(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
2015 年 11 月 12 日 (12.11.2015)

WIPO | PCT

(10) 国际公布号
WO 2015/169202 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01)

(21) 国际申请号: PCT/CN2015/078254

(22) 国际申请日: 2015 年 5 月 5 日 (05.05.2015)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201410187592.2 2014 年 5 月 6 日 (06.05.2014) CN

(71) 申请人: 夏普株式会社 (SHARP KABUSHIKI KAISHA) [JP/JP]; 日本大阪府大阪市阿倍野区长池町 22 番 22 号, Osaka 545-8522 (JP)。

(72) 发明人; 及

(71) 申请人 (仅对美国): 张庆久 (ZHANG, Qingjiu) [CN/CN]; 中国上海市浦东新区张江高科技园区集成电路产业区张东路 1387 号 2 栋, Shanghai 201203 (CN)。

(72) 发明人: 乐宁 (LE, Ning); 中国上海市浦东新区张江高科技园区集成电路产业区张东路 1387 号 2 栋, Shanghai 201203 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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)。

[见续页]

(54) **Title**: METHOD AND DEVICE FOR RECOGNIZING CHARACTER IN LIVE-ACTION IMAGE

(54) **发明名称** : 对实景图像中的字符进行识别的方法和设备

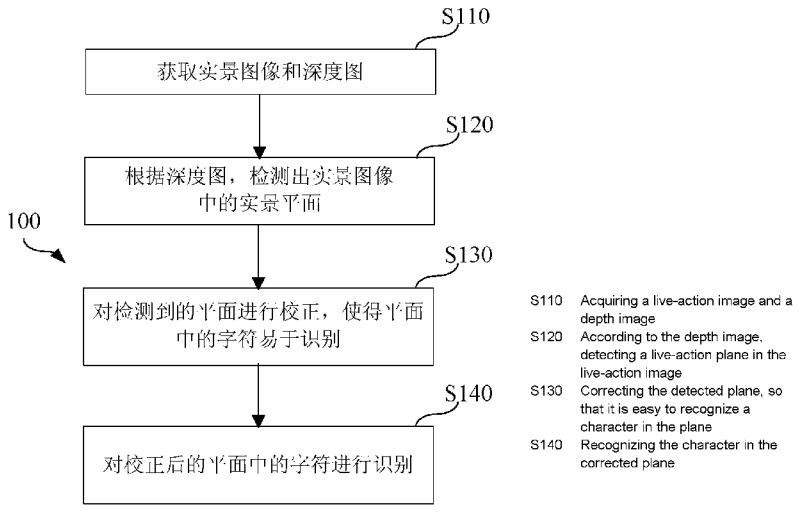


图 1 / FIG.1

(57) **Abstract**: Disclosed are a method for recognizing a character in a live-action image and a corresponding device. The method comprises: acquiring a live-action image and a depth image, each pixel in the depth image representing the distance from a live-action point corresponding to each pixel in the live-action image to a camera; according to the depth image, detecting a live-action plane in the live-action image; correcting the detected plane, so that it is easy to recognize a character in the plane; and recognizing the character in the corrected plane. By means of the disclosed method and device, the accuracy of recognizing a character in a live-action image can be improved.

(57) **摘要**: 公开了一种对实景图像中的字符进行识别的方法以及相应的设备。所述方法包括：获取实景图像和深度图，所述深度图中的每个像素表示与实景图像中的像素相对应的实景点到摄像机的距离；根据深度图，检测出实景图像中的实景平面；对检测到的平面进行校正，使得平面中的字符易于识别；以及对校正后的平面中的字符进行识别。采用所公开的方法和设备可以提高对实景图像中的字符进行识别的准确度。

WO 2015/169202 A1

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

对实景图像中的字符进行识别的方法和设备

技术领域

本申请总体涉及光学字符识别 OCR 技术，具体涉及对实景图像中的字符进行识别的方法和设备。

背景技术

除了识别算法本身的性能以外，传统 OCR 技术对图像中字符的识别准确度还取决于待处理图像的复杂度、文字所在平面与图像呈现平面的平行度、以及图像中文字的平正程度。为了使识别准确度达到可接受的程度，通常要求待处理图像中仅包含一页文字，该页文字与图像呈现平面基本平行，且图像中的文字基本平正（即，非歪斜）摆放。

在实际应用中，用户可能希望对日常拍摄的实景图像中的文字（例如街拍场景中某一透视投影的标牌上的文字）进行光学识别，并可能希望对其进行文字处理。然而，由于日常拍摄的实景图像的构图通常较为复杂、实体对象较多、实景平面与图像呈现平面夹角较大（由于透视投影导致）、文字倾斜度较大，识别准确度难以达到令人满意的程度。

题为“Generating Hi-Res Dewarped Book Images”的美国专利申请 US 20130242054 A1 提出了一种对书本立体图像去除扭曲的方法。该方法利用三维几何表面（例如圆柱表面）产生 3D 信息，而无需使用文本行或其他内容分析。具体地，在捕捉到立体图像后，首先根据立体照片产生三维几何平面。然后，计算相应的模型。接着，由模型对立体图像去除扭曲。

虽然上述方案能够提高对书页中文字的识别准确度，还是难于处理具有复杂场景的实体图像。

发明内容

鉴于现有技术的上述问题和缺陷，本发明的目的在于提出一种能够以较高的识别准确度对实景图像中的字符进行识别的方案。

根据本发明的第一方面，提供了一种对实景图像中的字符进行识别的方法，包括：获取实景图像和深度图，所述深度图中的每个像素表示与实景图像中的像素相对应的实景点到摄像机的距离；根据深度图，检测出实景图像中的实景平面；对检测到的平面进行校正，使得平面中的字符易于识别；以及对校正后的平面中的字符进行识别。

根据深度图检测出实景图像中的实景平面可以包括：基于深度图中相邻像素的深度差，检测实景图像中的连通体；利用检测出的连通体各像素的坐标及对应的深度值，进行平面拟合；滤除连通体的位于拟合平面以外的像素；以及确定包围连通体的剩余像素的凸多边形。如果相邻像素的深度差小于阈值，则可以将相邻像素确定为属于同一连通体。平面拟合可以采用最小二乘法。

所述方法还可以包括：基于检测出的连通体的面积、平均深度和/或在实景图像中的相对位置，滤除噪声连通体。

对检测到的平面进行校正可以包括：在三维空间中调整检测到的平面的倾斜姿态，使其处于与显示平面平行的状态；以及在显示平面所对应的二维空间中旋转经调整的平面，使其中的文字平正显示。

在三维空间中调整检测到的平面的倾斜姿态可以包括：基于平面拟合得到的平面参数，计算出使检测到的平面旋转为与显示平面平行的平面的透视变换矩阵；对实景图像的与检测到的平面相对应的部分的像素应用所述透视变换矩阵。

所述方法还可以包括：对检测到的平面进行校正前，使用者确认或选择实景图像中检测出的实景平面。

所述方法还可以包括：对校正后的平面中的字符进行识别前，使用者确认或选择校正后的平面。

根据本发明的第二方面，提供了一种对实景图像中的字符进行识别的设备，包括：图像获取装置，用于获取实景图像和深度图，所述深度图中的每个像素表示与实景图像中的像素相对应的实景点到摄像机的距离；平面检测装置，用于根据深度图，检测出实景图像中的实景平面；平面校正装置，用于对检测到的平面进行校正，使得平面中的字符易于识别；以及字符识别装置，用于对校正后的平面中的字符进行识别。

所述平面检测装置可以包括：连通体检测单元，被配置为基于深度

图中相邻像素的深度差，检测实景图像中的连通体；平面拟合单元，被配置为利用检测出的连通体各像素的坐标及对应的深度值，进行平面拟合；噪声过滤单元，被配置为滤除连通体的位于拟合平面以外的像素；以及确定单元，被配置为确定包围连通体的剩余像素的凸多边形。如果相邻像素的深度差小于阈值，则可以将相邻像素确定为属于同一连通体。平面拟合可以采用最小二乘法。

所述噪声过滤单元还可以被配置为：基于检测出的连通体的面积、平均深度和/或在实景图像中的相对位置，滤除噪声连通体。

所述平面校正装置可以包括：三维倾斜校正单元，被配置为在三维空间中调整检测到的平面的倾斜姿态，使其处于与显示平面平行的状态；以及二维歪斜校正单元，被配置为在显示平面所对应的二维空间中旋转经调整的平面，使其中的文字平正显示。

所述三维倾斜校正单元可以包括：透视变换矩阵计算子单元，被配置为基于平面拟合得到的平面参数，计算出使检测到的平面旋转为与显示平面平行的平面的透视变换矩阵；透视变换矩阵应用子单元，被配置为对实景图像中与检测到的平面相对应的部分的像素应用所述透视变换矩阵。

所述的设备还可以包括：实景平面确认/选择装置，用于使用者确认或选择实景图像中检测出的实景平面。

所述设备还可以包括：校正平面确认/选择装置，用于使用者确认或选择校正后的平面。

所述实景图像可以是彩色图像或灰度图像。

附图说明

通过下面结合附图说明本发明的优选实施例，将使本发明的上述及其它目的、特征和优点更加清楚，其中：

图 1 是示出了根据本发明的对实景图像中的字符进行识别的方法的流程图；

图 2 和 3 分别示出了所获得的实景图像和深度图的示例；

图 4 是示出了根据本发明的平面检测操作的示例实现的流程图；

图 5 示出了由图 3 所示的深度图获得的连通体的示意表示；

图 6 示出了移除噪声连通体后的示意效果；

图 7 示出了检测出的平面区域的示意表示；

图 8 是示出了根据本发明的屏幕校正操作的示例实现的流程图；

图 9 示出了三维倾斜校正的示意效果；

图 10 和 11 分别给出了对实景图像的与检测到的平面相对应的部分应用三维倾斜校正前后的示意图；

图 12 和 13 分别给出了对三维倾斜校正后的图像进行二值化处理和旋转处理后的示意图；

图 14～16 示出了根据本发明的对实景图像中的字符进行识别的设备的示例结构的框图；以及

图 17 示出了使用智能电话拍摄实景图像、对实景图像应用根据本发明的识别方法、并将识别出的中文文本翻译成英文文本的示例过程。

具体实施方式

下面参照附图对本发明的优选实施例进行详细说明，在描述过程中省略了对于本发明来说是不必要的细节和功能，以防止对本发明的理解造成混淆。

首先，参照图 1，描述根据本发明的对实景图像中的字符进行识别的方法 100。如图 1 所示，方法 100 起始于步骤 s110，获取实景图像和深度图，其中，深度图中的每个像素表示与实景图像中的像素相对应的实景点到摄像机的距离。接着，执行步骤 s120，根据深度图，检测出实景图像中的实景平面。然后，执行步骤 s130，对检测到的平面进行校正，使得平面中的字符易于识别。最后，执行步骤 s140，对校正后的平面中的字符进行识别。

优选地，在执行步骤 s130 前，使用者确认或选择实景图像中检测出的实景平面。使用者可以确认全部或部分实景图像中检测出的实景平面。或者使用者选择全部或部分实景图像中检测出的实景平面。

优选地，在执行步骤 s140 前，使用者确认或选择校正后的平面。使用者可以确认全部或部分校正后的平面。或者使用者选择全部或部分校正后的平面。

通过在对实景图像中的字符进行识别之前先借助深度图检测出实景

图像中的实景平面并对其校正，可以极大地提高对实景图像中的字符进行识别的准确度。

在具体实现中，深度图可由 Kinect 等设备直接捕捉，也可由立体系统通过对具有固定的相对空间位置的多个摄像机所拍摄的多个图像进行计算来生成。在图 2 和 3 中，分别给出了实景图像和深度图的示意表示。

图 4 进一步给出了步骤 s120 的一种示例实现。如图所示，过程起始于步骤 s1201，基于深度图中相邻像素的深度差，检测实景图像中的连通体，作为候选平面。在具体实现中，可以对整个深度图进行扫描，将深度差小于阈值的相邻像素确定为属于同一连通体，将深度差大于阈值的相邻像素确定为不同连通体的边界像素。在图 5 中，给出了由图 3 所示的深度图获得的连通体的示意表示。

优选地，可以基于连通体的多项特征（如基于检测出的连通体的面积、平均深度和/或在实景图像中的相对位置），对检测出的连通体进行过滤，移除不包含有字平面的噪声连通体，进而避免非必要的后续计算。在图 6 中，给出了移除噪声连通体后的示意效果。事实上，移除噪声连通体不仅能够节省大量的计算开销，还能够如图所示地移除图像中存在的可能导致较低文字识别率的大量噪声。

由图 6 可见，即便对噪声连通体进行了滤除，剩余连通体仍具有与平面无关的噪声点（图 6 中的噪声点对应于与纸张平面相连的人手的一部分）。为消除这些噪点，可以执行步骤 s1202，利用检测出的连通体各像素的坐标及对应的深度值，进行平面拟合，计算出平面的表征参数。为此，可以采用现有的各种平面拟合方法，包括但不限于最小二乘法。此后，执行步骤 s1203,将连通体的位于拟合平面以外的像素确定为噪声点加以滤除。最后，执行步骤 s1204，确定包围连通体的剩余像素的凸多边形，形成规则的平面区域。在图 7 中，给出了对图 6 所示的连通体执行上述操作后得到的平面区域的示意表示。

此后，可以进一步对检测到的平面进行校正，使得平面中的字符易于识别。图 8 给出了此操作的一种示例实现。如图所示，该操作可以细分为三维倾斜校正 s1301 以及二维歪斜校正 s1302 两个步骤。

三维倾斜校正 s1301 负责在三维空间中调整检测到的平面的倾斜姿态，使其处于与显示平面平行的状态。图 9 示出了三维倾斜校正的效果

示意。如图所示，通过校正，检测到的平面在三维空间中被旋转为与 XOY（Z=0）平行的目标平面，其中，(X，Y)表示像素在显示平面中的坐标，Z 表示其深度。

在实际实现中，可以选择检测到的平面中的一点作为参考点，以此确定目标平面的 Z 值。基于平面拟合得到的平面参数，计算出使检测到的平面旋转为与显示平面平行的平面的透视变换矩阵。然后，对实景图像的与检测到的平面相对应的部分的像素应用所述透视变换矩阵。图 10 和 11 分别给出了对实景图像的与检测到的平面相对应的部分应用三维倾斜校正前后的示意图。

三维倾斜校正后的图像中的文本可能是歪斜的，直接对其进行识别仍无法获得令人满意的识别度。为对此进行改善，在识别前应用二维歪斜校正 s1302，以在显示平面所对应的二维空间中旋转经三维倾斜校正的平面，使其中的文字平正显示。

在实际实现中，可以首先对三维倾斜校正后的图像进行二值化，并利用二值化图像的特征，计算使其中的文字平正显示所需的旋转角。然后，将三维倾斜校正后的图像旋转相应的角度。图 12 和 13 分别给出了对三维倾斜校正后的图像进行二值化处理和旋转处理后的示意图。由图 13 可见，经过平面检测和平面校正处理后得到的校正图像与将纸张直接放入扫描仪得到的扫描图像基本相近，对该图像进行文字识别能够确保令人满意的识别准确度。

与上述对实景图像中的字符进行识别的方法相对应地，本发明还提出了相关的设备。图 14～16 示出了这样的设备的示意结构方框图。

如图 14 所示，根据本发明的对实景图像中的字符进行识别的设备 1400 包括图像获取装置 1410、平面检测装置 1420、平面校正装置 1430 以及字符识别装置 1440。图像获取装置 1410 用于获取实景图像和深度图，所述深度图中的每个像素表示与实景图像中的像素相对应的实景点到摄像机的距离。平面检测装置 1420 用于根据深度图，检测出实景图像中的实景平面。平面校正装置 1430 用于对检测到的平面进行校正，使得平面中的字符易于识别。字符识别装置 1440 用于对校正后的平面中的字符进行识别。

优选地，在平面检测装置 1420 和平面校正装置 1430 之间，可以包

括实景平面确认/选择装置，用于使用者确认或选择实景图像中检测出的实景平面。所述实景平面确认装置可以用于使用者确认全部或部分实景图像中检测出的实景平面。或者所述实景平面选择装置可以用于使用者选择全部或部分实景图像中检测出的实景平面。

优选地，在平面校正装置 1430 和字符识别装置 1440 之间，可以包括校正平面确认/选择装置，用于使用者确认或选择校正后的平面。所述校正平面确认装置可以用于使用者确认全部或部分校正后的平面。或者校正平面选择装置可以用于使用者选择全部或部分校正后的平面。

如图 15 所示，所述平面检测装置 1420 可以包括连通体检测单元 1421、平面拟合单元 1422、噪声过滤单元 1423 和确定单元 1424。连通体检测单元 1421 被配置为基于深度图中相邻像素的深度差，检测实景图像中的连通体。平面拟合单元 1422 被配置为利用检测出的连通体各像素的坐标及对应的深度值，进行平面拟合。噪声过滤单元 1423 被配置为滤除连通体的位于拟合平面以外的像素。确定单元 1424 被配置为确定包围连通体的剩余像素的凸多边形。如果相邻像素的深度差小于阈值，则可以将相邻像素确定为属于同一连通体。平面拟合可以采用最小二乘法。所述噪声过滤单元还可以被配置为：基于检测出的连通体的面积、平均深度和/或在实景图像中的相对位置，滤除噪声连通体。

如图 16 所示，所述平面校正装置 1430 可以包括：三维倾斜校正单元 1431，被配置为在三维空间中调整检测到的平面的倾斜姿态，使其处于与显示平面平行的状态；以及二维歪斜校正单元 1432，被配置为在显示平面所对应的二维空间中旋转经调整的平面，使其中的文字平正显示。

所述三维校正单元 1431 可以包括：透视变换矩阵计算子单元，被配置为基于平面拟合得到的平面参数，计算出使检测到的平面旋转为与显示平面平行的平面的透视变换矩阵；以及透视变换矩阵应用子单元，被配置为对实景图像中与检测到的平面相对应的部分的像素应用所述透视变换矩阵。

根据本发明的对实景图像中的字符进行识别的方法和设备可以有利地应用于自身配备了两个以上的摄像机因而能够重建 3D 场景的手持设备（如智能电话、平板计算机等）。购买了该手持设备的用户在国外旅游期间可以方便地用手持设备拍摄实景图像，以较高的识别度对实景图

像中的外国文字（例如，街拍场景中透视投影的路标、布告栏、广告牌等上的文字）进行识别，并将其翻译成本国语。图 17 示出了使用智能电话拍摄实景图像、对实景图像应用根据本发明的识别方法、并将识别出的中文文本翻译成英文文本的过程。在图 17 中，第一步，使用智能电话拍摄获取实景图像和深度图；第二步，根据深度图，检测出实景图像中的实景平面，此处，以“码头”为例，检测出带有“码头”字符的实景平面；第三步和第四步，对检测到的平面进行校正，其中，第三步为透视变换，第四步为倾斜校正；第五步，对校正后的平面中的字符进行识别，得到“码头”的中文文本；第六步，将识别出的中文文本“码头”翻译成英文文本“Pier”。

应当注意的是，在以上的描述中，仅以示例的方式，示出了本发明的技术方案，但并不意味着本发明局限于上述步骤和单元结构。在可能的情形下，可以根据需要对步骤和单元结构进行调整和取舍。因此，某些步骤和单元并非实施本发明的总体发明思想所必需的元素。因此，本发明所必需的技术特征仅受限于能够实现本发明的总体发明思想的最低要求，而不受以上具体实例的限制。

至此已经结合优选实施例对本发明进行了描述。应该理解，本领域技术人员在不脱离本发明的精神和范围的情况下，可以进行各种其它的改变、替换和添加。因此，本发明的范围不局限于上述特定实施例，而应由所附权利要求所限定。

权 利 要 求

1. 一种对实景图像中的字符进行识别的方法，包括：

获取实景图像和深度图，所述深度图中的每个像素表示与实景图像中的像素相对应的实景点到摄像机的距离；

根据深度图，检测出实景图像中的实景平面；

对检测到的平面进行校正，使得平面中的字符易于识别；以及

对校正后的平面中的字符进行识别。

2. 根据权利要求 1 所述的方法，其中，根据深度图检测出实景图像中的实景平面包括：

基于深度图中相邻像素的深度差，检测实景图像中的连通体；

利用检测出的连通体各像素的坐标及对应的深度值，进行平面拟合；

滤除连通体的位于拟合平面以外的像素；以及

确定包围连通体的剩余像素的凸多边形。

3. 根据权利要求 2 所述的方法，其中，如果相邻像素的深度差小于阈值，则将相邻像素确定为属于同一连通体。

4. 根据权利要求 2 所述的方法，其中，平面拟合采用最小二乘法。

5. 根据权利要求 2 所述的方法，还包括：基于检测出的连通体的面积、平均深度和/或在实景图像中的相对位置，滤除噪声连通体。

6. 根据权利要求 1 或 2 所述的方法，其中，对检测到的平面进行校正包括：

在三维空间中调整检测到的平面的倾斜姿态，使其处于与显示平面平行的状态；以及

在显示平面所对应的二维空间中旋转经调整的平面，使其中的文字平正显示。

7. 根据权利要求 6 所述的方法，其中，在三维空间中调整检测到的平面的倾斜姿态包括：

基于平面拟合得到的平面参数，计算出使检测到的平面旋转为与显示平面平行的平面的透视变换矩阵；

对实景图像的与检测到的平面相对应的部分的像素应用所述透视变换矩阵。

8. 根据权利要求 1 所述的方法，还包括：对检测到的平面进行校正前，使用者确认或选择实景图像中检测出的实景平面。

9. 根据权利要求 1 所述的方法，还包括：对校正后的平面中的字符进行识别前，使用者确认或选择校正后的平面。

10. 根据权利要求 1 或 2 所述的方法，其中，所述实景图像是彩色图像或灰度图像。

11. 一种对实景图像中的字符进行识别的设备，包括：

图像获取装置，用于获取实景图像和深度图，所述深度图中的每个像素表示与实景图像中的像素相对应的实景点到摄像机的距离；

平面检测装置，用于根据深度图，检测出实景图像中的实景平面；

平面校正装置，用于对检测到的平面进行校正，使得平面中的字符易于识别；以及

字符识别装置，用于对校正后的平面中的字符进行识别。

12. 根据权利要求 11 所述的设备，其中，所述平面检测装置包括：

连通体检测单元，被配置为基于深度图中相邻像素的深度差，检测实景图像中的连通体；

平面拟合单元，被配置为利用检测出的连通体各像素的坐标及对应的深度值，进行平面拟合；

噪声过滤单元，被配置为滤除连通体的位于拟合平面以外的像素；以及

确定单元，被配置为确定包围连通体的剩余像素的凸多边形。

13. 根据权利要求 12 所述的设备，其中，如果相邻像素的深度差小于阈值，则将相邻像素确定为属于同一连通体。

14. 根据权利要求 12 所述的设备，其中，平面拟合采用最小二乘法。

15. 根据权利要求 12 所述的设备，其中，所述噪声过滤单元还被配置为：基于检测出的连通体的面积、平均深度和/或在实景图像中的相对位置，滤除噪声连通体。

16. 根据权利要求 11 或 12 所述的设备，其中，所述平面校正装置包括：

三维倾斜校正单元，被配置为在三维空间中调整检测到的平面的倾斜姿态，使其处于与显示平面平行的状态；以及

二维歪斜校正单元，被配置为在显示平面所对应的二维空间中旋转经调整的平面，使其中的文字平正显示。

17. 根据权利要求 16 所述的设备，其中，所述三维倾斜校正单元包括：

透视变换矩阵计算子单元，被配置为基于平面拟合得到的平面参数，计算出使检测到的平面旋转为与显示平面平行的平面的透视变换矩阵；

透视变换矩阵应用子单元，被配置为对实景图像中与检测到的平面相对应的部分的像素应用所述透视变换矩阵。

18. 根据权利要求 11 所述的设备，还包括：实景平面确认/选择装置，用于使用者确认或选择实景图像中检测出的实景平面。

19. 根据权利要求 11 所述的设备，还包括：校正平面确认/选择装置，用于使用者确认或选择校正后的平面。

20. 根据权利要求 11 或 12 所述的设备，其中，所述实景图像是彩色图像或灰度图像。

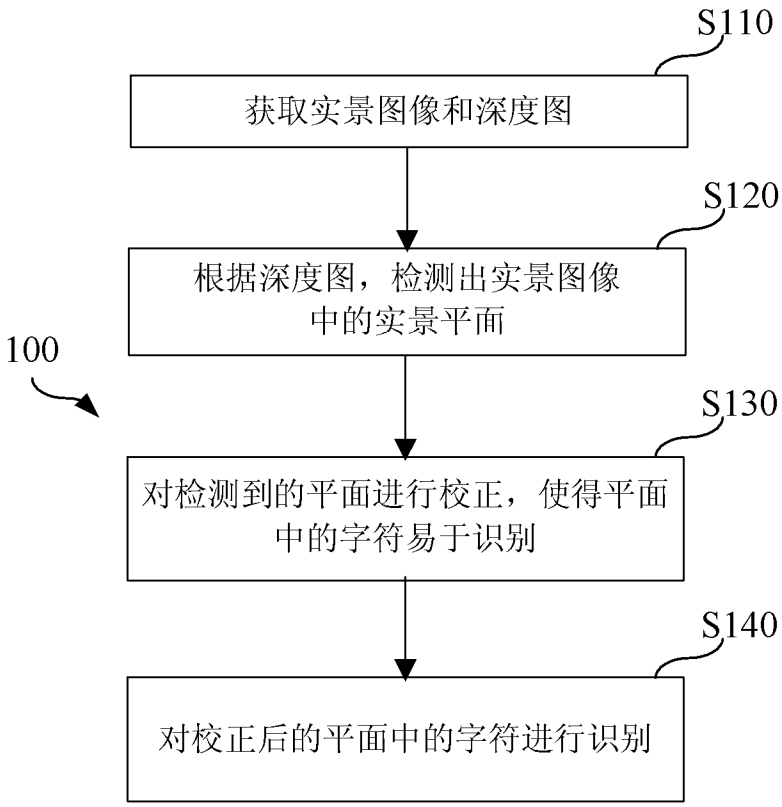

S110
获取实景图像和深度图
S120
根据深度图，检测出实景图像中的实景平面
100
S130
对检测到的平面进行校正，使得平面中的字符易于识别
S140
对校正后的平面中的字符进行识别

图 1

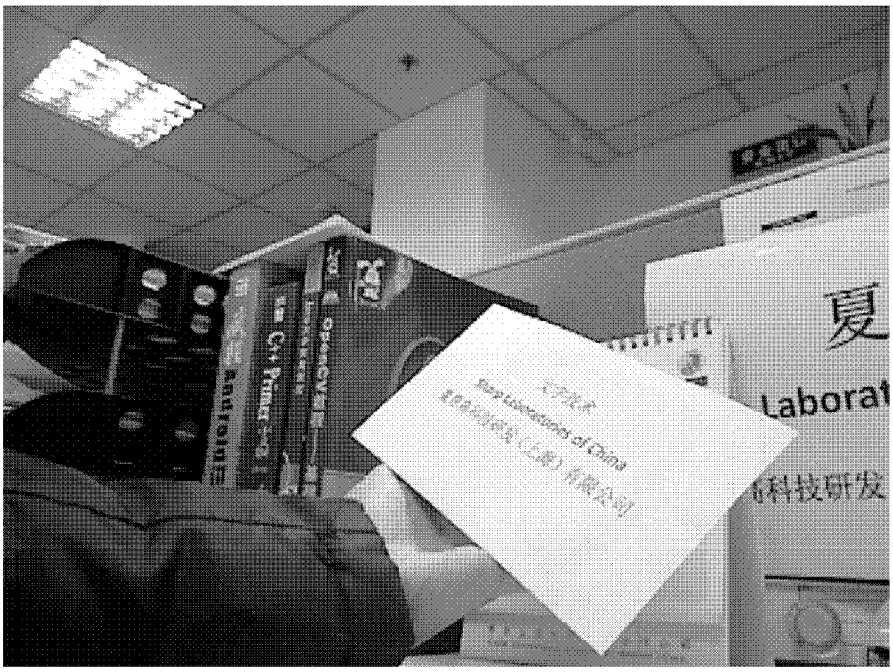


图 2

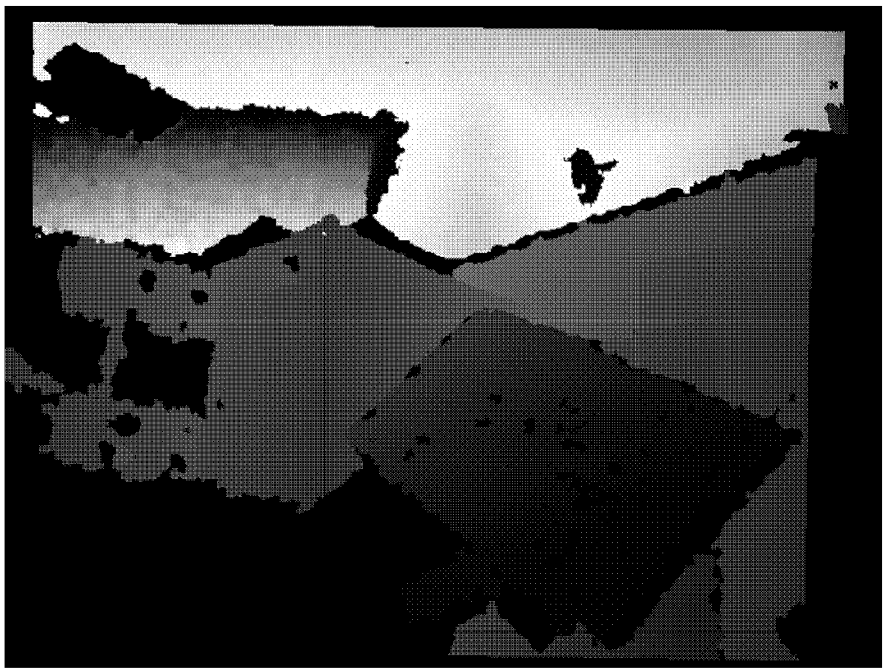


图 3

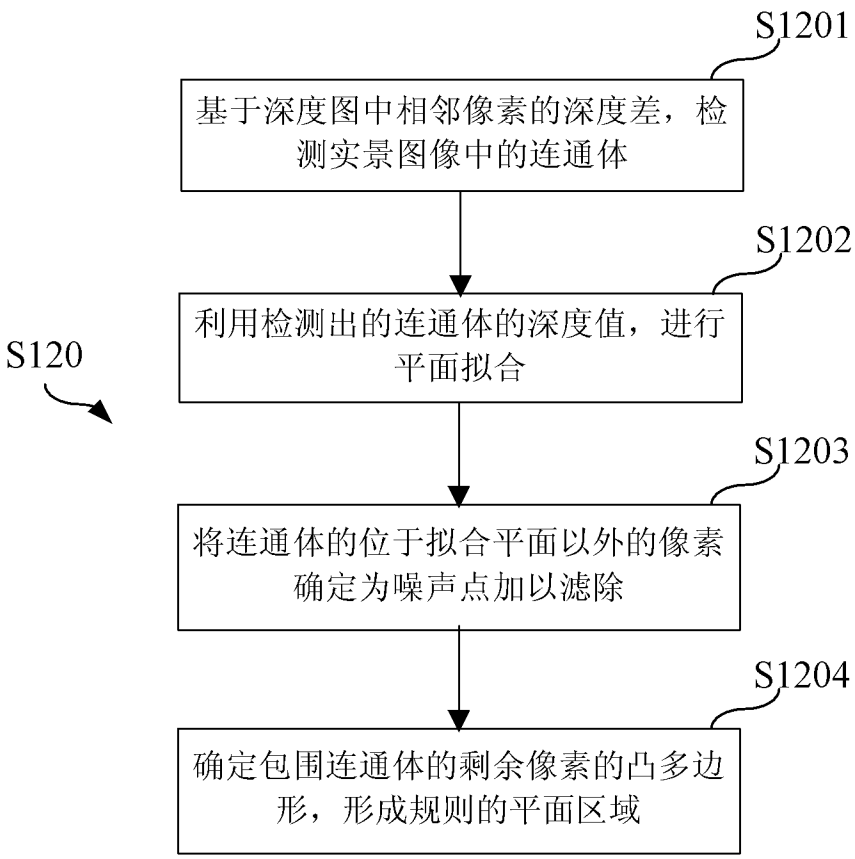

S1201
基于深度图中相邻像素的深度差，检测实景图像中的连通体
S1202
利用检测出的连通体的深度值，进行平面拟合
S120
S1203
将连通体的位于拟合平面以外的像素确定为噪声点加以滤除
S1204
确定包围连通体的剩余像素的凸多边形，形成规则的平面区域

图 4

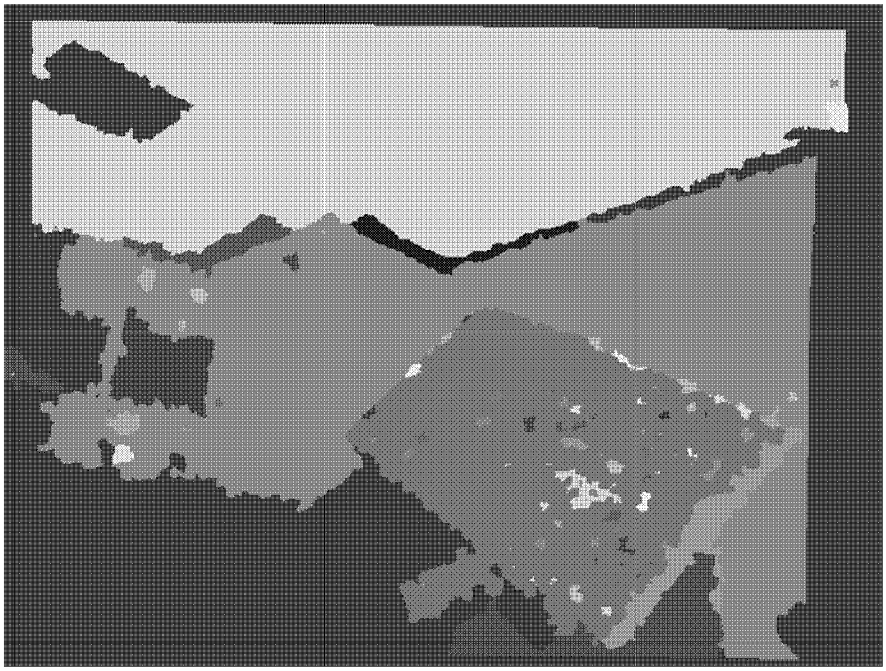


图 5

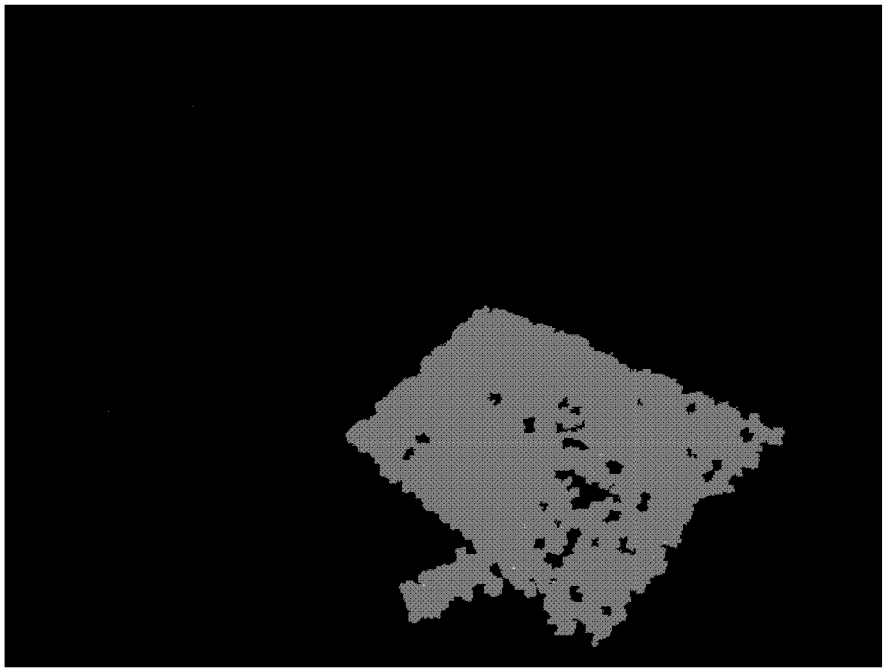


图 6

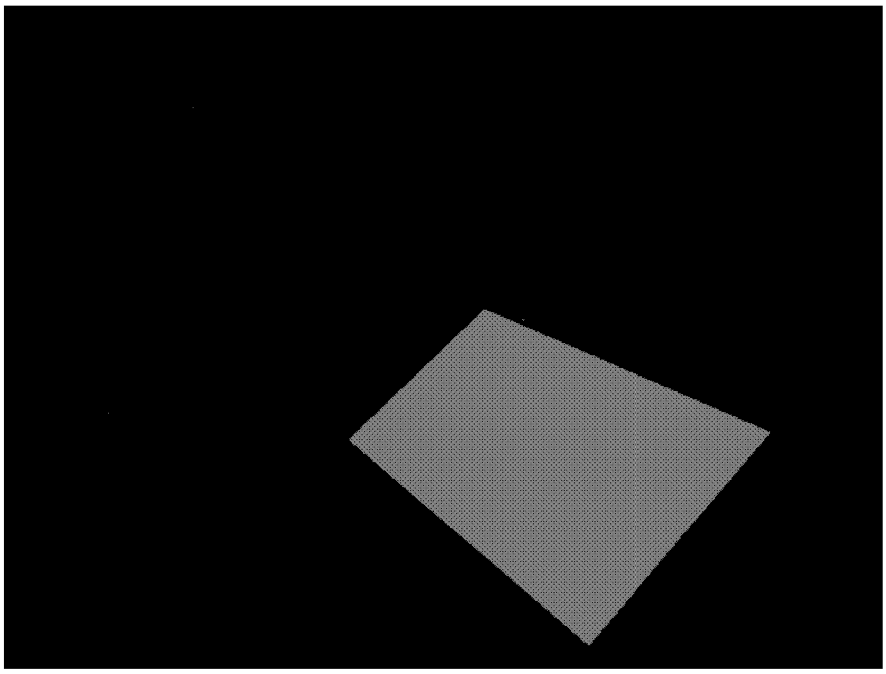


图 7

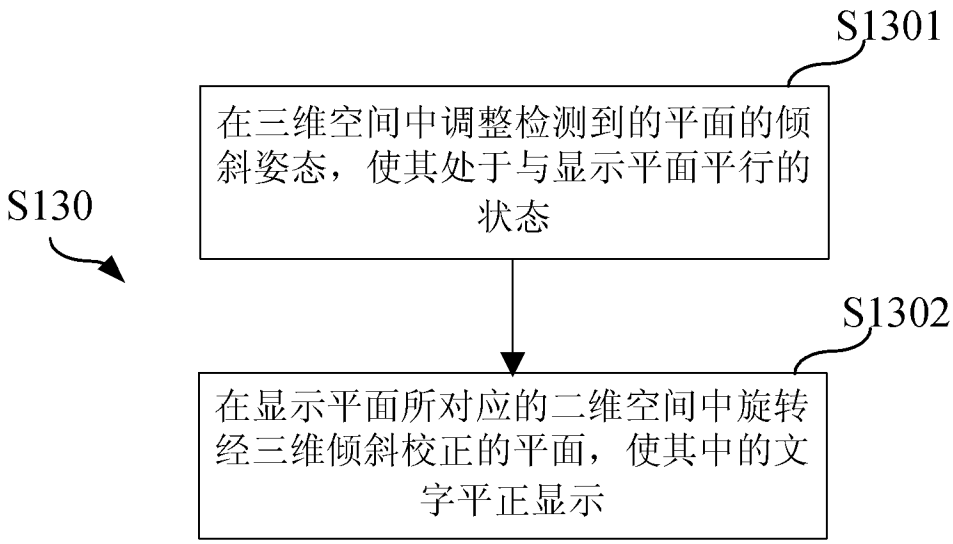

S1301
在三维空间中调整检测到的平面的倾斜姿态，使其处于与显示平面平行的状态
S130
S1302
在显示平面所对应的二维空间中旋转经三维倾斜校正的平面，使其中的文字平正显示

图 8

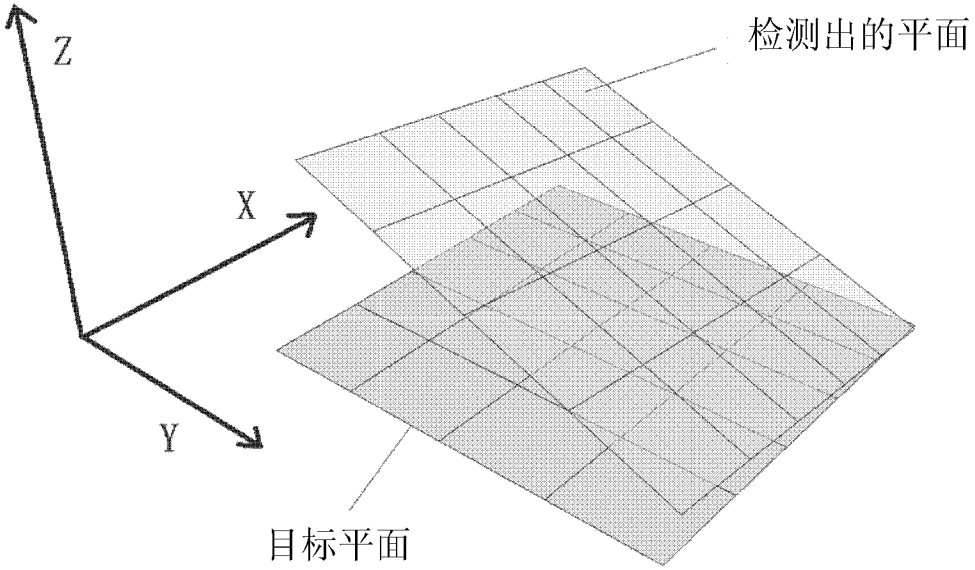


图 9

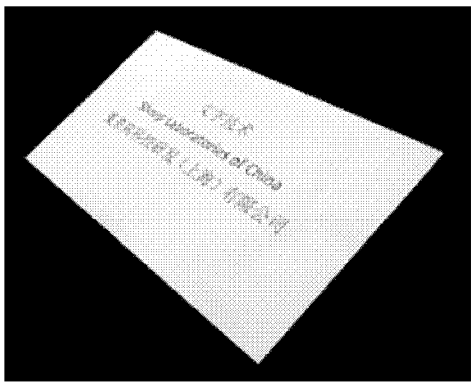


图 10



图 11



图 12



图 13

设备1400

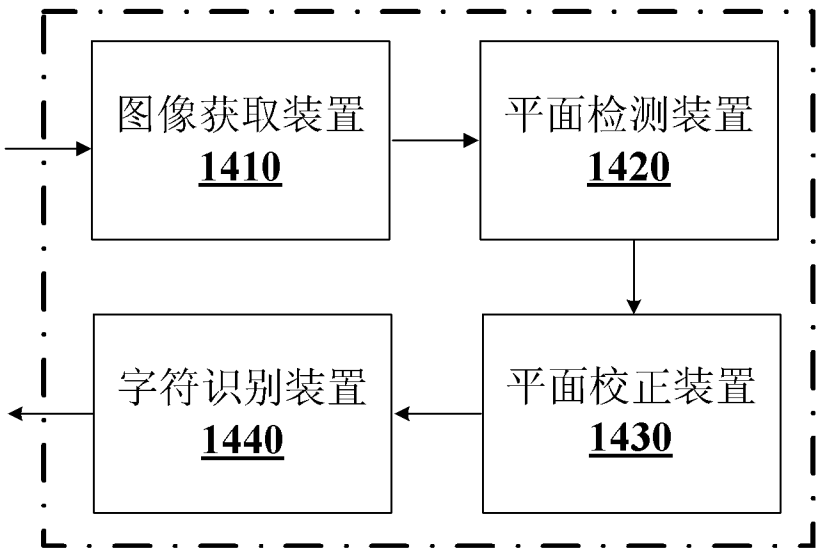


图 14

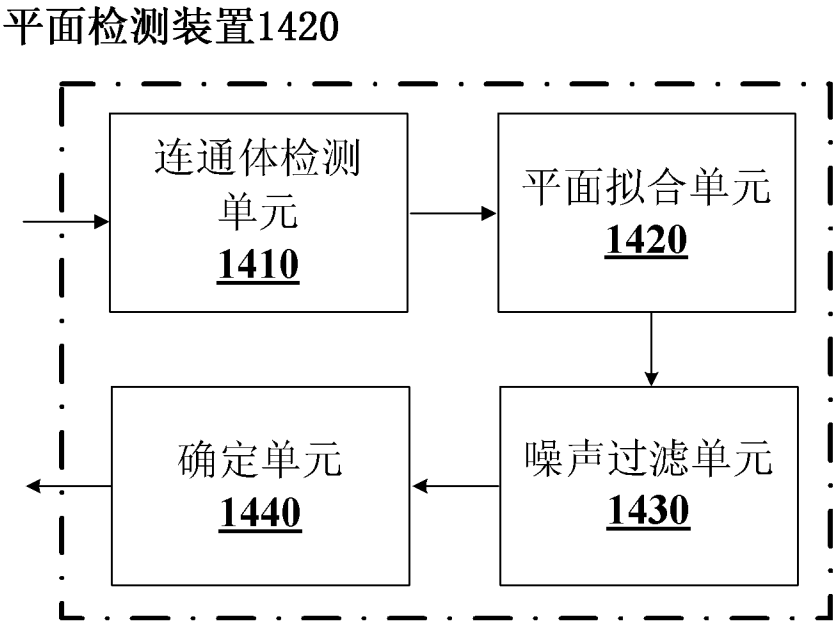

平面检测装置1420
连通体检测单元 1410
平面拟合单元 1420
确定单元 1440
噪声过滤单元 1430

图 15

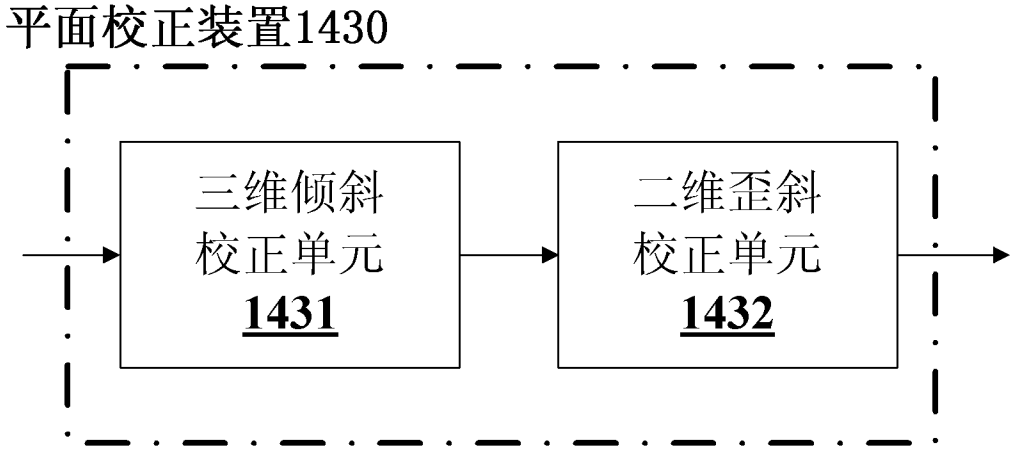

平面校正装置1430
三维倾斜校正单元 1431
二维歪斜校正单元 1432

图 16

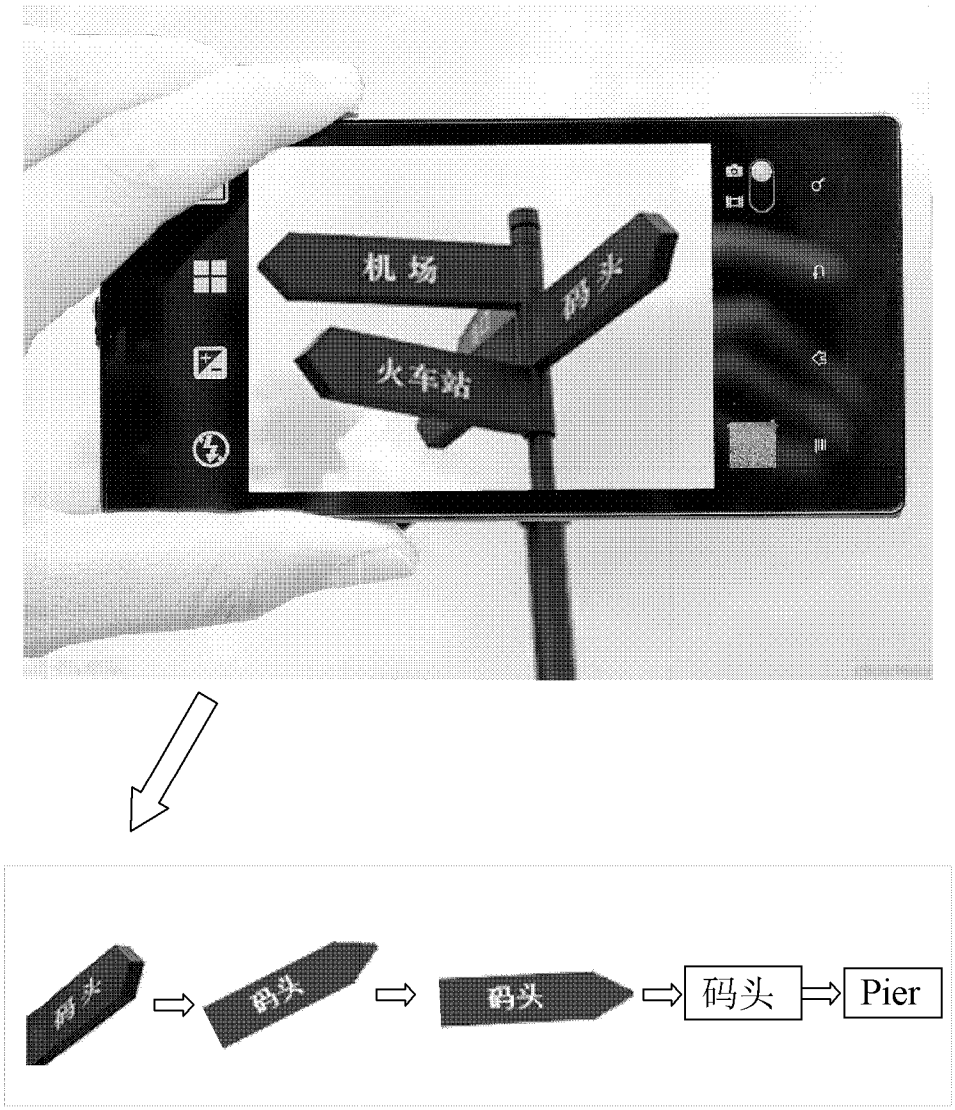


图 17

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/078254

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00 (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)

G06K, G06T

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)

CNPAT, CNKI: SHARP; ZHANG, Qingjiu; YUE, Ning; image, character, recognition, depth map, segmentation, correction, inclination, connected component, connected domain, plane

WPI, EPODOC: sharp, image, character, recogniz+, OCR, recognition, depth, range, segment+, inclination, correct+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 103065149 A (SHANGHAI JIAO TONG UNIVERSITY), 24 April 2013 (24.04.2013), description, page 3, paragraph [0039] to page 4, paragraph [0055], and figure 1	1-20
Y	CN 102609934 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES), 25 July 2012 (25.07.2012), description, page 2, paragraph [0026] to page 5, paragraph [0077], and figure 4	1-20
Y	CN 101887521 A (ZTE CORP.), 17 November 2010 (17.11.2010), description, page 3, paragraphs [0033]-[0040]	1-20
A	CN 101625762 A (SHENZHEN WISION TECHNOLOGY HOLDINGS CO., LTD.), 13 January 2010 (13.01.2010), the whole document	1-20
A	JP 2003051005 A (FUJI PHOTO FILM CO., LTD.), 21 February 2003 (21.02.2003), the whole 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

"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
25 June 2015 (25.06.2015)	**05 August 2015 (05.08.2015)**
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer **WANG, Conglei** Telephone No.: (86-10) **62413236**

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2015/078254

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103065149 A	24 April 2013	None	
CN 102609934 A	25 July 2012	WO 2013091369 A1	27 June 2013
CN 101887521 A	17 November 2010	None	
CN 101625762 A	13 January 2010	None	
JP 2003051005 A	21 February 2003	None	

国际检索报告

国际申请号
PCT/CN2015/078254

A. 主题的分类

G06K 9/00(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)

G06K，G06T

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））

CNPAT, CNKI：夏普，张庆久，乐宁，图像，字符，识别，深度图，分割，校正，倾斜，连通体，连通域，平面 WPI, EPODOC: sharp, image, character, recogniz+, OCR, recognition, depth, range, segment+, inclination, correct+

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
Y	CN 103065149 A (上海交通大学) 2013年 4月 24日 (2013-04-24) 说明书第3页第[0039]段-第4页第[0055]段、附图1	1-20
Y	CN 102609934 A (中国科学院自动化研究所) 2012年 7月 25日 (2012-07-25) 说明书第2页第[0026]段-第5页第[0077]段、附图4	1-20
Y	CN 101887521 A (中兴通讯股份有限公司) 2010年 11月 17日 (2010-11-17) 说明书第3页第[0033]-[0040]段	1-20
A	CN 101625762 A (深圳市中瀛鑫科技发展有限公司) 2010年 1月 13日 (2010-01-13) 全文	1-20
A	JP 2003051005 A (FUJI PHOTO FILM CO., LTD.) 2003年 2月 21日 (2003-02-21) 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
2015年 6月 25日	2015年 8月 5日

ISA/CN的名称和邮寄地址	受权官员
中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国	王从雷
传真号 (86-10)62019451	电话号码 (86-10)62413236

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2015/078254

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 103065149 A	2013年 4月 24日	无	
CN 102609934 A	2012年 7月 25日	WO 2013091369 A1	2013年 6月 27日
CN 101887521 A	2010年 11月 17日	无	
CN 101625762 A	2010年 1月 13日	无	
JP 2003051005 A	2003年 2月 21日	无	