(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2020 年 5 月 7 日 (07.05.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/088092 A1

(51) 国际专利分类号:
G06T 7/70 (2017.01) *G06N 3/02* (2006.01)
G06T 17/00 (2006.01)

(21) 国际申请号: PCT/CN2019/104231

(22) 国际申请日: 2019 年 9 月 3 日 (03.09.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811295915.4 2018年11月1日 (01.11.2018) CN

(71) 申请人: 北京达佳互联信息技术有限公司(**BEIJING DAJIA INTERNET INFORMATION TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区上地西路6号1幢1层101D1-7, Beijing 100085 (CN)。

(72) 发明人: 张雄(**ZHANG, Xiong**); 中国北京市海淀区上地西路6号1幢1层101D1-7, Beijing 100085 (CN)。 李强(**LI, Qiang**); 中国北京市海淀区上地西路6号1幢1层101D1-7, Beijing 100085 (CN)。 郑文(**ZHENG, Wen**); 中国北京市海淀区上地西路6号1幢1层101D1-7, Beijing 100085 (CN)。

(74) 代理人: 北京柏杉松知识产权代理事务所（普通合伙）(**PATENTSINO IP FIRM**); 中国北京市西城区北三环中路27号商房大厦413室, Beijing 100029 (CN)。

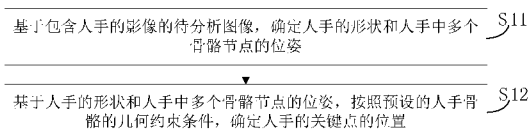
(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(54) **Title:** KEY POINT POSITION DETERMINING METHOD AND APPARATUS, AND ELECTRONIC DEVICE

(54) 发明名称: 关键点位置确定方法、装置及电子设备



S11 Determine the shape of a hand and poses of a plurality of skeleton nodes in the hand on the basis of an image to be analyzed comprising the image of the hand

S12 On the basis of the shape of the hand and the poses of the plurality of skeleton nodes, determine the positions of key points of the hand according to a preset geometric constraint condition of the skeleton of the hand

图 1

(57) **Abstract:** The present invention relates to a key point position determining method and apparatus, and an electronic device. The method comprises: determining the shape of a hand and poses of a plurality of skeleton nodes in the hand on the basis of an image to be analyzed comprising the image of the hand; and on the basis of the shape and the poses, determining the positions of key points of the hand according to a preset geometric constraint condition of the skeleton of the hand. According to the present invention, the position of a shielded key point can be accurately calculated by utilizing the geometric constraint condition of the hand bone of the hand in combination with hand parameters extracted from the image to be analyzed.

(57) 摘要: 本公开是关于关键点位置确定方法、装置及电子设备。其中方法包括：基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。本公开可以利用人手手骨自身所存在的几何约束条件，结合从待分析图像中提取到的人手的参数，较为准确地推算出被遮挡的关键点的位置。

WO 2020/088092 A1

关键点位置确定方法、装置及电子设备

本申请要求于 2018 年 11 月 01 日提交中国专利局、申请号为 201811295915.4，发明名称为“关键点位置确定方法、装置及电子设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本公开涉及图像处理技术领域，尤其涉及关键点位置确定方法、装置及电子设备。

背景技术

相关技术中，可以利用人手的多个关键点所处的三维空间位置，对人手的姿态进行估计。针对每个关键点，可以基于人手色彩图像如人手 RGB（Red Green Blue，红绿蓝）图像，计算该人手色彩图像中各个像素点为该关键点的概率，得到该关键点在二维图像上的概率分布，并基于该概率分布，利用预先经过训练的神经网络，计算得到该关键点的三维空间位置。

但是，在人手色彩图像中，关键点可能处于被遮挡的状态，在这种情况下，关键点在人手色彩图像中所处的位置可能呈现的是障碍物的图像，可能导致在计算概率分布时，得到的该位置为关键点所处的位置的概率较低，与实际情况存在较大偏差。进而造成无法计算出该关键点的三维空间位置，或者计算出的该关键点的三维空间位置准确性较低，从而影响人手的姿态估计。

发明内容

为克服相关技术中存在的问题，本公开提供一种关键点位置确定方法、装置及电子设备。

根据本公开实施例的第一方面，提供一种关键点位置确定方法，包括：

基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

结合第一方面，在第一种可能的实现方式中，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置，

包括：

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

结合第一方面的第一种可能的实现方式，在第二种可能的实现方式中，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型，包括：

将所述形状和所述位姿作为模型参数输入至预设的骨骼动画框架中，得到所述人手的三维骨骼模型，所述骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

结合第一方面，在第三种可能的实现方式中，所述基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿，包括：

将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

结合第一方面的第三种可能的实现方式，在第四种可能的实现方式中，所述参数提取网络为移动端神经网络 MobileNet。

根据本公开实施例的第二方面，提供一种关键点位置确定装置，包括：

人手分析单元，被配置为执行基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；

位置确定单元，被配置为执行基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

结合第二方面，在第一种可能的实现方式中，所述位置确定单元，包括：

构建模块，具体被配置为执行基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

读取模块，被配置为执行从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

结合第二方面的第一种可能的实现方式，在第二种可能实现方式中，所述构建模块，具体被配置为执行将所述形状和所述位姿作为模型参数输入至预设的骨骼动画框架中，得到所述人手的三维骨骼模型，所述骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

结合第二方面，在第三种可能的实现方式中，所述人手分析单元，具体被配置为执行将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

结合第二方面的第三种可能的实现方式，在第四种可能的实现方式中，所述参数提取网络为移动端神经网络 MobileNet。

根据本公开实施例的第三方面，提供了一种电子设备，包括：

处理器；

用于存储处理器可执行指令的存储器;

其中，所述处理器被配置为：

基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

结合第三方面，在第一种可能的实现方式中，所述处理器，具体被配置为执行：

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

结合第三方面的第一种可能的实现方式，在第二种可能的实现方式中，所述处理器具体被配置为执行：

将所述形状和所述位姿作为模型参数输入至预设的骨骼动画框架中，得到所述人手的三维骨骼模型，所述骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

结合第三方面，在三种可能的实现方式中，所述处理具体被配置为执行：

将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

结合第三方面的第三种可能的实现方式，在第四种可能的实现方式中，所述参数提取网络为移动端神经网络 MobileNet。

根据本公开实施例的第四方面，提供了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关键点位置确定方法，所述方法包括：

基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

结合第四方面，在第一种可能的实现方式中，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置，包括：

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

结合第四方面的第一种可能的实现方式，在第二种可能的实现方式中，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置，包括：

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

结合第四方面，在第三种可能的实现方式中，所述基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿，包

括：

将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

结合第四方面的第三种可能的实现方式中，在第四种可能的实现方式中，所述参数提取网络为移动端神经网络 MobileNet。

根据本公开实施例的第五方面，提供一种计算机程序产品，当所述计算机程序产品由用户终端的处理器执行时，使得用户终端能够执行一种关键点位置确定方法，所述方法包括：

基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

结合第五方面，在第一种可能的实现方式中，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置，包括：

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

结合第五方面的第一种可能的实现方式，在第二种可能的实现方式中，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型，包括：

将所述形状和所述位姿作为模型参数输入至预设的骨骼动画框架中，得到所述人手的三维骨骼模型，所述骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

结合第五方面，在第三种可能的实现方式中，所述基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿，包括：

将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

结合第五方面的第三种可能的实现方式，在第四种可能的实现方式中，所述参数提取网络为移动端神经网络 MobileNet。

本公开的实施例提供的技术方案可以包括以下有益效果：利用人手手骨自身所存在的几何约束条件，结合从待分析图像中提取到的人手的参数，可以较为准确地推算出被遮挡的关键点的位置。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

为了更清楚地说明本发明实施例和现有技术的技术方案，下面对实施例和现有技术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发明的一些实施例，对于本领域普通技术人员来讲，在不付出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。

图 1 是根据一示例性实施例示出的一种关键点位置确定方法的流程图；

图 2a 是根据一示例性实施例示出的一种人手手骨骨骼结构图；

图 2b 是根据一示例性实施例示出的一种关键点分布图；

图 3 是根据一示例性实施例示出的另一种关键点位置确定方法的流程图；

图 4 是根据一示例性实施例示出的一种关键点位置确定装置的结构图；

图 5 是根据一示例性实施例示出的一种用于关键点位置确定的装置的框图；

图 6 是根据一示例性实施例示出的另一种用于关键点位置确定的装置的框图。

具体实施方式

为使本发明的目的、技术方案、及优点更加清楚明白，以下参照附图并举实施例，对本发明进一步详细说明。显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

图 1 是根据一示例性实施例示出的一种关键点位置确定方法的流程图，如图 1 所示，关键点位置确定方法可以应用于终端中，包括以下步骤：

在步骤 S11 中，基于包含人手的影像的待分析图像，确定人手的形状和人手中多个骨骼节点的位姿。

其中，待分析图像可以是拍摄人手得到的色彩图像（如 RGB 图像），该待分析图像中的人手的影像可以部分被遮挡，也可以没有被遮挡，对于没有被遮挡的情况，由于不存在本公开实施例所需要解决的技术问题，此处不做讨论。每个骨骼节点均为预设的人手中的一个点，示例性的，可以预先将人手中食指的掌指关节设置为一个骨骼节点。骨骼节点的数量可以根据实际需求或者用户经验进行设置，例如为了更准确地确定关键点的位置，可以预先设置数量相对较多的骨骼节点，又例如为了降低确定关键点的位置所占用的计算资源，可以预先设置数量相对较少的骨骼节点。在本步骤中，一个骨骼节点的位姿可以包括该骨骼节点在三维空间坐标系中所处的位置以及角度。可以理解，虽然不同的人的人手的形状整体上相近，但是存在一些因人而异的变量，例如人手的长短、宽窄、胖瘦等，可以通过确定的人手的形状体现这些因人而异的变量。

具体的，可以是利用预设的图像识别算法对待分析图像进行处理，得到人手的形状和多个骨骼节点的位姿，也可以是利用预先经过训练的神经网络，实现待分析图像到人手的形状以及骨骼节点的位姿之间的端对端映射。

示例性的，在一种实施例中，还可以是将包含人手的影像的待分析图像，输入预设的参数提取网络，得到人手的形状和人手中多个骨骼节点的位姿。其中，参数提取网络预先经过多个标注有真值的样本图像的训练，每个样本图像中包括有人手的影像，并且标注的真值可以为该人手的形状以及多个骨骼节点的位姿，参数提取网络的训练方式可以是上述将样本图像输入至参数提取网络，并计算参数提取网络的输出与标注的真值之间的损失函数，基于损失函数利用随机梯度下降法调整参数提取网络的网络参数。

参数提取网络可以是大型的卷积神经网络，如 U-net、HourGlass，也可以是轻量型的卷积神经网络 MobileNet，MobileNet 相比于 U-net 或者 HourGlass，网络的结构复杂度更低，所占用的计算资源也相对较少，更容易在移动终端

上运行。而 U-net 或 HourGlass 由于网络的结构复杂度相对较高，因此占用的计算资源相对较多，难以或者无法在移动终端上运行。

在步骤 S12 中，基于人手的形状和人手中多个骨骼节点的位姿，按照预设的人手骨骼的几何约束关系几何约束条件，确定人手的关键点的位置。

人手中存在手骨，手骨包括多块骨头，可以如图 2a 所示，这些骨头能够在手部肌肉的带动下运动，但是由于这些骨头之间可能存在相互连接的关系，因此骨头的运动受到了限制。例如，中指的中节指骨可以相对于第一关节向手心方向和/或手背方向弯曲，而对于正常人无法或者难以相对于第一关节向食指方向或者无名指方向弯曲（以下称这两个方向为侧向），因此可以预先设置几何约束条件：中指的近节指骨和中节指骨的侧向角相同。假设中指的近节指骨（即图中 201）的长度为 3cm，由于中指的近节指骨与中指的中节指骨（即图中 203）相关节（以下称该相关节处为第一关节），并且与掌骨头（即图中 202）相关节（以下称该相关节处为第二关节），如果将中指的近节指骨视为刚体，则可以预先设置几何约束条件：第一关节与第二关节之间的距离为 3cm。

在人手姿态估计中所使用的关键点可以为手骨的关节处，例如常用的人手姿态估计算法中，使用手骨的 21 个关节处作为 21 个关键点，具体的分布可以参见图 2b 所示，其中 1-21 分别表示这 21 个关键点所处的位置，因此关键点的位置并非是任意分布的，而是受到人手骨骼的几何约束条件影响的。示例性的，假设在 S11 中确定得到中指的第一关节在三维空间坐标系（假设该三维空间坐标系的单位为 cm，手心方向为 x 轴的正方向）中的坐标为(0,0,0)，中指的近节指骨长度为 3cm，并且相对中指的第二关节向手心方向弯曲 90°，第一关节为关键点，则可以确定得到该关键点的坐标为（3，0，0），并且由于该关键点的位置是基于人手的形状（近节指骨长 3cm）以及第二关节和近节指骨的位姿确定得到的，因此即使在待分析图像中第一关节被遮挡，也能够确定出第一关节的位置。

选用该实施例，可以利用人手骨骼固有的几何约束条件，结合待分析图像中提取到的人手的形状以及多个骨骼节点的位姿，对关键点的位置进行推算，并不直接从待分析图像中识别并定位关键点，因此即使关键点被遮挡住，

仍然能够确定出关键点的位置。

参见图 3，图 3 所示为本公开实施例提供的另一种关键点位置确定方法的流程示意图，包括以下步骤：

在步骤 S31 中，基于包含人手的影像的待分析图像，确定人手的形状和人手中多个骨骼节点的位姿。

该步骤与 S11 相同，可以参见前述关于 S11 的描述，在此不再赘述。

在步骤 S32 中，将得到的形状和位姿作为模型参数输入至预设的骨骼动画框架中，得到人手的三维骨骼模型。

其中，骨骼动画框架可以为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。不同人的人手的三维骨骼模型之间可能存在一定的差异，但是同样作为人类，不同人的人手的三维骨骼模型之间也存在一定的共性，如手骨中各块骨头之间的连接关系、能够进行的运动方式，因此不同人的人手的三维骨骼模型中可能存在一些相同的几何约束条件，在本实施例中，可以从多个样本人手三维骨骼模型中提取这些几何约束条件。

进一步的，样本人手三维骨骼模型可以根据实际需求进行选取，例如，为了扩展该实施例的适用人群，可以选用不同种族、不同年龄段、不同性别的人手的三维骨骼模型作为样本人手三维骨骼模型。又例如，已经确定该实施例的面向群体为亚洲青年、中年群体，则可以选用亚裔并且年龄段位于 12-40 岁的男性以及女性的人手的三维骨骼模型作为样本人手三维骨骼模型。

虽然不同的人的人手的三维骨骼模型之间存在一定的共性，但也存在一些因人而异的差异，或者随时间变化的变量，例如一部分人的中指近节指骨可能长 3cm，另外一部分人的中指近节指骨可能长 3.5cm。又例如在某一时刻人手处于紧握状态，中指近节指骨相对第二关节向掌心弯曲，在另一个时刻人手处于平摊状态，中指近节指骨可能没有相对第二关节向掌心弯曲。因此在构建人手的三维骨骼模型时，需要输入人手的形状以及多个骨骼节点的位姿作为模型参数。

在步骤 S33 中，从三维骨骼模型中读取人手的关键点的三维空间坐标，作为关键点的位置。

在构建完成人手的三维骨骼模型后，可以视为已知人手的手骨上任意点

的位置，因此可以从三维骨骼模型中读取人手的关键点的三维空间坐标。选用该实施例，关键点是从人手的三维骨骼模型中读取得到的，因此关键点之间符合从多个样本人手的三维骨骼模型中提取到的几何约束条件，因此可以认为关键点的分布符合人手中各个关节处的真实分布情况，进而可以认为此时得到的关键点的位置更加准确。

图 4 是根据一示例性实施例示出的一种关键点位置确定装置框图。参照图 4，该装置包括人手分析单元 401，位置确定单元 402。

该人手分析单元 401 被配置为执行基于包含人手的影像的待分析图像，确定人手的形状和人手中多个骨骼节点的位姿；

该位置确定单元 402 被配置为执行基于形状和位姿，按照预设的人手骨骼的几何约束条件，确定人手的关键点的位置。

在一种示例性实施例中，该位置确定单元 402，包括：

构建模块，被配置为执行基于形状和位姿，按照预设的人手骨骼的几何约束条件，构建人手的三维骨骼模型；

读取模块，被配置为执行从三维骨骼模型中读取人手的关键点的三维空间坐标，作为关键点的位置。

在一种示例性实施例中，该构建模块，可以具体被配置为执行将形状和位姿作为模型参数输入至预设的骨骼动画框架中，得到人手的三维骨骼模型，骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

在一种示例性实施例中，该人手分析单元 401，可以具体被配置为执行将包含人手的影像的待分析图像，输入预设的参数提取网络，得到人手的形状和人手中多个骨骼节点的位姿，参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

在一种示例性实施例中，参数提取网络为移动端神经网络 MobileNet。

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。

图 5 是根据一示例性实施例示出的一种用于关键点位置确定的装置 500 的框图。例如，装置 500 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。

参照图 5，装置 500 可以包括以下一个或多个组件：处理组件 502，存储器 504，电源组件 506，多媒体组件 508，音频组件 510，输入/输出（I/O）接口 512，传感器组件 514，以及通信组件 516。

处理组件 502 通常控制装置 500 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 502 可以包括一个或多个处理器 520 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 502 可以包括一个或多个模块，便于处理组件 502 和其他组件之间的交互。例如，处理组件 502 可以包括多媒体模块，以方便多媒体组件 508 和处理组件 502 之间的交互。

存储器 504 被配置为存储各种类型的数据以支持在设备 500 的操作。这些数据的示例包括用于在装置 500 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 504 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。

电源组件 506 为装置 500 的各种组件提供电力。电源组件 506 可以包括电源管理系统，一个或多个电源，及其他与为装置 500 生成、管理和分配电力相关联的组件。

多媒体组件 508 包括在装置 500 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。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 508 包括一个前置摄像头和/或后置摄像头。当设备 500 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。

音频组件 510 被配置为输出和/或输入音频信号。例如，音频组件 510 包括一个麦克风（MIC），当装置 500 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 504 或经由通信组件 516 发送。在一些实施例中，音频组件 510 还包括一个扬声器，用于输出音频信号。

I/O 接口 512 为处理组件 502 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。

传感器组件 514 包括一个或多个传感器，用于为装置 500 提供各个方面的状态评估。例如，传感器组件 514 可以检测到设备 500 的打开/关闭状态，组件的相对定位，例如组件为装置 500 的显示器和小键盘，传感器组件 514 还可以检测装置 500 或装置 500 一个组件的位置改变，用户与装置 500 接触的存在或不存在，装置 500 方位或加速/减速和装置 500 的温度变化。传感器组件 514 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 514 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 514 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。

通信组件 516 被配置为便于装置 500 和其他设备之间有线或无线方式的通信。装置 500 可以接入基于通信标准的无线网络，如 WiFi，运营商网络（如 2G、3G、4G 或 5G），或它们的组合。在一个示例性实施例中，通信组件 516 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 516 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。

在示例性实施例中，装置 500 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。

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 504，上述指令可由装置 500 的处理器 520 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。

在示例性实施例中，还提供了一种计算机程序产品，该计算机程序产品可以存储于计算机可读存储介质中，例如存储器 504，该计算机程序产品可由 500 的处理器 520 执行以完成上述方法。

图 6 是根据一示例性实施例示出的另一种用于关键点位置确定的装置 600 的框图。例如，装置 600 可以被提供为一服务器。参照图 6，装置 600 包括处理组件 622，其进一步包括一个或多个处理器，以及由存储器 632 所代表的存储器资源，用于存储可由处理组件 622 的执行的指令，例如应用程序。存储器 632 中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外，处理组件 622 被配置为执行指令，以执行上述方法。

装置 600 还可以包括一个电源组件 626 被配置为执行装置 600 的电源管理，一个有线或无线网络接口 650 被配置为将装置 600 连接到网络，和一个输入输出（I/O）接口 658。装置 600 可以操作基于存储在存储器 632 的操作系统，例如 Windows ServerTM，Mac OS XTM，UnixTM, LinuxTM，FreeBSDTM 或类似。

需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，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、物品或者设备不仅包括那些要素，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、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

本说明书中的各个实施例均采用相关的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同

之处。尤其，对于装置、电子设备以及存储介质实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

以上所述仅为本发明的较佳实施例而已，并不用以限制本发明，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明保护的范围之内。

权 利 要 求

1、一种关键点位置确定方法，其特征在于，包括：

基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

2、根据权利要求 1 所述的方法，其特征在于，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置，包括：

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

3、根据权利要求 2 所述的方法，其特征在于，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型，包括：

将所述形状和所述位姿作为模型参数输入至预设的骨骼动画框架中，得到所述人手的三维骨骼模型，所述骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

4、根据权利要求 1 所述的方法，其特征在于，所述基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿，包括：

将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

5、根据权利要求 4 所述的方法，其特征在于，所述参数提取网络为移动端神经网络 MobileNet。

6、一种关键点位置确定装置，其特征在于，包括：

人手分析单元，被配置为执行基于包含人手的影像的待分析图像，确定

所述人手的形状和所述人手中多个骨骼节点的位姿；

位置确定单元，被配置为执行基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

7、根据权利要求 6 所述的装置，其特征在于，所述位置确定单元，包括：

构建模块，被配置为执行基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

读取模块，被配置为执行从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

8、根据权利要求 7 所述的装置，其特征在于，所述构建模块，具体被配置为执行将所述形状和所述位姿作为模型参数输入至预设的骨骼动画框架中，得到所述人手的三维骨骼模型，所述骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

9、根据权利要求 6 所述的装置，其特征在于，所述人手分析单元，具体被配置为执行将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

10、根据权利要求 9 所述的装置，其特征在于，所述参数提取网络为移动端神经网络 MobileNet。

11、一种电子设备，其特征在于，包括：

处理器；

用于存储处理器可执行指令的存储器;

其中，所述处理器被配置为执行：

基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

12、根据权利要求 11 所述的电子设备，其特征在于，所述处理器，具体被配置为执行：

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建

所述人手的三维骨骼模型；

从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

13、根据权利要求 12 所述的电子设备，其特征在于，所述处理器具体被配置为执行：

将所述形状和所述位姿作为模型参数输入至预设的骨骼动画框架中，得到所述人手的三维骨骼模型，所述骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

14、根据权利要求 11 所述的电子设备，其特征在于，所述处理具体被配置为执行：

将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

15、根据权利要求 14 所述的电子设备，其特征在于，所述参数提取网络为移动端神经网络 MobileNet。

16、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关键点位置确定方法，所述方法包括：

基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿；

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置。

17、根据权利要求 16 所述的非临时性计算机可读存储介质，其特征在于，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，确定所述人手的关键点的位置，包括：

基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型；

从所述三维骨骼模型中读取所述人手的关键点的三维空间坐标，作为所述关键点的位置。

18、根据权利要求 17 所述的非临时性计算机可读存储介质，其特征在于，所述基于所述形状和所述位姿，按照预设的人手骨骼的几何约束条件，构建所述人手的三维骨骼模型，包括：

将所述形状和所述位姿作为模型参数输入至预设的骨骼动画框架中，得到所述人手的三维骨骼模型，所述骨骼动画框架为从多个样本人手三维骨骼模型中提取几何约束条件得到的参数化的框架。

19、根据权利要求 16 所述的非临时性计算机可读存储介质，其特征在于，所述基于包含人手的影像的待分析图像，确定所述人手的形状和所述人手中多个骨骼节点的位姿，包括：

将包含人手的影像的待分析图像，输入预设的参数提取网络，得到所述人手的形状和所述人手中多个骨骼节点的位姿，所述参数提取网络预先经过标注有人手的形状和各个骨骼节点的位姿的样本图像的训练。

20、根据权利要求 19 所述的非临时性计算机可读存储介质，其特征在于，所述参数提取网络为移动端神经网络 MobileNet。

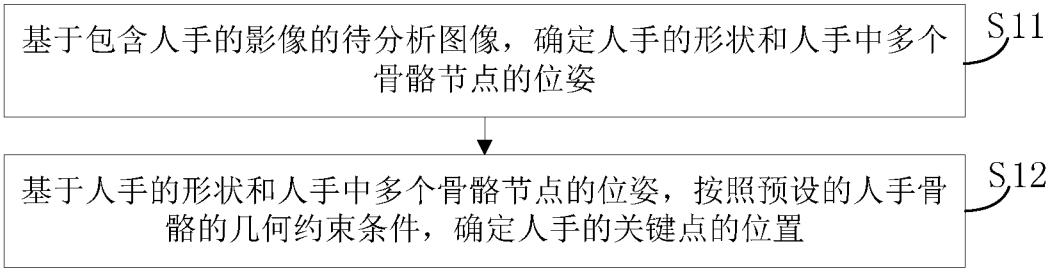

基于包含人手的影像的待分析图像，确定人手的形状和人手中多个骨骼节点的位姿
S11
基于人手的形状和人手中多个骨骼节点的位姿，按照预设的人手骨骼的几何约束条件，确定人手的关键点的位置
S12

图 1

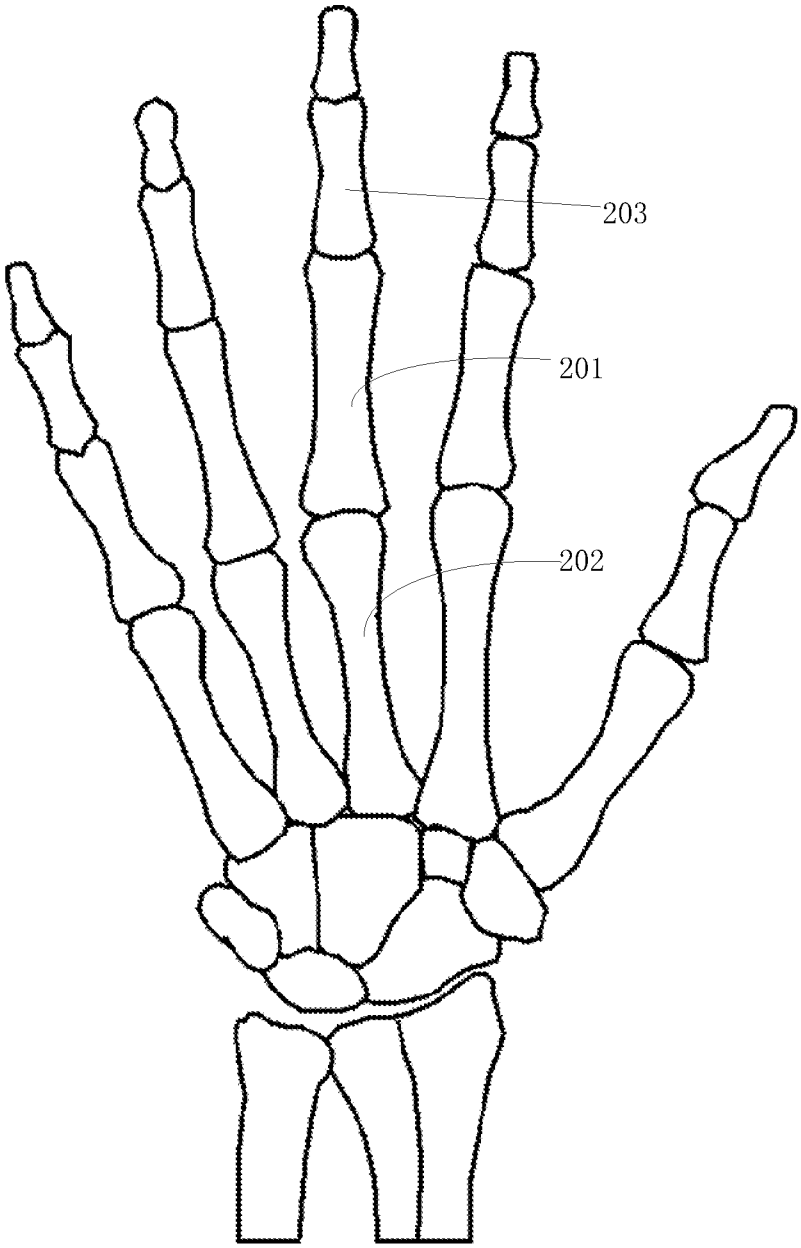


图 2a

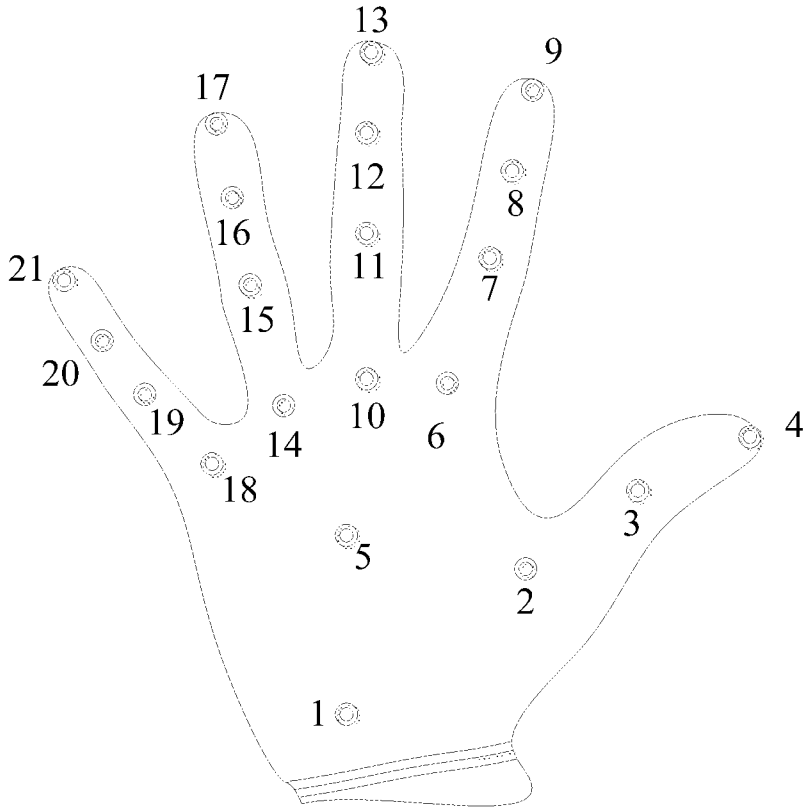

13
9
17
12
8
16
11
7
21
15
20
10
6
4
19
14
18
3
5
2
1

图 2b

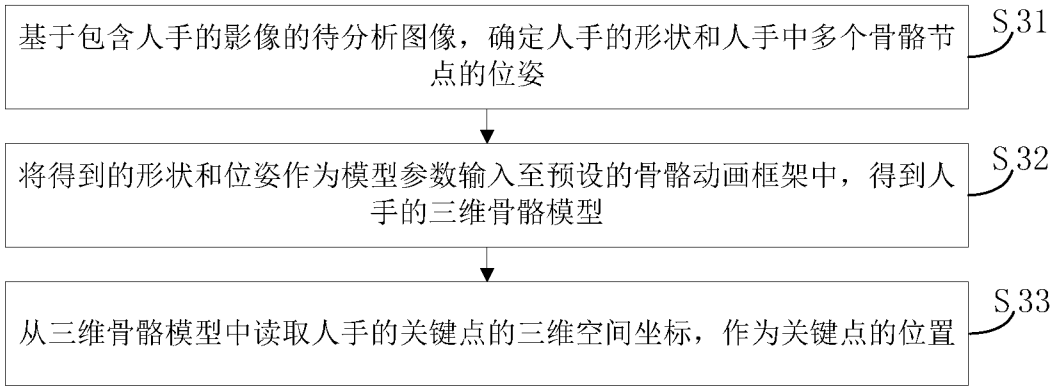

基于包含人手的影像的待分析图像，确定人手的形状和人手中多个骨骼节点的位姿
S31
将得到的形状和位姿作为模型参数输入至预设的骨骼动画框架中，得到人手的三维骨骼模型
S32
从三维骨骼模型中读取人手的关键点的三维空间坐标，作为关键点的位置
S33

图 3

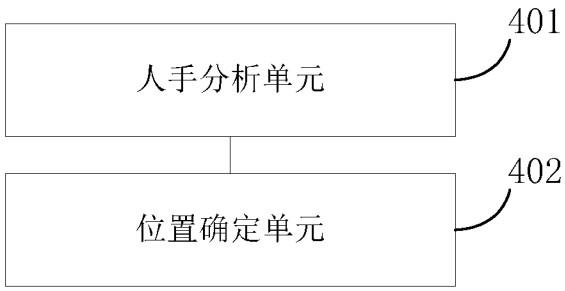

401
人手分析单元
402
位置确定单元

图 4

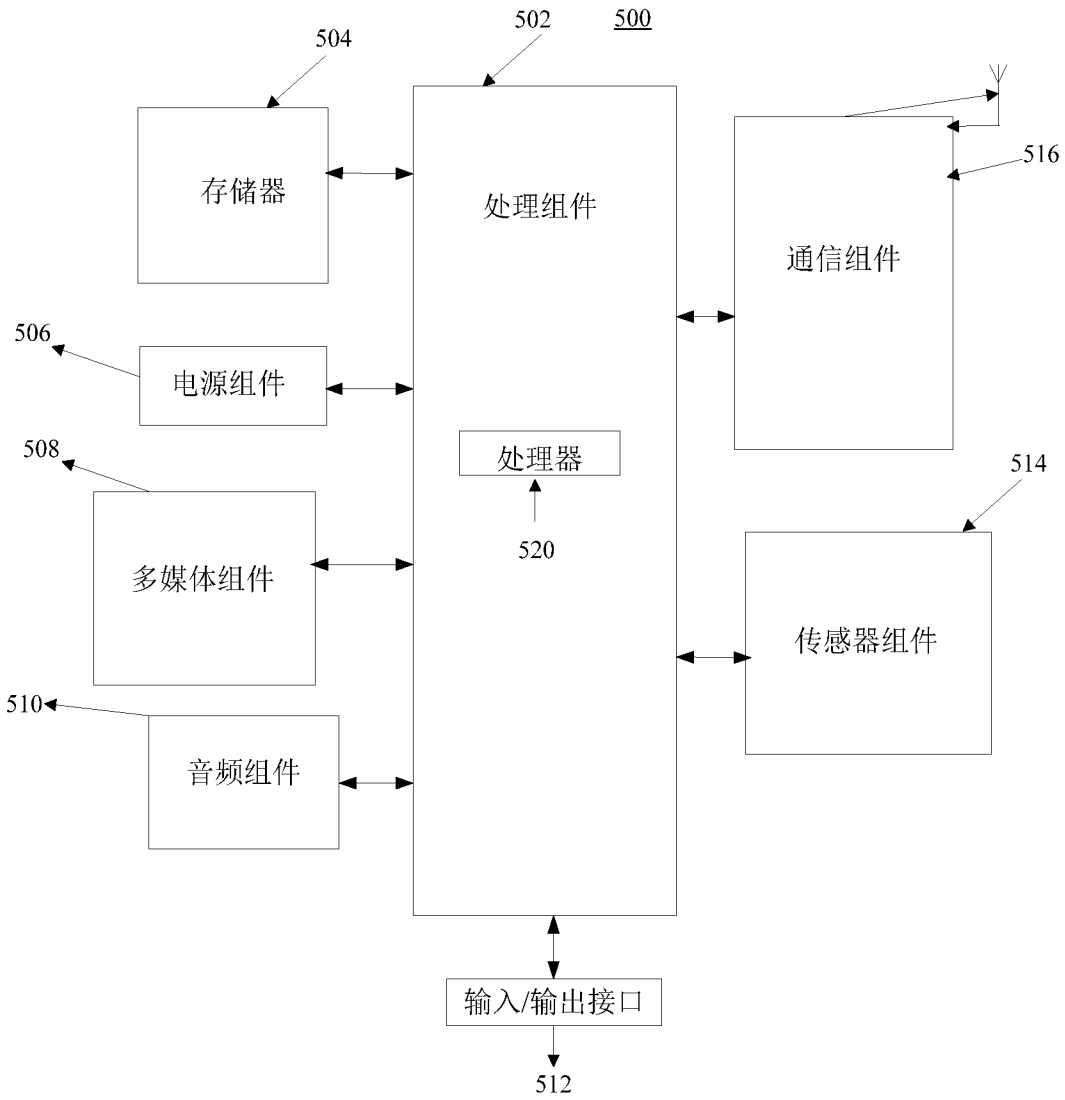

502
500
504
存储器
处理组件
516
通信组件
506
电源组件
处理器
520
508
多媒体组件
514
传感器组件
510
音频组件
输入/输出接口
512

图 5

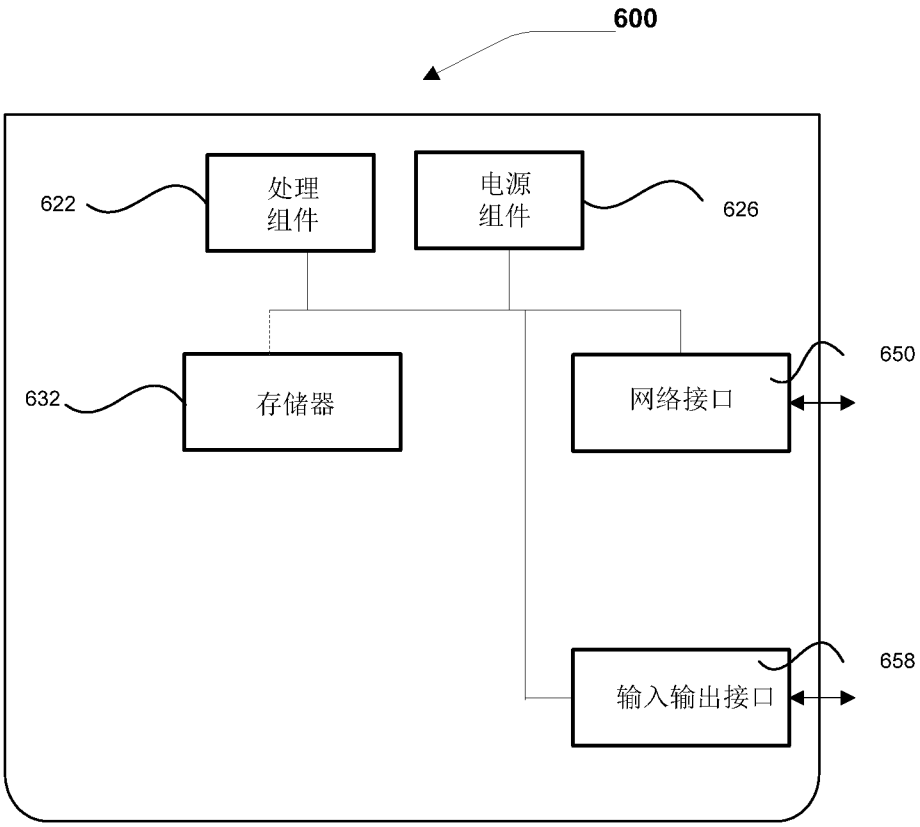


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/104231

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/70(2017.01)i; G06T 17/00(2006.01)i; G06N 3/02(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T; G06K; G06N; G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNAPT; CNKI; EPODOC; WPI; IEEE: 手, 手势, 识别, 骨骼, 关节, 关键点, 几何, 大小, 长度, 尺寸, 模型, 3D, 三维, 神经网络, 参数, hand, gesture, skeleton, bone, geometrical, length, relationship, model, neural network, parameter, dimension, angle

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109410276 A (BEIJING DAJIA INTERCONNECTION INFORMATION TECHNOLOGY CO., LTD.) 01 March 2019 (2019-03-01) claims 1-10	1-20
Y	CN 108399367 A (SHENZHEN AXMTEC TECHNOLOGY CO., LTD.) 14 August 2018 (2018-08-14) description, paragraphs [0050]-[0092], [0104]-[0115], [0120], [0133] and [0134]	1-20
Y	WO 2012095756 A2 (EXTREME REALITY LTD.) 19 July 2012 (2012-07-19) description, paragraphs [0041]-[0050], and figures 3-5C	1-20
A	CN 107767419 A (GUANGZHOU DEEP DOMAIN INFORMATION TECHNOLOGY CO., LTD.) 06 March 2018 (2018-03-06) entire document	1-20
A	CN 106886741 A (YUTOU TECHNOLOGY (HANGZHOU) CO., LTD.) 23 June 2017 (2017-06-23) entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 November 2019	**28 November 2019**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/104231

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109410276	A	01 March 2019	None			
CN	108399367	A	14 August 2018	None			
WO	2012095756	A2	19 July 2012	JP	2014501011	A	16 January 2014
				EP	2605844	A2	26 June 2013
				CA	2804902	A1	19 July 2012
				US	2011129124	A1	02 June 2011
				IL	224026	A	30 June 2016
CN	107767419	A	06 March 2018	None			
CN	106886741	A	23 June 2017	None			

国际检索报告

国际申请号
PCT/CN2019/104231

A. 主题的分类

G06T 7/70(2017.01)i; G06T 17/00(2006.01)i; G06N 3/02(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06T; G06K; G06N; G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNAPT; CNKI; EPODOC; WPI; IEEE: 手，手势，识别，骨骼，关节，关键点，几何，大小，长度，尺寸，模型，3D，三维，神经网络，参数，hand, gesture, skeleton, bone, geometrical, length, relationship, model, neural network, parameter, dimension, angle

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 109410276 A (北京达佳互联信息技术有限公司) 2019年 3月 1日 (2019-03-01) 权利要求第1-10项	1-20
Y	CN 108399367 A (深圳市阿西莫夫科技有限公司) 2018年 8月 14日 (2018-08-14) 说明书第[0050]-[0092]、[0104]-[0115]、[0120]、[0133]-[0134]段	1-20
Y	WO 2012095756 A2 (EXTREME REALITY LTD.) 2012年 7月 19日 (2012-07-19) 说明书第[0041]-[0050]段，附图3-5C	1-20
A	CN 107767419 A (广州深域信息科技有限公司) 2018年 3月 6日 (2018-03-06) 全文	1-20
A	CN 106886741 A (芋头科技杭州有限公司) 2017年 6月 23日 (2017-06-23) 全文	1-20

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

"A" 认为不特别相关的表示了现有技术一般状态的文件

"E" 在国际申请日的当天或之后公布的在先申请或专利

"L" 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

"O" 涉及口头公开、使用、展览或其他方式公开的文件

"P" 公布日先于国际申请日但迟于所要求的优先权日的文件

"T" 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

"X" 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

"Y" 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

"&" 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2019年 11月 7日	2019年 11月 28日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 向琳 电话号码 86-10-53961566

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/104231

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	109410276	A	2019年 3月 1日	无			
CN	108399367	A	2018年 8月 14日	无			
WO	2012095756	A2	2012年 7月 19日	JP	2014501011	A	2014年 1月 16日
				EP	2605844	A2	2013年 6月 26日
				CA	2804902	A1	2012年 7月 19日
				US	2011129124	A1	2011年 6月 2日
				IL	224026	A	2016年 6月 30日
CN	107767419	A	2018年 3月 6日	无			
CN	106886741	A	2017年 6月 23日	无			