线(如图b所示)。曲线局部极大值点为指尖点,局部极小值点为指根点。由于 rp 是基于手掌方向来设置的,因此其对应的距离曲线的极值点更加接近真实的指尖和指根点;记食指-中指对应的指根点为 $vp1$,无名指-中指对应的指根点为 $vp2$ 。由 $vp1$ 及 $vp2$ 的连线以及其连线的垂线可以建立手掌直角坐标系,进而可以根据预先设定的ROI偏移量和ROI边长求取方形ROI的四个角点的坐标。对手掌图像采用基于线扫描的手掌感兴趣区域定位方法具体为图9-c、9-d所示,首先对手掌区域进行分割,得到二值化的手掌区域图(如图c所示),不妨设图中手掌区域像素灰度值为255,背景区域像素灰度值为0;再竖直扫描线(记为 $1s$)由上向下进行扫描,检测当前位置的像素值;由上至下扫描过程中,当首次发现像素值由255跳变为0时,当前像素点记为 $p1$;继续向下检测,当检测到像素值由0跳变为255时,当前像素点记为 $p2$;继续向下检测,当检测到像素值由255跳变为0时当前像素点记为 $p3$;继续向下检测,当检测到像素值由0跳变为255时,当前像素点记为 $p4$;继续向下检测,当检测到像素值由255跳变为0时,当前像素点记为 $p5$;继续向下检测,当检测到像素值由0跳变为255时,当前像素点记为 $p6$;如果单次竖直扫描可以成功检测到点 $p1$ 至 $p6$,则:通过边缘跟踪算法获取指谷边缘 $p1-p2$ 及 $p5-p6$;检测两段指谷边缘曲线 $p1-p2$ 及 $p5-p6$ 的切线,切线对应的切点分别记为 $vp1$ 和 $vp2$,由 $vp1$ 及 $vp2$ 的连线以及其连线的垂线可以建立手掌直角坐标系,进而可以根据预先设定的ROI偏移量 d 和ROI边长 s 求取方形ROI的四个角点的坐标,之后程序返回成功退出;如果单次竖直扫描无法成功检测到完整的点 $p1$ 至 $p6$,则:移动固定步长至下一个横坐标位置,如果新的位置超出了预设范围(如图像宽度的 $1/2$ 等预先设置的阈值),则返回失败退出,如果新的位置未超出预设范围,则返回上述步骤再次执行竖直扫描,尝试定位;其中对图9-d中的图像进行提取得到图9-e中的ROI图像。

[0152] 进一步的,如图8所示,为本发明实施例提供的迭代手掌感兴趣区域定位算法流程图,该图实现上述迭代手掌ROI定位算法过程,其执行如下内容:

[0153] 步骤S81、具体包括:依次进行操作,开始,输入手掌图像,进行手掌区域分割,进行最大连通域检测,手掌中心点检测,边缘距离队列生成,手掌轮廓起始点检测,边缘距离队列重排列,指尖关键点检测,手掌方向检测,参考点检测,再次进行边缘距离队列重排序,对指尖、指谷关键点检测,校验关键点个数;

[0154] 步骤S82、判断是否检测到5个指尖点及4个指谷点,当是进入步骤S83,当否进入步骤S84;

[0155] 步骤S83、具体包括:执行基于距离极值点的ROI定位方法,对指谷边缘切线检测,进行ROI定位及图像提取,对ROI几何及图像质量异常检测,判断ROI是否异常,当是进入步骤S84,当否进入步骤S85;

[0156] 步骤S84、具体包括:对图像进行基于线扫描的手指边缘检测,检测白-黑-白-黑-白-黑-白变化序列,判断是否检测成功,若是则进入步骤S86,若否进入步骤S87;

[0157] 步骤S85、输出ROI图像及坐标,成功返回;

[0158] 步骤S86、具体包括:指谷边缘切线检测。ROI定位及图像提取,ROI几何及质量异常

检测,判断ROI是否异常,若是则进入步骤S87,若否则进入步骤S85;

[0159] 步骤S87、线扫描移动到下一个扫描位置,判断是否触发扫描停止条件,若是则进入步骤S88,若否则进入步骤S84;

[0160] 步骤S88、失败退出。

[0161] 其中,最大连通域检测:仅保留手掌区域,排除背景噪声干扰;

[0162] 手掌中心点检测:手掌最大内接圆的圆心即为手掌中心点;

[0163] 边缘距离队列:即参考点与逆时针顺序的手掌边缘点之间构成的距离队列(最开始的参考点,可以选则掌根侧图像边框的中点);

[0164] 边缘距离队列重排列:手掌轮廓检测所返回的轮廓点序列往往从图中最左上角的手掌轮廓点开始,依此生成的距离队列的极值点之间的顺序受手掌旋转的影响无法唯一确定。因此,为了对首个轮廓点的位置进行约束,我们约定:手掌轮廓上距离参照点最近的那个点为首个轮廓点,然后按照逆时针顺序依次排列手掌轮廓点再求取对应的距离队列。此过程视为边缘距离队列重排列;

[0165] 指尖关键点检测:距离队列的局部极大值所对应的手掌边缘上的点为指尖点;

[0166] 手掌方向检测:由中指指尖与手掌中心点构成的直线的方向视为手掌方向;

[0167] 参考点检测:由上述中指指尖与手掌中心点构成的直线向掌根方向延长与轮廓相交的点即可作为新的参考点(视为一种对起始参考点进行的优化操作,使得参考点距离四指指尖较远);

[0168] 指谷点检测:距离队列的局部极小值所对应的手掌边缘上的点为指谷点

[0169] 校验关键点个数:正常的完整的手掌图像应该能够检测到5个指尖点、4个指谷点;当大拇指没有拍到时,应该能够检测到4个指尖点、3个指谷点;因此,如果能够检测到完整的5个指尖点和4个指谷点,则执行基于距离极值点的ROI定位方法。如果仅仅检测到4个指尖点、3个指谷点则执行基于线扫描的ROI定位方法;

[0170] 如图11所示,为本发明实施例提供的迭代手掌感兴趣区域定位算法在不同数据集上自动生成的标注效果图,由本发明迭代手掌ROI定位算法对受约束的理想手掌进行自动标注后,均能够得到包含清晰、完整的手掌区域、指根关键点及ROI的训练用标记图像。

[0171] 基于线扫描的ROI定位方法具体如下:

[0172] 假设手掌横放,指尖向左;在起始位置(如图像原点)竖直地由上到下竖直地检测手掌区域图的亮度值,当完整出现了(白-黑-白-黑-白-黑-白)的变化模式时,即检测到了正常的四指手指边缘。通过检测指谷边缘切线即可定位切点 t_1 、 t_2 ;如果没有检测到完整的跃变模式,则本次竖直扫描失败,向右移动一个步长 $step$,继续进行下一次竖直扫描,探测(白-黑-白-黑-白-黑-白)的变化模式。如果一直失败,则一直向右移动 $step$ 步长,以至超过了预先设定的扫描区域(如图像宽度的 $3/4$),则触发终止扫描条件,失败退出。

[0173] 标签生成方法结合上述迭代手掌ROI定位算法具体为:在向手掌区域弱分割器输入一幅待标注的第一手掌图像 I ,通过上述区域弱分割器提取手掌区域 Seg_palm ,对手掌区域 Seg_palm 执行边缘检测得到手掌轮廓 $Boundary$;使用迭代手掌感兴趣区域定位算法处理手掌轮廓 $Boundary$ 得到指根关键点 t_1 和 t_2 的坐标(其中 t_1 、 t_2 对应于两指根边缘最长切线的两个切点),再对 $Boundary$ 进行加粗5个像素得到边缘区域 Seg_edge ,对食指-中指及无名指-小指的指根关键点 t_1 、 t_2 执行基于迭代手掌感兴趣区域定位算法,其效果如图5中所示

得到清晰完整的ROI区域,能够确定ROI的四个角点坐标 $r_1 \sim r_4$,及ROI中心点坐标 c ;基于指根关键点 t_1 和 t_2 生成指根边缘点序列 $edge_1$ 和 $edge_2$,其效果如图4中指谷处出现的加粗轮廓;至此,获得对应于输入的第一手掌图像I的标注信息,具体为:手掌区域 Seg_palm ,手掌轮廓区域 Seg_edge ,指根关键点坐标 t_1, t_2 ,ROI关键点坐标角点 $r_1 \sim r_4$ 及中心点 c ,指根边缘点序列 $edge_1$ 和 $edge_2$;由以上,得到的标注信息记为第一标注信息L,且 $L = [Seg_palm, Seg_edge, t_1, t_2, c, r_1 \sim r_4, edge_1, edge_2]$ 。

[0174] 所述数据增强模块,包括背景合成和数据增广。经过标签生成方法迭代ROI定位算法能够自动的生成受约束手掌图像的第一标注信息L,但是我们的目标是自然场景下的手掌ROI定位,为了充分训练本方案所提出的定位网络模型,通过数据增强的方法利用姿态受限的手掌图像和对应的标注数据来生成更加复杂的自然状态下的训练样本。其中背景合成具体包括:输入姿态及背景受限第一手掌图像I和对应的第一标注信息L,从自然图像库中随机提取一张自然背景图像B,根据手掌区域 Seg_palm 从第一手掌图像I中提取手掌像素 $PalmPixel$,将手掌像素 $PalmPixel$ 粘贴到背景图像B中生成自然环境下的第二手掌图像I2;数据增广具体包括:输入第二手掌图像I2及第一标注信息L,对第二手掌图像I2及第一标注信息L同时执行随机空间变换操作如:平移、旋转、裁切、缩放、透视变换等,得到第三手掌图像I3和第二标注信息L2,对第三手掌图像I3执行图像质量随机扰动如:对比度、亮度、模糊度、色彩通道转换、灰度化等,得到第四手掌图像I4。通过上述背景合成和数据增广可以得到新的训练样本(第四手掌图像I4,第二标注信息L2)。

[0175] 具体的,如图12所示,为本发明实施例提供的数据增强的过程示意图,本发明所述数据增强过程通过获取背景受到限制的RGB手掌图像后,再通过手掌轮廓的区域蒙版对带有自然背景的背景场景图像进行处理得到背景图,通过合成背景图和手掌图像得到背景合成后的图像,之后对该图像进行灰度比、对比度、亮度、饱和度、通道置换、模糊、镜像、裁切、旋转、缩放等方面的随机变换,得到数据增强后的手掌图像。

[0176] 进一步的,本发明提供的弱监督训练策略图如图7所示,本发明所述弱监督训练先对手掌区域分割器进行训练,所述训练通过输入姿态受限单色背景标准手掌图像集,截取背景图像块和手掌皮肤图像块,通过两类像素值的分布训练二分类器,实现手掌区域分割器;对手掌区域分割器训练结束后,进行图像标注,即对手掌区域分割及手掌轮廓生成,再对指根及ROI关键点定位和指根边缘点序列定位;之后进行数据增强,即对手掌区域提取及复杂背景合成后对图像及标注信息执行空间变化,再对图像质量执行扰动;数据增强后即生成训练样本及标签,通过样本及标签训练手掌关键点回归网络,即训练所述手掌感兴趣区域定位网络模型。

[0177] 本发明就整体的手掌关键点坐标回归网络结构配置进行了图示,如图2所示,在本发明中,先获取输入的手掌图像,通过所述主干网络对输入的手掌图像进行处理获取初级特征;之后通过语义分割对手掌图像进行区域分割,得到包含有3个通道的手掌区域特征图,特征图的3个通道输出分别代表了当前像素位置隶属于三个类别的概率(即图1中的BREmap),再对生成的3通道特征图进行通道间的softmax操作得到6个通道的特征图;基于得到的6通道特征图,通过定位网络得到校正参数,基于参数对特征图进行手掌姿态调整,并进行相应的下采样,得到BRE注意力特征图,即调整后的特征图;对调整后的特征图进行基于全局信息的回归和基于局部信息的回归,其中基于全局信息的回归通过多头注意力模

块MHSA和多层感知机模块MLP得到预测坐标,基于局部信息的回归通过Vision Transformer网络模块、CNN卷积模块和DSNT得到预测坐标,之后对得到的基于全局信息和基于局部信息的坐标进行融合回归得到最终的坐标,并通过该坐标进行变换得到ROI图像。进一步的,过程中对手掌及边界分割Focal loss损失、坐标回归MSE损失、边缘最小距离MSE损失进行计算,并通过相应的权重求和得到最终的损失项。此外,在图中所示的①—⑨用于指示网络运行大致方式。

[0178] 进一步的,本发明就手掌关键点坐标回归网络逻辑通过图3中框图进行简述,先像本发明输入手掌图像,通过语义分割提取手掌区域、手掌轮廓区域,从而生成区域特征图(其中包含有手掌区域Seg palm和手掌轮廓区域Segedge),根据区域特征图回归手掌姿态校正参数,根据校正参数对区域特征图执行空间变换并进行指根区域缩放及手掌姿态调整,之后进行关键点坐标的回归,通过关键点坐标进行射影,逆映射变换,从而实现ROI图像提取。进一步,所述进行关键点坐标回归为输入调整后的区域特征图后,进行全局信息和局部信息的坐标回归,并进行基于全局信息的融合权重预测,通过权重对坐标进行融合得到第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标。

[0179] 本发明就所述的方法对多个手掌图像进行ROI提取得到本发明实施例提供的网络模型在自然手掌图像上的定位效果图,如图14所示,通过该图可以得出本发明首先实现手掌区域的语义分割,生成手掌区域及手掌轮廓区域特征图,然后,对上述得到的区域特征图执行姿态变换,对手掌的几何位置、旋转倾斜角度、缩放尺度进行自适应的调整和校正,使得不同用户的手掌特征图在处理之后能够在偏移、旋转、尺度等方面达到归一化的目的;上述手掌及轮廓区域特征图被降采样缩小以减小后端MLP全连接回归网络的计算量,同时,手掌的指根区域被放大以保证回归精度;自适应姿态校正之后,基于全局信息和局部信息对关键点的坐标进行加权回归,自适应加权系数由网络根据区域特征图进行预测得到,上述关键点包括,指根关键点 t_1 、 t_2 ,ROI中心点 c 及四个角点 $r_1 \sim r_4$;得到坐标后生成的ROI图像即在图14中体现出来,其获取的ROI图像都较为清晰、准确。

[0180] 示例性设备

[0181] 如图15中所示,基于上述自然场景下鲁棒的手掌感兴趣区域定位方法,本发明实施例提供了一种自然场景下鲁棒的手掌感兴趣区域定位装置,该装置包括:

[0182] 手掌区域及轮廓提取模块151,用于获取输入的手掌图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;

[0183] 手掌姿态调整模块152,用于将所述手掌区域和轮廓区域特征图输入由弱监督训练策略训练的手掌感兴趣区域定位网络模型,通过训练的手掌感兴趣区域定位网络模型的定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大操作,生成调整后的特征图;

[0184] 手掌关键点坐标融合回归模块153,用于对调整后的特征图进行关键点坐标的融合回归,得到融合回归的第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标;

[0185] 手掌感兴趣区域图像提取模块154,用于基于得到的第二关键点坐标,提取得到手

掌感兴趣区域图像。

[0186] 其中,进一步的解释中,所述手掌区域及轮廓提取模块151实现本发明方法对应的网络模型中的主干网络和手掌区域及轮廓提取模块的效果,手掌姿态调整模块152实现本发明方法对应的网络模型中的手掌姿态调整模块的效果,手掌关键点坐标融合回归模块153实现本发明方法对应的网络模型中的手掌关键点坐标融合回归模块的效果,手掌感兴趣区域图像提取模块154实现本发明方法对应的网络模型中的手掌感兴趣区域图像提取模块的效果。

[0187] 本发明还提供了一种特征提取识别器,其特征在于,所述特征提取器对上述所述的自然场景下鲁棒的手掌感兴趣区域定位方法获取的手掌感兴趣区域图像进行手掌生物信息检测,基于检测到的手掌生物信息实现用户身份识别及核验。

[0188] 具体的,手掌ROI区域包含丰富的掌纹特征包括:三大主线(感情线、智慧线、生命线)、褶皱线、类似指纹的谷线、脊线、及细节点、皮肤层纹理、以及皮下静脉血管纹路。通过特征提取识别器能够对纹线的主方向、纹线的交叉点位置、皮肤纹理的局部梯度分布、静脉血管的空间分布等信息进行提取,再通过提取到的信息实现高精度的用户身份识别及核验。

[0189] 由本发明所述方法得到的ROI区域的定位能够有效地降低手掌姿态变化和手掌空间位置变化给后续特征识别带来的干扰,使得同一用户在不同场景下所采集得到的掌纹特征具有高度的一致性,作为掌纹识别的先决条件,ROI定位是非接触式掌纹识别系统落地的有力保证,从而使得本发明所述的特征提取识别器能够更加准确的对用户身份进行识别及核验。

[0190] 基于上述实施例,本发明还提供了一种终端设备。所述终端设备包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的自然场景下鲁棒的手掌感兴趣区域定位程序,所述处理器执行所述自然场景下鲁棒的手掌感兴趣区域定位程序时,实现上述自然场景下鲁棒的手掌感兴趣区域定位方法的步骤。

[0191] 综上所述,本发明提供了一种自然场景下鲁棒的手掌感兴趣区域定位方法、装置、特征提取识别器及设备,所述方法包括获取输入的手掌图片,对输入图像中的手掌区域及手掌轮廓区域进行检测,得到手掌区域和轮廓区域特征图;将所述手掌区域和轮廓区域特征图输入训练的手掌感兴趣区域定位网络模型,通过由弱监督训练策略训练的手掌感兴趣区域定位网络模型的定位网络对所述手掌区域和轮廓区域特征图进行回归,得到手掌的姿态校正参数,基于所述姿态校正参数对手掌区域特征图进行空间变化,并对图像进行降采样和指根放大操作,生成调整后的特征图;对调整后的特征图进行关键点坐标的融合回归,得到融合回归的第一关键点坐标,并将该坐标基于上述手掌姿态校正参数逆变换回原输入图像的坐标系中,得到对应于输入手掌图像的手掌感兴趣区域第二关键点坐标;基于得到的第二关键点坐标,提取得到手掌感兴趣区域图像。这样本发明能够通过语义分割网络实现对手掌区域及手掌轮廓区域特征图的区分,再对得到的图像特征图进行姿态调整校正使得不同手掌姿态校正归一化,为后续网络模块提供高质量且姿态一致的输入样本,之后对得到的姿态进行校正过后的图像特征图进行基于全局信息和局部信息的关键点的坐标加权回归,使得全局和局部两种信息得以融合,实现坐标回归的鲁棒性,最后通过坐标执行射影变换得到手掌感兴趣区域图像,达到不需要红外相机的辅助的情况下对包含有复杂背景

的手掌图像的处理;同时在本发明实现的过程中,网络在特征提取时,关注指根边缘区域,得出关键点回归的指根边缘距离损失,为后续坐标回归提供高质量的特征输入,提升回归精度及鲁棒性;此外本发明在训练手掌感兴趣区域定位网络模型的过程中,采用迭代手掌感兴趣区域定位算法,其中结合了异常检测功能模块,且使用基于距离极值点检测的手掌感兴趣区域定位方法以及基于线扫描的手掌感兴趣区域定位方法,使得其既能够剔除错误的手掌感兴趣区域定位,又可以迭代式搜索来提升整体的手掌感兴趣区域定位成功率,而在采用迭代手掌感兴趣区域定位算法得到标定信息后,本发明一种自然场景手掌数据集扩增方法,对原图像进行背景合成后,再对背景合成后得到的图像进行数据增广,从而实现为定位网络提供更加逼真和丰富的训练样本、进一步增强网络的泛化能力的效果;同时本发明对由自然场景下鲁棒的手掌感兴趣区域定位方法得到的手掌感兴趣区域图像进行手掌生物信息检测,可以实现高精度的用户身份识别及核验。

[0192] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者装置不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者装置所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者装置中还存在另外的相同要素。

[0193] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一非易失性计算机可读存储介质中,该计算机程序在执行时,可包括如上述各方法的实施例的流程。其中,本发明所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。作为说明而非局限,RAM以多种形式可得,诸如静态RAM(SRAM)、动态RAM(DRAM)、同步DRAM(SDRAM)、双数据率SDRAM(DDRSDRAM)、增强型SDRAM(ESDRAM)、同步链路(Synchlink)DRAM(SLDRAM)、存储器总线(Rambus)直接RAM(RDRAM)、直接存储器总线动态RAM(DRDRAM)、以及存储器总线动态RAM(RDRAM)等。

[0194] 应当理解的是,本发明的应用不限于上述的举例,对本领域的技术人员来说,可以根据上述说明加以改进或变换,所有这些改进和变化都属于本发明所附权利要求的保护范围。

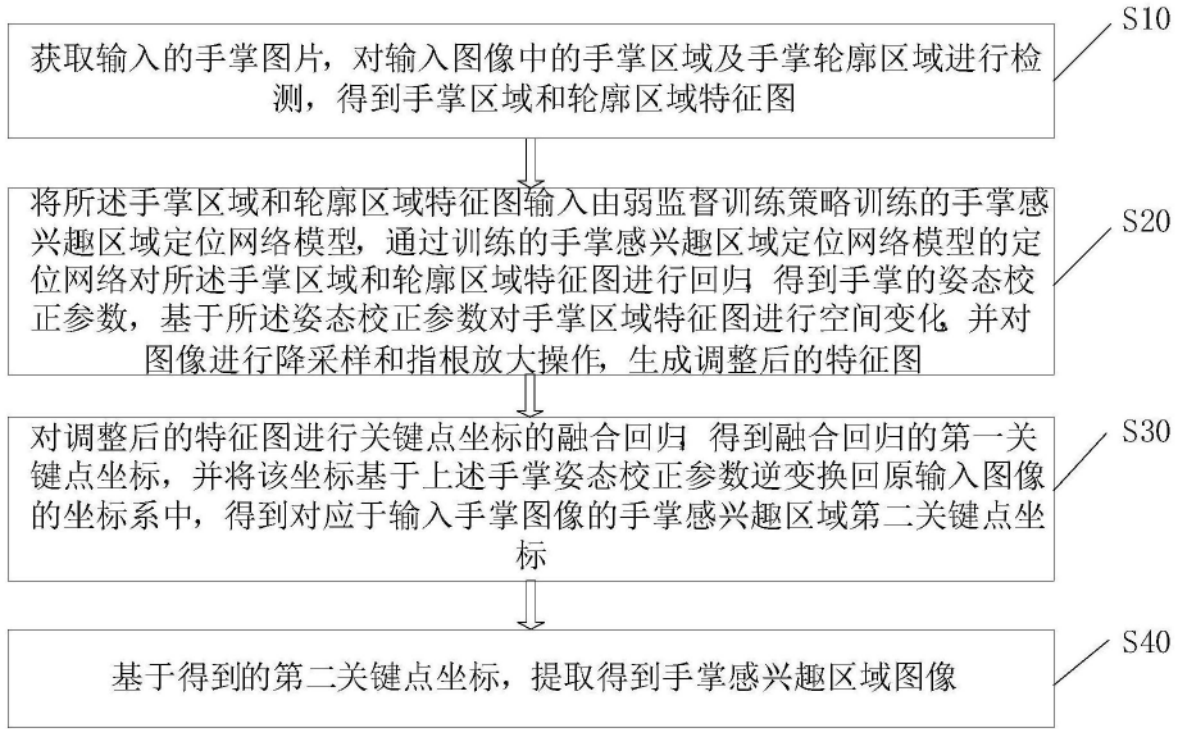


图1

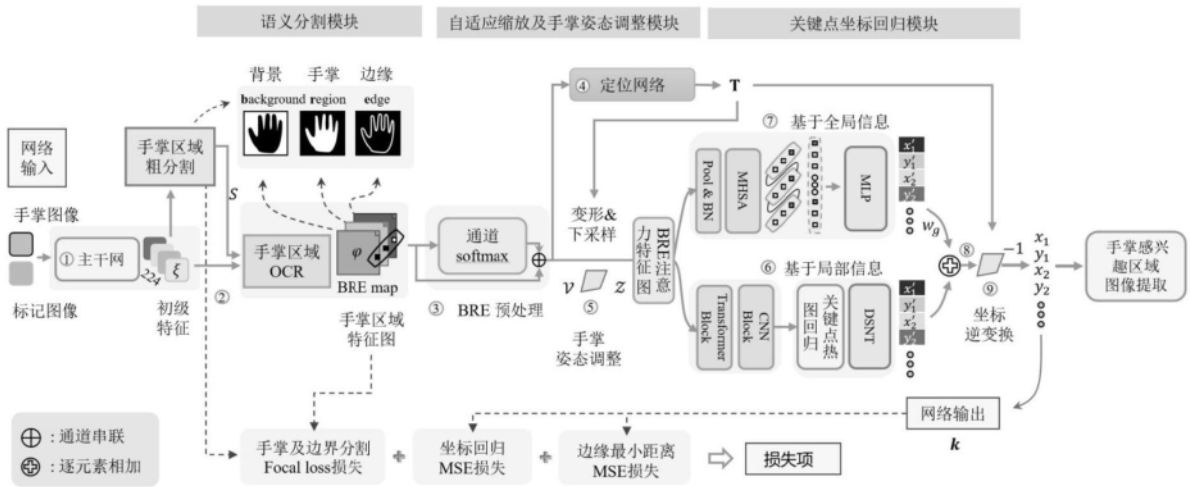


图2

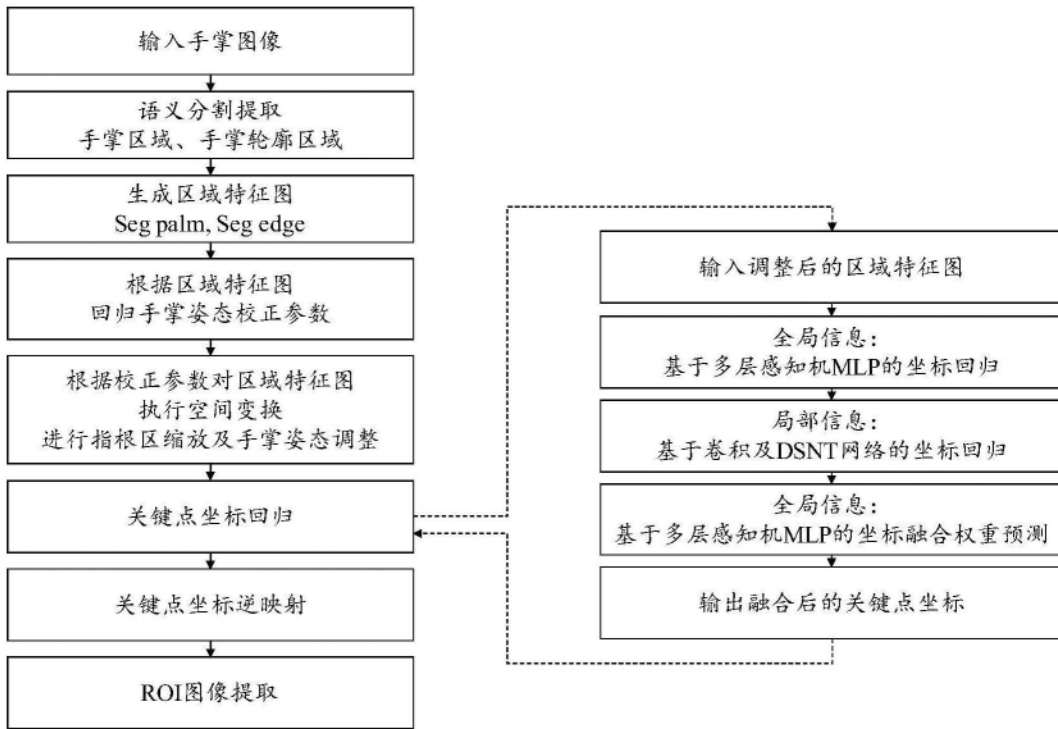


图3

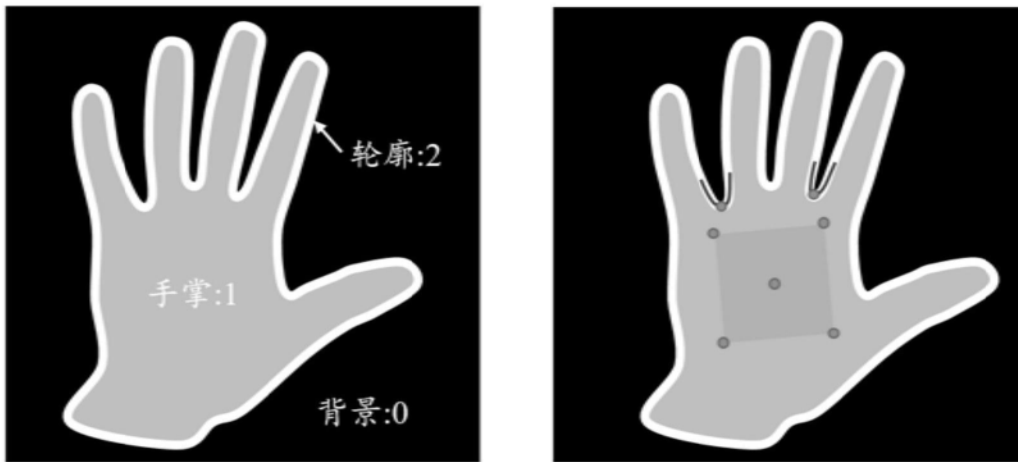


图4

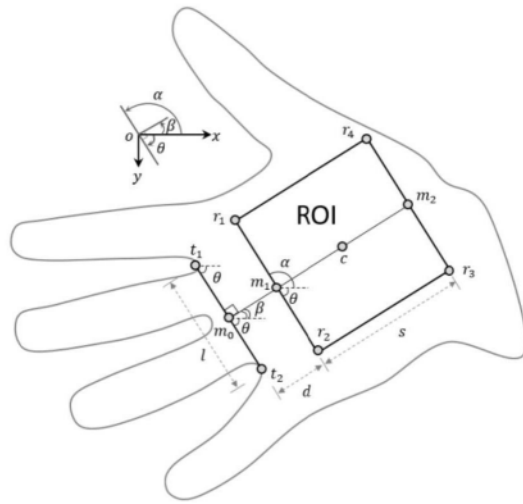


图5

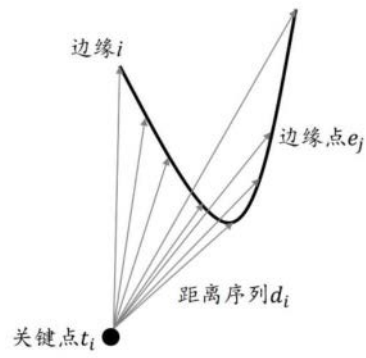


图6

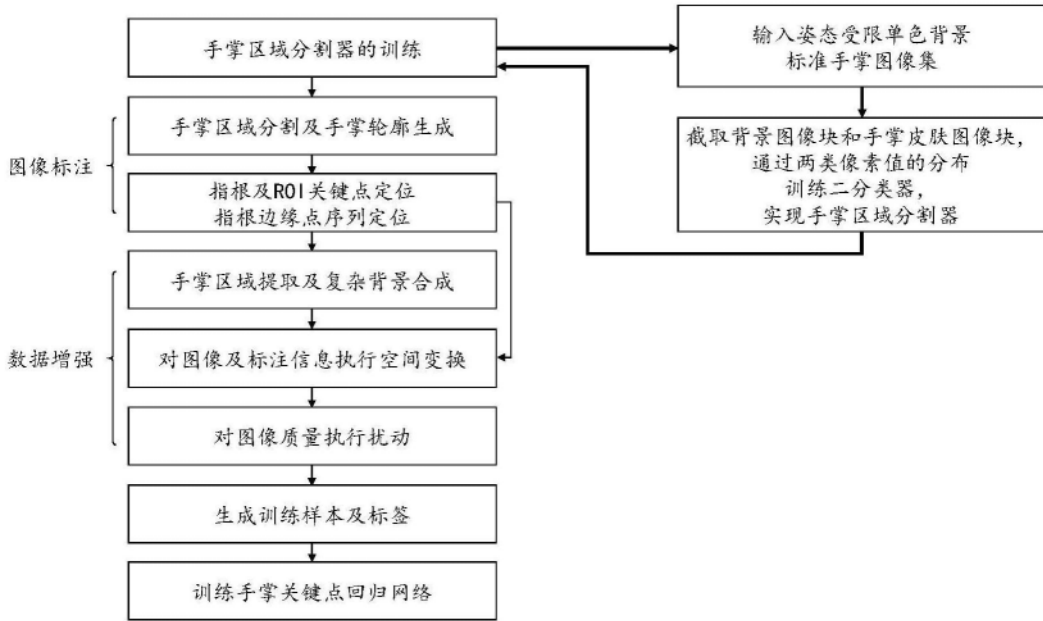


图7

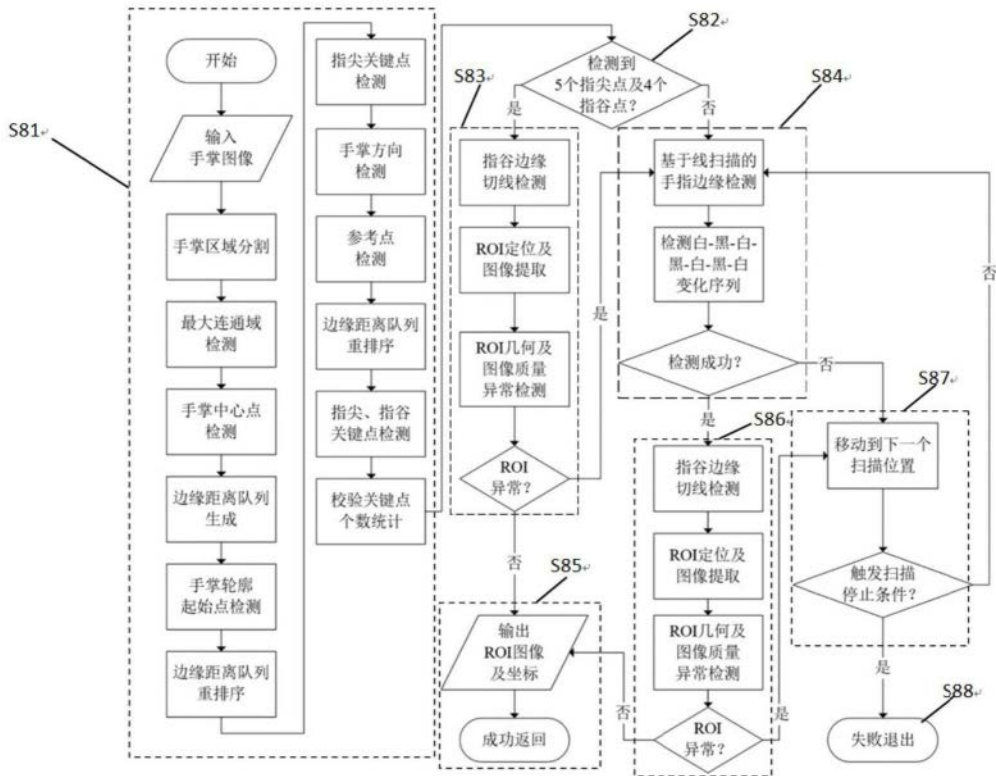


图8

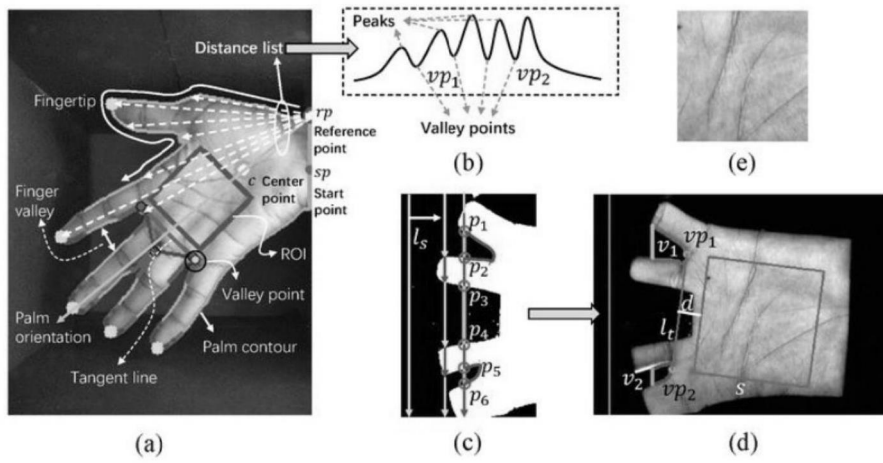


图9

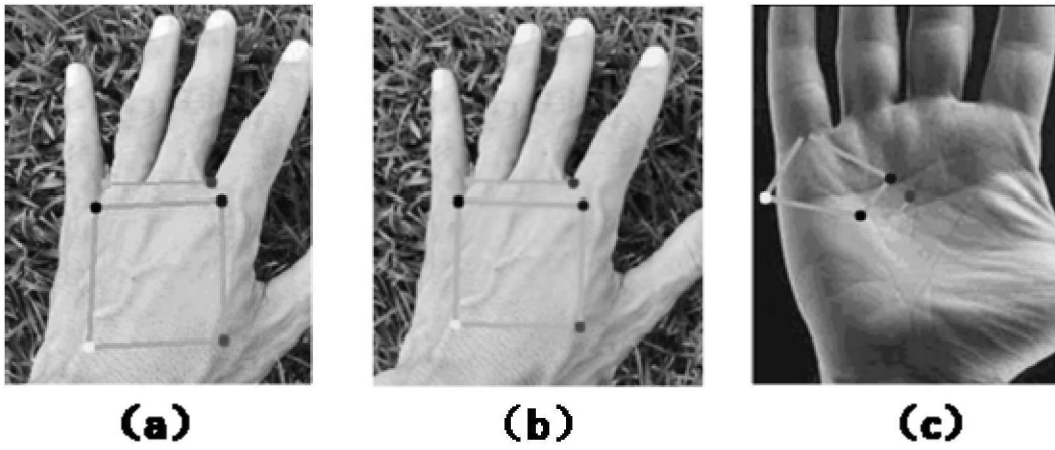


图10

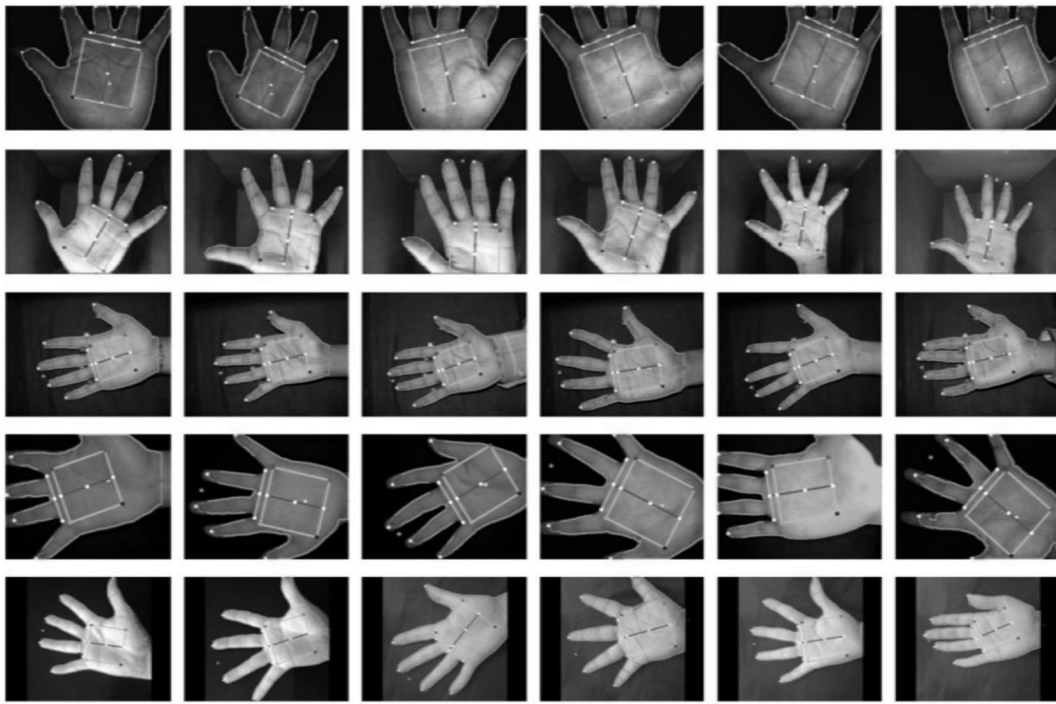


图11

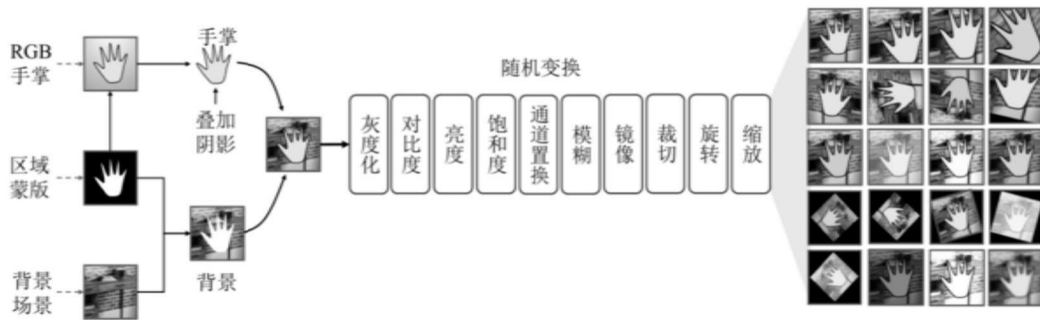


图12

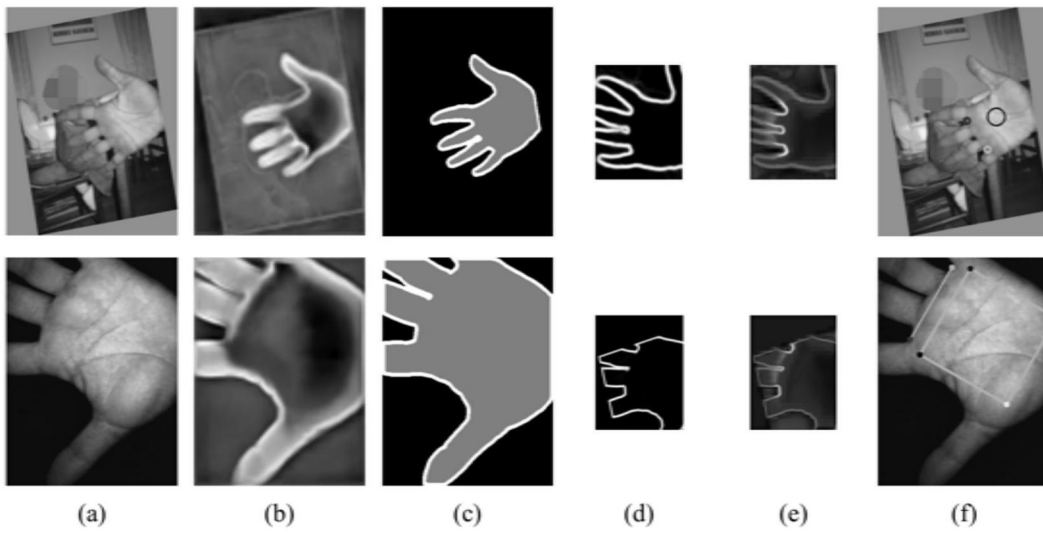


图13



图14

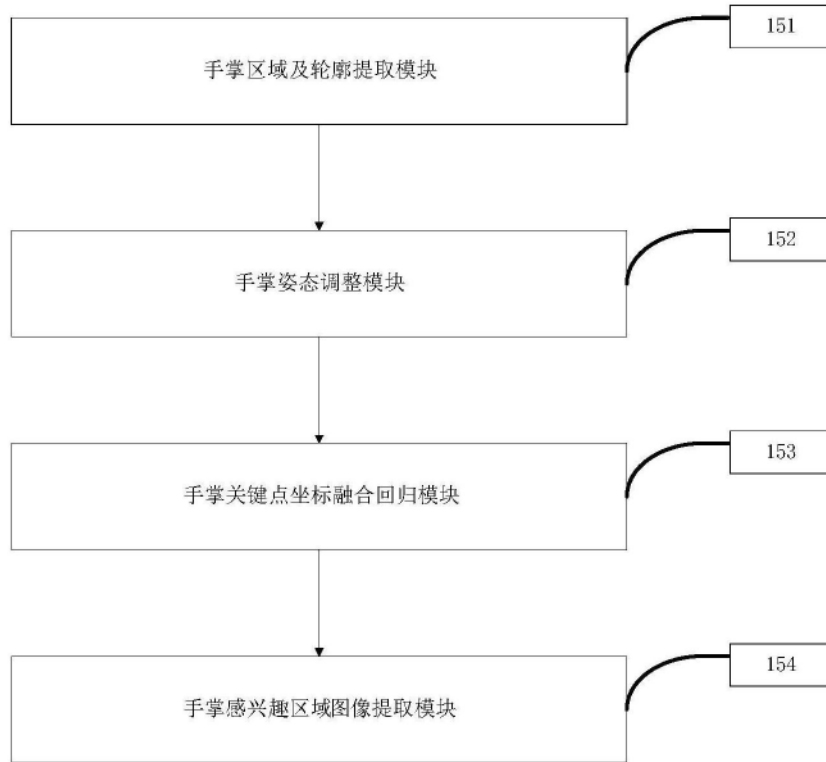


图15