



(12) 发明专利申请

(10) 申请公布号 CN 117745826 A

(43) 申请公布日 2024. 03. 22

(21) 申请号 202311775598.7

G06T 7/50 (2017.01)

(22) 申请日 2023.12.21

G06T 5/50 (2006.01)

(71) 申请人 武汉星巡智能科技有限公司

地址 430000 湖北省武汉市洪山区洪山街  
马湖村武汉创意天地三、四期第4号楼  
11层1号房

(72) 发明人 陈辉 熊章 张智 胡国湖  
张青军

(74) 专利代理机构 成都恪睿信专利代理事务所  
(普通合伙) 51303

专利代理师 陈兴强

(51) Int. Cl.

G06T 7/73 (2017.01)

G06T 5/90 (2024.01)

G06T 7/13 (2017.01)

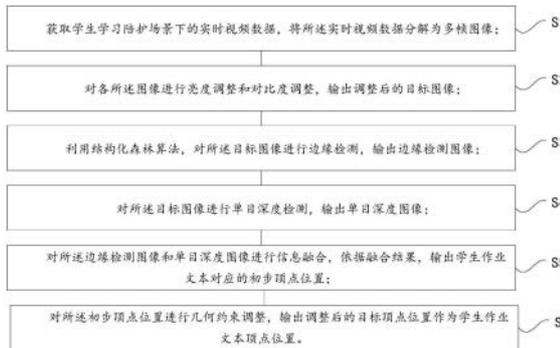
权利要求书2页 说明书13页 附图4页

(54) 发明名称

基于文本角点检测的学习陪护方法、装置、  
机器人及介质

(57) 摘要

本发明涉及学习陪护技术领域,解决了现有技术中检测作业文本四周角点位置准确度低、成本高的问题,提供了一种基于文本角点检测的学习陪护方法、装置、机器人及介质。该方法包括:获取学生学习陪护场景下的多帧图像;对各所述图像进行亮度调整和对比度调整,输出目标图像;利用结构化森林算法,对所述目标图像进行边缘检测,输出边缘检测图像;对所述目标图像进行单目深度检测,输出单目深度图像;对所述边缘检测图像和单目深度图像进行信息融合,输出初步顶点位置;对所述初步顶点位置进行几何约束调整,输出目标顶点位置作为学生作业文本顶点位置。本发明提高了对学生作业文本的检测和定位的准确性和可靠性,降低了检测成本。



1. 一种基于文本角点检测的学习陪护方法,其特征在于,所述方法包括:  
S1:获取学生学习陪护场景下的实时视频数据,将所述实时视频数据分解为多帧图像;  
S2:对各所述图像进行亮度调整和对比度调整,输出调整后的目标图像;  
S3:利用结构化森林算法,对所述目标图像进行边缘检测,输出边缘检测图像;  
S4:对所述目标图像进行单目深度检测,输出单目深度图像;  
S5:对所述边缘检测图像和单目深度图像进行信息融合,依据融合结果,输出学生作业文本对应的初步顶点位置;

S6:对所述初步顶点位置进行几何约束调整,输出调整后的目标顶点位置作为学生作业文本顶点位置。

2. 根据权利要求1所述的基于文本角点检测的学习陪护方法,其特征在于,所述S2包括:

S21:对各所述图像进行预处理,输出预处理之后灰度化的图像;

S22:对灰度化的图像中亮度分量和对比度分量分别进行调整,输出调整后的亮度分量和调整后的对比度分量;

S23:依据调整后的亮度分量和调整后的对比度分量,结合预处理之前图像的颜色信息,输出所述目标图像。

3. 根据权利要求1所述的基于文本角点检测的学习陪护方法,其特征在于,所述S3包括:

S31:获取多帧学生学习陪护场景下的训练图像和对应的边缘图像;

S32:依据所述训练图像,获取训练图像中的目标图像块和目标图像块中的多个目标像素点;

S33:依据各所述目标像素点,确定各目标像素点与相邻像素点之间的像素对比度特征;

S34:依据所述像素对比度特征,获取各目标像素点的结构化边缘响应;

S35:对各目标像素点进行非极大抑制,确定各目标像素点的边缘方向,完成随机森林训练;

S36:将所述目标图像输入训练完成的随机森林模型进行多尺度边缘检测,输出所述边缘检测图像。

4. 根据权利要求1所述的基于文本角点检测的学习陪护方法,其特征在于,所述S5包括:

S51:对所述单目深度图像进行梯度计算,提取出高于预设深度梯度阈值的区域图像;

S52:对所述区域图像和所述边缘检测图像进行逐像素融合,得出融合后的二值图像;

S53:对所述二值图像进行霍夫变换,获取初始边缘线段;

S54:对所述初始边缘线段进行二次筛选,输出筛选后的目标线段;

S55:依据所述目标线段,连接各目标线段,输出各目标线段之间的交点位置作为所述初步顶点位置。

5. 根据权利要求4所述的基于文本角点检测的学习陪护方法,其特征在于,所述S54包括:

S541:获取预设的线段长度阈值和预期边缘方向;

S542:将初始边缘线段的线段长度与所述线段长度阈值进行比较,输出大于所述线段长度阈值的对应合格线段;

S543:获取预设的角度阈值,将合格线段的方向与预期边缘方向之间的夹角与所述角度阈值进行比较,输出夹角小于所述角度阈值的对应线段作为所述目标线段。

6.根据权利要求1至5任一项所述的基于文本角点检测的学习陪护方法,其特征在于,所述S6包括:

S61:依据所述初步顶点位置,对初步顶点之间的几何特征进行特征提取,输出关键特征信息;

S62:依据所述关键特征信息和预设的学生作业文本对应的标准特征信息,对所述初步顶点位置进行调整和校验,输出所述学生作业文本顶点位置。

7.根据权利要求6所述的基于文本角点检测的学习陪护方法,其特征在于,所述S62包括:

S621:获取所述关键特征信息,其中,所述关键特征信息至少包括:对角线长度、对角线与水平线之间的目标夹角、边缘线段到顶点之间的距离、对边长度和内角角度;

S622:依据所述对角线长度和所述目标夹角,对初步顶点进行预筛选,输出预筛选后符合预设的对角线特征的第一候选顶点;

S623:依据最小二乘法,结合所述边缘线段到顶点之间的距离,对第一候选顶点进行补充筛选,输出筛选后的第二候选顶点;

S624:依据所述对边长度和所述内角角度,对所述第二候选顶点进行校验,将校验成功的顶点位置作为所述目标顶点位置。

8.一种基于文本角点检测的学习陪护装置,其特征在于,所述装置包括:

图像获取模块,用于获取学生学习陪护场景下的实时视频数据,将所述实时视频数据分解为多帧图像;

预处理模块,用于对各所述图像进行亮度调整和对比度调整,输出调整后的目标图像;

边缘检测模块,用于利用结构化森林算法,对所述目标图像进行边缘检测,输出边缘检测图像;

深度检测模块,用于对所述目标图像进行单目深度检测,输出单目深度图像;

信息融合模块,用于对所述边缘检测图像和单目深度图像进行信息融合,依据融合结果,输出学生作业文本对应的初步顶点位置;

几何约束调整模块,用于对所述初步顶点位置进行几何约束调整,输出调整后的目标顶点位置作为学生作业文本顶点位置。

9.一种学习陪护机器人,其特征在于,包括:至少一个处理器、至少一个存储器以及存储在所述存储器中的计算机程序指令,当所述计算机程序指令被所述处理器执行时实现如权利要求1-7中任一项所述的方法。

10.一种存储介质,其上存储有计算机程序指令,其特征在于,当所述计算机程序指令被处理器执行时实现如权利要求1-7中任一项所述的方法。

## 基于文本角点检测的学习陪护方法、装置、机器人及介质

### 技术领域

[0001] 本发明涉及学习陪护技术领域,尤其涉及一种基于文本角点检测的学习陪护方法、装置、机器人及介质。

### 背景技术

[0002] 在现实生活的学习陪护中,学生作业通常具有四个顶点,通过准确检测学生作业的四个顶点,可以实现多个重要功能:例如,首先,在作业边缘裁剪方面,通过获取精确的边缘范围,实现了对作业区域的准确裁剪,这为后续的图像处理和分析提供了精确的输入;其次,在自动文档识别方面,检测作业的四个顶点能够自动识别文档的区域,在处理大量学生作业时尤为关键,能够提高处理效率;此外,图像校正功能通过检测四个顶点,能够对作业进行角度和倾斜程度的校正,使得作业图像在一个标准的平面上,更易于后续处理和分析;在提高兼容性方面,学生使用不同设备拍摄或扫描作业,通过检测四个顶点,更具兼容性,能够适应不同格式和角度的作业图片;最后,准确检测作业矩形框的四个顶点有助于减少因不准确截取或图像畸变引起的误差,这对于后续的作业评估和处理至关重要。总体而言,检测学生作业顶点的过程提高了自动化处理的准确性和效率,使学生提交作业更为便捷,同时为学术评估和数据分析提供了可靠基础。目前,检测学生作业的四个顶点主要依赖于计算机视觉和图像处理技术,常见的方法包括边缘检测算法、角点检测算法、霍夫变换、深度学习方法及图像预处理,然而,这些方法在应对不同光照、背景噪音、遮挡等情况时存在一些技术问题,如准确性受到影响、角点检测失败或误差、霍夫变换参数调整难度等,解决这些问题需要综合考虑图像特性、算法鲁棒性、参数设置等因素,并在实际应用中平衡算法的实时性和可扩展性,可能需要对不同技术进行组合或调整,以适应特定场景下的需求。

[0003] 现有中国专利CN110108231A公开了一种基于角点检测的矩体箱体三维尺寸测量方法及存储介质,所述方法包括:利用RGB摄像头所获取的彩色图信息,找到矩体箱体的端点在图像上可能存在的位置,再有效利用深度摄像头所获取的深度图信息及其内参,通过结合彩色图信息和深度图信息结合实现快速高效精确地获取到矩体箱体的三维尺寸信息;对彩色图CS0进行数据预处理和灰度处理得出灰度图GS3,以及进行边缘提取操作和二值化处理得出灰度图GS5;然而,首先,上述专利采用的深度摄像头造价高昂,并不适用于学习陪护这种低成本的实际场景,并且在实际应用中,确保RGB摄像头和深度摄像头的信息同步涉及硬件和软件的协同工作,需要确保两种类型的图像在时间和空间上的匹配,匹配技术难度大,导致最终角点检测的准确度低。

[0004] 为此,如何在学习陪护场景下,低成本、准确地检测作业文本四周角点位置是亟待解决的问题。

### 发明内容

[0005] 有鉴于此,本发明提供了一种基于文本角点检测的学习陪护方法、装置、机器人及介质,用以解决现有技术中检测作业文本四周角点位置准确度低、成本高的问题。

[0006] 本发明采用的技术方案是：

[0007] 第一方面,本发明提供了一种基于文本角点检测的学习陪护方法,所述方法包括：

[0008] S1:获取学生学习陪护场景下的实时视频数据,将所述实时视频数据分解为多帧图像；

[0009] S2:对各所述图像进行亮度调整和对比度调整,输出调整后的目标图像；

[0010] S3:利用结构化森林算法,对所述目标图像进行边缘检测,输出边缘检测图像；

[0011] S4:对所述目标图像进行单目深度检测,输出单目深度图像；

[0012] S5:对所述边缘检测图像和单目深度图像进行信息融合,依据融合结果,输出学生作业文本对应的初步顶点位置；

[0013] S6:对所述初步顶点位置进行几何约束调整,输出调整后的目标顶点位置作为学生作业文本顶点位置。

[0014] 优选地,所述S2包括：

[0015] S21:对各所述图像进行预处理,输出预处理之后灰度化的图像；

[0016] S22:对灰度化的图像中亮度分量和对比度分量分别进行调整,输出调整后的亮度分量和调整后的对比度分量；

[0017] S23:依据调整后的亮度分量和调整后的对比度分量,结合预处理之前图像的颜色信息,输出所述目标图像。

[0018] 优选地,所述S3包括：

[0019] S31:获取多帧学生学习陪护场景下的训练图像和对应的边缘图像；

[0020] S32:依据所述训练图像,获取训练图像中的目标图像块和目标图像块中的多个目标像素点；

[0021] S33:依据各所述目标像素点,确定各目标像素点与相邻像素点之间的像素对比度特征；

[0022] S34:依据所述像素对比度特征,获取各目标像素点的结构化边缘响应；

[0023] S35:对各目标像素点进行非极大抑制,确定各目标像素点的边缘方向,完成随机森林训练；

[0024] S36:将所述目标图像输入训练完成的随机森林模型进行多尺度边缘检测,输出所述边缘检测图像。

[0025] 优选地,所述S5包括：

[0026] S51:对所述单目深度图像进行梯度计算,提取出高于预设深度梯度阈值的区域图像；

[0027] S52:对所述区域图像和所述边缘检测图像进行逐像素融合,得出融合后的二值图像；

[0028] S53:对所述二值图像进行霍夫变换,获取初始边缘线段；

[0029] S54:对所述初始边缘线段进行二次筛选,输出筛选后的目标线段；

[0030] S55:依据所述目标线段,连接各目标线段,输出各目标线段之间的交点位置作为所述初步顶点位置。

[0031] 优选地,所述S54包括：

[0032] S541:获取预设的线段长度阈值和预期边缘方向；

[0033] S542:将初始边缘线段的线段长度与所述线段长度阈值进行比较,输出大于所述线段长度阈值的对应合格线段;

[0034] S543:获取预设的角度阈值,将合格线段的方向与预期边缘方向之间的夹角与所述角度阈值进行比较,输出夹角小于所述角度阈值的对应线段作为所述目标线段。

[0035] 优选地,所述S6包括:

[0036] S61:依据所述初步顶点位置,对初步顶点之间的几何特征进行特征提取,输出关键特征信息;

[0037] S62:依据所述关键特征信息和预设的学生作业文本对应的标准特征信息,对所述初步顶点位置进行调整和校验,输出所述学生作业文本顶点位置。

[0038] 优选地,所述S62包括:

[0039] S621:获取所述关键特征信息,其中,所述关键特征信息至少包括:对角线长度、对角线与水平线之间的目标夹角、边缘线段到顶点之间的距离、对边长度和内角角度;

[0040] S622:依据所述对角线长度和所述目标夹角,对初步顶点进行预筛选,输出预筛选后符合预设的对角线特征的第一候选顶点;

[0041] S623:依据最小二乘法,结合所述边缘线段到顶点之间的距离,对第一候选顶点进行补充筛选,输出筛选后的第二候选顶点;

[0042] S624:依据所述对边长度和所述内角角度,对所述第二候选顶点进行校验,将校验成功的顶点位置作为所述目标顶点位置。

[0043] 第二方面,本发明提供了一种基于文本角点检测的学习陪护装置,所述装置包括:

[0044] 图像获取模块,用于获取学生学习陪护场景下的实时视频数据,将所述实时视频数据分解为多帧图像;

[0045] 预处理模块,用于对各所述图像进行亮度调整和对比度调整,输出调整后的目标图像;

[0046] 边缘检测模块,用于利用结构化森林算法,对所述目标图像进行边缘检测,输出边缘检测图像;

[0047] 深度检测模块,用于对所述目标图像进行单目深度检测,输出单目深度图像;

[0048] 信息融合模块,用于对所述边缘检测图像和单目深度图像进行信息融合,依据融合结果,输出学生作业文本对应的初步顶点位置;

[0049] 几何约束调整模块,用于对所述初步顶点位置进行几何约束调整,输出调整后的目标顶点位置作为学生作业文本顶点位置。

[0050] 第三方面,本发明实施例还提供了一种学习陪护机器人,包括:至少一个处理器、至少一个存储器以及存储在存储器中的计算机程序指令,当计算机程序指令被处理器执行时实现如上述实施方式中第一方面的方法。

[0051] 第四方面,本发明实施例还提供了一种存储介质,其上存储有计算机程序指令,当计算机程序指令被处理器执行时实现如上述实施方式中第一方面的方法。

[0052] 综上所述,本发明的有益效果如下:

[0053] 本发明提供的基于文本角点检测的学习陪护方法、装置、机器人及介质,所述方法包括:获取学生学习陪护场景下的实时视频数据,将所述实时视频数据分解为多帧图像;对各所述图像进行亮度调整和对比度调整,输出调整后的目标图像;利用结构化森林算法,对

所述目标图像进行边缘检测,输出边缘检测图像;对所述目标图像进行单目深度检测,输出单目深度图像;对所述边缘检测图像和单目深度图像进行信息融合,依据融合结果,输出学生作业文本对应的初步顶点位置;对所述初步顶点位置进行几何约束调整,输出调整后的目标顶点位置作为学生作业文本顶点位置。本发明通过对图像进行亮度调整和对比度调整,模拟了人眼感知机制,使得输出的目标图像更符合人眼的视觉体验,有助于提高图像的观感质量,使得后续的处理更为准确;使用结构化森林算法进行边缘检测,有效地提取出图像中的边缘信息,尤其是对于学生作业文本等目标区域的边缘,有助于后续的深度检测和顶点位置的准确定位;利用单目深度检测获取深度图像,有助于理解场景中物体的距离信息,降低了检测成本,在学生作业学习场景下,深度信息能够帮助更好地理解和分析学习材料,同时也有助于后续的文本检测和顶点位置的调整;通过将边缘检测图像和单目深度图像进行信息融合,综合利用两者的优势,提高对学生作业文本位置的判别能力,深度信息有助于排除背景干扰,结合边缘信息可以更准确地定位文本区域;对初步顶点位置进行几何约束调整,更好地符合矩形结构,提高了最终顶点位置的精度,这对于后续的文本分析和理解提供了更可靠的基础。总的来说,本发明通过多阶段的图像处理和信息融合,结合了亮度调整、对比度调整、边缘检测、深度检测和几何约束调整等多种技术手段,提高了对学生作业文本的检测和定位的准确性和可靠性。

## 附图说明

[0054] 为了更清楚地说明本发明实施例的技术方案,下面将对本发明实施例中所需要使用的附图作简单地介绍,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,这些均在本发明的保护范围内。

[0055] 图1为本发明实施例1中基于文本角点检测的学习陪护方法的整体工作的流程示意图;

[0056] 图2为本发明实施例1中对各所述图像进行亮度调整和对比度调整的流程示意图;

[0057] 图3为本发明实施例1中对所述目标图像进行边缘检测的流程示意图;

[0058] 图4为本发明实施例1中对所述边缘检测图像和单目深度图像进行信息融合的流程示意图;

[0059] 图5为本发明实施例1中对所述初始边缘线段进行二次筛选的流程示意图;

[0060] 图6为本发明实施例1中对所述初步顶点位置进行几何约束调整的流程示意图;

[0061] 图7为本发明实施例1中对所述初步顶点位置进行调整和校验的流程示意图;

[0062] 图8为本发明实施例2中基于文本角点检测的学习陪护装置的结构框图;

[0063] 图9为本发明实施例3中学习陪护机器人的结构示意图。

## 具体实施方式

[0064] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述。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。在本发明的描述中,需要理解的是,术语“中心”、“上”、“下”、“前”、“后”、“左”、“右”、“竖

直”、“水平”、“顶”、“底”、“内”、“外”等指示的方位或位置关系为基于附图所示的方位或位置关系,仅是为了便于描述本申请和简化描述,而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作,因此不能理解为对本发明的限制。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……”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。如果不冲突,本发明实施例以及实施例中的各个特征可以相互结合,均在本发明的保护范围之内。

[0065] 实施例1

[0066] 请参见图1,本发明实施例1公开了一种基于文本角点检测的学习陪护方法,所述方法包括:

[0067] S1:获取学生学习陪护场景下的实时视频数据,将所述实时视频数据分解为多帧图像;

[0068] 具体地,通过摄像设备包括摄像头或其他视频采集器,获取学生学习陪护场景下的实时视频数据,包含学生学习环境的连续视频信号;将获取的实时视频数据分解为多帧图像,将连续的视频数据划分为一系列的静止图像帧,形成图像序列,每一帧都代表了视频数据在特定时间点的静态画面;将连续的视频数据拆分为离散的图像,为后续的处理步骤提供静态图像数据,有助于对每一时刻的场景进行更详细的分析和处理。

[0069] S2:对各所述图像进行亮度调整和对比度调整,输出调整后的目标图像;

[0070] 在一实施例中,请参见图2,所述S2包括:

[0071] S21:对各所述图像进行预处理,输出预处理之后灰度化的图像;

[0072] 具体地,首先对各个图像进行预处理,预处理的目的是为了简化图像,将其转化为更易处理的形式,其中,预处理包括将彩色图像转换为灰度图像,即去除颜色信息,只保留亮度信息,有助于后续的亮度和对比度调整。

[0073] S22:对灰度化的图像中亮度分量和对比度分量分别进行调整,输出调整后的亮度分量和调整后的对比度分量;

[0074] 具体地,利用Retinex理论,通过调整灰度化的图像的亮度分量来增强或减弱整体亮度水平,亮度调整通过以下方式实现:通过应用Gamma函数调整图像的灰度映射,以改变图像的整体亮度;通过调整图像的对比度分量来增加或减少图像中不同区域之间的灰度差异,对比度调整通过以下方式实现:通过重新分配图像灰度级的分布,增强或减弱图像中不同亮度级别的对比度。

[0075] S23:依据调整后的亮度分量和调整后的对比度分量,结合预处理之前图像的颜色信息,输出所述目标图像。

[0076] 具体地,将调整后的亮度分量和对比度分量与预处理之前的图像的颜色信息结合,生成最终的目标图像,确保调整后的图像在保留原始颜色信息的同时,经过亮度和对比度的优化,以满足人眼感知的需求,其中所述原始颜色信息包括红、绿、蓝的通道信息,结合方式至少包括线性组合的方式。

[0077] S3:利用结构化森林算法,对所述目标图像进行边缘检测,输出边缘检测图像;

- [0078] 在一实施例中,请参见图3,所述S3包括:
- [0079] S31:获取多帧学生学习陪护场景下的训练图像和对应的边缘图像;
- [0080] 具体地,从学习陪护场景中获取多帧训练图像,每个训练图像都有对应的边缘图像,提供了真实的边缘信息,用于训练模型。
- [0081] S32:依据所述训练图像,获取训练图像中的目标图像块和目标图像块中的多个目标像素点;
- [0082] 具体地,在每个训练图像中,选择包含学生作业文本边缘的目标图像块,从选定的目标图像块中采样多个目标像素点,这些目标像素点用于后续的特征计算。
- [0083] S33:依据各所述目标像素点,确定各目标像素点与相邻像素点之间的像素对比度特征;
- [0084] 具体地,对每个目标像素点及其相邻像素点计算像素对比度特征,该特征用于描述目标区域内的灰度变化,其计算公式为: $f(x) = |I(x+p) - I(x-p)|$ ,其中 $I(x+p)$ 表示像素点 $x+p$ 的灰度值, $I(x-p)$ 表示像素点 $x-p$ 的灰度值, $p$ 表示像素点 $x$ 周围8邻域内的一个点。
- [0085] S34:依据所述像素对比度特征,获取各目标像素点的结构化边缘响应;
- [0086] 具体地,利用像素对比度特征,得到的 $f(x)$ 值反映了目标像素点和其相邻像素之间的灰度变化,结构化边缘响应考虑将像素对比度特征在不同方向上进行加权求和,以捕捉更丰富的结构信息,使用结构化滤波等方法,根据不同方向的权重对像素对比度特征进行组合。结构化边缘响应的计算旨在突出目标区域内的结构化边缘信息,对于学生作业文本而言,有助于提取文本边缘的特征。
- [0087] S35:对各目标像素点进行非极大抑制,确定各目标像素点的边缘方向,完成随机森林训练;
- [0088] 具体地,对于计算得到的结构化边缘响应,应用非极大抑制(NMS)操作,以保留具有最大响应的边缘方向,在每个像素点上,选择具有最大结构化边缘响应的方向作为该点的主要边缘方向,针对每个目标像素点,完成结构化边缘响应的训练。
- [0089] S36:将所述目标图像输入训练完成的随机森林模型进行多尺度边缘检测,输出所述边缘检测图像。
- [0090] 具体地,将所述目标图像输入训练得到的随机森林以在多尺度上执行边缘检测,使用集成学习方法融合多个决策树的预测结果,得到最终的边缘检测图像。
- [0091] S4:对所述目标图像进行单目深度检测,输出单目深度图像;
- [0092] 具体地,单目深度计算模块是一种用于从单张图像中估计场景深度信息的模型或算法,在这里,所使用的计算模块是MobileDepth模型,该模型是一种轻量级、适用于移动设备的深度学习模型,专门设计用于单目深度估计任务;MobileDepth模型基于深度学习架构,例如卷积神经网络(CNN),被训练用于从单目图像中预测场景的深度信息,经过大量的深度图像数据集的训练,以学习图像中物体的尺寸和距离等深度信息,将目标图像输入MobileDepth模型,通过模型的前向传播,得到对应场景的深度图像,深度图像中的每个像素值表示图像中对应点到摄像头的距离,形成了场景的深度信息。通过获得目标图像的单目深度图,提供了关于图像中物体距离的信息,这对于后续的信息融合步骤是重要的,因为结合深度信息可以更准确地理解和定位学生作业文本的位置。
- [0093] S5:对所述边缘检测图像和单目深度图像进行信息融合,依据融合结果,输出学生

作业文本对应的初步顶点位置；

[0094] 在一实施例中,请参见图4,所述S5包括:

[0095] S51:对所述单目深度图像进行梯度计算,提取出高于预设深度梯度阈值的区域图像;

[0096] 具体地,对单目深度图进行梯度计算,获取每个像素点深度的变化率;通过计算深度图中每个像素点的梯度,得到深度变化的方向和强度;预设深度梯度阈值,筛选出高于预设深度梯度阈值的深度变化显著的区域,这些区域对应于作业文本的边缘。

[0097] S52:对所述区域图像和所述边缘检测图像进行逐像素融合,得出融合后的二值图像;

[0098] 具体地,获取之前通过结构化森林算法得到的边缘检测图像,逐像素地将边缘特征图与深度梯度显著区域进行融合,确保只有同时在边缘特征图和深度梯度图中显著的点被认为是作业文本的潜在边缘。这一步有助于减少误检,只保留深度和边缘信息都显著的区域。

[0099] S53:对所述二值图像进行霍夫变换,获取初始边缘线段;

[0100] 具体地,将融合后的图像进行二值化,得到二值图像,应用霍夫变换,提取表示作业边缘的线段,其中,霍夫变换可以识别图像中的直线,这里用于检测作业文本的边缘线段。

[0101] S54:对所述初始边缘线段进行二次筛选,输出筛选后的目标线段;

[0102] 在一实施例中,请参见图5,所述S54包括:

[0103] S541:获取预设的线段长度阈值和预期边缘方向;

[0104] 具体地,获取预设的线段长度阈值和预期边缘方向;例如从系统参数中获取线段长度阈值为10像素,预期边缘方向为水平方向。

[0105] S542:将初始边缘线段的线段长度与所述线段长度阈值进行比较,输出大于所述线段长度阈值的对应合格线段;

[0106] 具体地,将初始边缘线段的线段长度与所述线段长度阈值进行比较,假设初始提取的边缘线段中有一条长度为15像素,另一条长度为8像素,与预设的线段长度阈值比较后,输出长度大于10像素的线段,即长度为15像素的线段。

[0107] S543:获取预设的角度阈值,将合格线段的方向与预期边缘方向之间的夹角与所述角度阈值进行比较,输出夹角小于所述角度阈值的对应线段作为所述目标线段。

[0108] 具体地,取角度阈值为20度,将合格线段的方向与预期边缘方向之间的夹角与角度阈值进行比较:对于已筛选出的长度合格的线段,计算每条线段与水平方向的夹角,假设计算得到第一条线段与水平方向的夹角为15度,第二条线段为25度,与预设的角度阈值20度比较后,输出夹角小于20度的线段,即第一条线段。通过线段长度和角度的筛选后,得到了符合条件的目标线段,为后续的顶点计算提供了可靠的输入。

[0109] S55:依据所述目标线段,连接各目标线段,输出各目标线段之间的交点位置作为所述初步顶点位置。

[0110] 具体地,获取提取的边缘线段,计算线段之间的交点,这些交点被认为是作业文本的初步顶点。这一步为后续的几何约束调整提供了初步的顶点位置。

[0111] S6:对所述初步顶点位置进行几何约束调整,输出调整后的目标顶点位置作为学

生作业文本顶点位置。

[0112] 具体地,对通过霍夫变换得到的初步顶点集合进行几何约束调整,输出调整后的目标顶点位置,以提高学生作业文本顶点位置的准确性和一致性。

[0113] 在一实施例中,请参见图6,所述S6包括:

[0114] S61:依据所述初步顶点位置,对初步顶点之间的几何特征进行特征提取,输出关键特征信息;

[0115] 具体地,利用初步检测的顶点集合,计算顶点之间的几何特征,例如边长、角度,对于每一对顶点,可以计算其之间的距离作为边长,以及两条边之间的夹角,输出提取得到的关键特征信息,至少包括各边的长度、角度。

[0116] S62:依据所述关键特征信息和预设的学生作业文本对应的标准特征信息,对所述初步顶点位置进行调整和校验,输出所述学生作业文本顶点位置。

[0117] 具体地,利用关键特征信息和预设的学生作业文本标准特征信息进行调整和校验,对于每个顶点,根据标准特征信息,调整其位置,以使得提取的几何特征更加接近标准特征,输出校验和调整后的学生作业文本顶点位置,这一位置更符合预期的几何特征,提高了整个系统对学生作业文本位置的准确性。

[0118] 在一实施例中,请参见图7,所述S62包括:

[0119] S621:获取所述关键特征信息,其中,所述关键特征信息至少包括:对角线长度、对角线与水平线之间的目标夹角、边缘线段到顶点之间的距离、对边长度和内角角度;

[0120] S622:依据所述对角线长度和所述目标夹角,对初步顶点进行预筛选,输出预筛选后符合预设的对角线特征的第一候选顶点;

[0121] 具体地,对于每个检测到的交点,计算由该点和其他点组成的两条对角线与水平线的夹角,以两两组合的方式获得不同的对角线,计算对角线长度:对于每组对角线,计算其长度,即交点之间的距离,设定预设的阈值,包括夹角阈值 $\theta_{\text{threshold}}$ 和最小对角线长度阈值 $d_{\text{min}}$ 、最大对角线长度阈值 $d_{\text{max}}$ 。剔除那些明显不符合矩形对角线关系的顶点对,保留满足以下条件的交点:

[0122]  $\theta_{\text{threshold}} < |\theta_{\text{AB}} - \theta_{\text{CD}}| < 180^\circ - \theta_{\text{threshold}}$

[0123]  $d_{\text{min}} < ||\text{AB}|| < d_{\text{max}}$  和  $d_{\text{min}} < ||\text{CD}|| < d_{\text{max}}$

[0124] 其中,AB和CD是对角线, $\theta_{\text{AB}}$ 和 $\theta_{\text{CD}}$ 是这两条线段与水平线的夹角, $||\text{AB}||$ 是线段AB的长度。

[0125] S623:依据最小二乘法,结合所述边缘线段到顶点之间的距离,对第一候选顶点进行补充筛选,输出筛选后的第二候选顶点;

[0126] 具体地,针对筛选后的第一候选顶点,通过几何知识提取关键特征信息,包括边缘线段到各个顶点的距离,使用最小二乘法,建立目标函数,即边缘线段到各个顶点的距离之和的平方,作为需要最小化的误差,通过迭代优化过程,逐步调整各个顶点的位置,减小目标函数的值,使得边缘线段与顶点的距离更加符合最小二乘法的要求,即满足:

[0127] 
$$\min_{V_i} \sum_{j=1}^n ||V_i - P_j||^2$$

[0128] 其中, $V_i$ 是顶点的位置, $P_j$ 是第j个边缘线段上的点, $||V_i - P_j||$ 是点到顶点的欧式距离,输出筛选后的第二候选顶点。通过最小二乘法的优化,顶点的位置得到了精确的调

整,使得边缘线段与顶点的距离达到最小,有助于进一步提高检测到的矩形的几何精度,确保最终的顶点位置更符合实际的作业文本形状。

[0129] S624:依据所述对边长度和所述内角角度,对所述第二候选顶点进行校验,将校验成功的顶点位置作为所述目标顶点位置。

[0130] 具体地,校验优化后的第二候选顶点对之间的距离,确保对边距离一致性,预先设置距离比例因子 $\alpha$ ,对任意第二候选顶点连接形成的对边AB和CD,校验其长度比例是否满足以下关系:

$$[0131] \quad \frac{1}{\alpha} < \frac{\|AB\|}{\|CD\|} < \alpha$$

[0132] 如果不满足,则重新调整顶点位置,使对边长度比接近于1;

[0133] 并对优化后的顶点集合进行角度校验,由于学生作业文本近似矩形,检验四个内角是否接近 $90^\circ$

[0134]  $90^\circ - \delta < \angle ABC < 90^\circ + \delta$ 对所有角ABC

[0135] 其中, $\delta$ 是允许的预设角度偏差,如果实际角度与 $90^\circ$ 的偏差过大,则需要重新调整顶点,使其更加接近真实的矩形顶点;通过顶点间距离一致性校验和角度一致性校验,确保了优化后的顶点集合形成的边长比例和角度接近于理想的作业文本实际形状,提高了检测的几何精度。

[0136] 实施例2

[0137] 请参见图8,本发明实施例2还提供了一种基于文本角点检测的学习陪护装置,所述装置包括:

[0138] 图像获取模块,用于获取学生学习陪护场景下的实时视频数据,将所述实时视频数据分解为多帧图像;

[0139] 预处理模块,用于对各所述图像进行亮度调整和对比度调整,输出调整后的目标图像;

[0140] 在一实施例中,所述预处理模块包括:

[0141] 图像灰度化处理子模块,用于对各所述图像进行预处理,输出预处理之后灰度化的图像;

[0142] 亮度分量和对比度分量调整子模块,用于对灰度化的图像中亮度分量和对比度分量分别进行调整,输出调整后的亮度分量和调整后的对比度分量;

[0143] 目标图像获取子模块,用于依据调整后的亮度分量和调整后的对比度分量,结合预处理之前图像的颜色信息,输出所述目标图像。

[0144] 边缘检测模块,用于利用结构化森林算法,对所述目标图像进行边缘检测,输出边缘检测图像;

[0145] 在一实施例中,所述边缘检测模块包括:

[0146] 训练图像和边缘图像获取子模块,用于获取多帧学生学习陪护场景下的训练图像和对应的边缘图像;

[0147] 目标像素点获取子模块,用于依据所述训练图像,获取训练图像中的目标图像块和目标图像块中的多个目标像素点;

[0148] 像素对比度特征确定子模块,用于依据各所述目标像素点,确定各目标像素点与

相邻像素点之间的像素对比度特征；

[0149] 结构化边缘响应获取子模块,用于依据所述像素对比度特征,获取各目标像素点的结构化边缘响应；

[0150] 非极大抑制子模块,用于对各目标像素点进行非极大抑制,确定各目标像素点的边缘方向,完成随机森林训练；

[0151] 多尺度边缘检测模块,用于将所述目标图像输入训练完成的随机森林模型进行多尺度边缘检测,输出所述边缘检测图像。

[0152] 深度检测模块,用于对所述目标图像进行单目深度检测,输出单目深度图像；

[0153] 信息融合模块,用于对所述边缘检测图像和单目深度图像进行信息融合,依据融合结果,输出学生作业文本对应的初步顶点位置；

[0154] 在一实施例中,所述信息融合模块包括：

[0155] 梯度计算子模块,用于对所述单目深度图像进行梯度计算,提取出高于预设深度梯度阈值的区域图像；

[0156] 图像融合子模块,用于对所述区域图像和所述边缘检测图像进行逐像素融合,得出融合后的二值图像；

[0157] 霍夫变换子模块,用于对所述二值图像进行霍夫变换,获取初始边缘线段；

[0158] 二次筛选子模块,用于对所述初始边缘线段进行二次筛选,输出筛选后的目标线段；

[0159] 在一实施例中,所述二次筛选子模块包括：

[0160] 线段长度阈值和预期边缘方向获取单元,用于获取预设的线段长度阈值和预期边缘方向；

[0161] 合格线段确定单元,用于将初始边缘线段的线段长度与所述线段长度阈值进行比较,输出大于所述线段长度阈值的对应合格线段；

[0162] 目标线段获取单元,用于获取预设的角度阈值,将合格线段的方向与预期边缘方向之间的夹角与所述角度阈值进行比较,输出夹角小于所述角度阈值的对应线段作为所述目标线段。

[0163] 初步顶点位置获取子模块,用于依据所述目标线段,连接各目标线段,输出各目标线段之间的交点位置作为所述初步顶点位置。

[0164] 几何约束调整模块,用于对所述初步顶点位置进行几何约束调整,输出调整后的目标顶点位置作为学生作业文本顶点位置。

[0165] 在一实施例中,所述几何约束调整模块包括：

[0166] 特征提取子模块,用于依据所述初步顶点位置,对初步顶点之间的几何特征进行特征提取,输出关键特征信息；

[0167] 初步顶点位置调整校验子模块,用于依据所述关键特征信息和预设的学生作业文本对应的标准特征信息,对所述初步顶点位置进行调整和校验,输出所述学生作业文本顶点位置。

[0168] 在一实施例中,所述初步顶点位置调整校验子模块包括：

[0169] 关键特征信息获取单元,用于获取所述关键特征信息,其中,所述关键特征信息至少包括:对角线长度、对角线与水平线之间的目标夹角、边缘线段到顶点之间的距离、对边

长度和内角角度；

[0170] 预筛选单元,用于依据所述对角线长度和所述目标夹角,对初步顶点进行预筛选,输出预筛选后符合预设的对角线特征的第一候选顶点；

[0171] 补充筛选单元,用于依据最小二乘法,结合所述边缘线段到顶点之间的距离,对第一候选顶点进行补充筛选,输出筛选后的第二候选顶点；

[0172] 第二候选顶点校验单元,用于依据所述对边长度和所述内角角度,对所述第二候选顶点进行校验,将校验成功的顶点位置作为所述目标顶点位置。

[0173] 具体地,采用本发明实施例提供的基于文本角点检测的学习陪护装置,所述装置包括:图像获取模块,用于获取学生学习陪护场景下的实时视频数据,将所述实时视频数据分解为多帧图像;预处理模块,用于对各所述图像进行亮度调整和对比度调整,输出调整后的目标图像;边缘检测模块,用于利用结构化森林算法,对所述目标图像进行边缘检测,输出边缘检测图像;深度检测模块,用于对所述目标图像进行单目深度检测,输出单目深度图像;信息融合模块,用于对所述边缘检测图像和单目深度图像进行信息融合,依据融合结果,输出学生作业文本对应的初步顶点位置;几何约束调整模块,用于对所述初步顶点位置进行几何约束调整,输出调整后的目标顶点位置作为学生作业文本顶点位置。本装置通过对图像进行亮度调整和对比度调整,模拟了人眼感知机制,使得输出的目标图像更符合人眼的视觉体验,有助于提高图像的观感质量,使得后续的处理更为准确;使用结构化森林算法进行边缘检测,有效地提取出图像中的边缘信息,尤其是对于学生作业文本等目标区域的边缘,有助于后续的深度检测和顶点位置的准确定位;利用单目深度检测获取深度图像,有助于理解场景中物体的距离信息,降低了检测成本,在学生学习场景下,深度信息能够帮助更好地理解和分析学习材料,同时也有助于后续的文本检测和顶点位置的调整;通过将边缘检测图像和单目深度图像进行信息融合,综合利用两者的优势,提高对学生作业文本位置的判别能力,深度信息有助于排除背景干扰,结合边缘信息可以更准确地定位文本区域;对初步顶点位置进行几何约束调整,更好地符合矩形结构,提高了最终顶点位置的精度,这对于后续的文本分析和理解提供了更可靠的基础。总的来说,本装置通过多阶段的图像处理和信息融合,结合了亮度调整、对比度调整、边缘检测、深度检测和几何约束调整等多种技术手段,提高了对学生作业文本的检测和定位的准确性和可靠性。

[0174] 实施例3

[0175] 另外,结合图1描述的本发明实施例1的基于文本角点检测的学习陪护方法可以由学习陪护机器人来实现。图9示出了本发明实施例3提供的学习陪护机器人的硬件结构示意图。

[0176] 学习陪护机器人可以包括处理器以及存储有计算机程序指令的存储器。

[0177] 具体地,上述处理器可以包括中央处理器(CPU),或者特定集成电路(Application Specific Integrated Circuit,ASIC),或者可以被配置成实施本发明实施例的一个或多个集成电路。

[0178] 存储器可以包括用于数据或指令的大容量存储器。举例来说而非限制,存储器可包括硬盘驱动器(Hard Disk Drive,HDD)、软盘驱动器、闪存、光盘、磁光盘、磁带或通用串行总线(Universal Serial Bus,USB)驱动器或者两个或更多个以上这些的组合。在合适的情况下,存储器可包括可移除或不可移除(或固定)的介质。在合适的情况下,存储器可在数

据处理装置的内部或外部。在特定实施例中,存储器是非易失性固态存储器。在特定实施例中,存储器包括只读存储器(ROM)。在合适的情况下,该ROM可以是掩模编程的ROM、可编程ROM(PROM)、可擦除PROM(EPROM)、电可擦除PROM(EEPROM)、电可改写ROM(EAROM)或闪存或者两个或更多个以上这些的组合。

[0179] 处理器通过读取并执行存储器中存储的计算机程序指令,以实现上述实施例中的任意一种基于文本角点检测的学习陪护方法。

[0180] 在一个示例中,学习陪护机器人还可包括通信接口和总线。其中,如图9所示,处理器、存储器、通信接口通过总线连接并完成相互间的通信。

[0181] 通信接口,主要用于实现本发明实施例中各模块、装置、单元和/或设备之间的通信。

[0182] 总线包括硬件、软件或两者,将所述学习陪护机器人的部件彼此耦接在一起。举例来说而非限制,总线可包括加速图形端口(AGP)或其他图形总线、增强工业标准架构(EISA)总线、前端总线(FSB)、超传输(HT)互连、工业标准架构(ISA)总线、无限带宽互连、低引脚数(LPC)总线、存储器总线、微信道架构(MCA)总线、外围组件互连(PCI)总线、PCI-Express(PCI-X)总线、串行高级技术附件(SATA)总线、视频电子标准协会局部(VLB)总线或其他合适的总线或者两个或更多个以上这些的组合。在合适的情况下,总线可包括一个或多个总线。尽管本发明实施例描述和示出了特定的总线,但本发明考虑任何合适的总线或互连。

[0183] 实施例4

[0184] 另外,结合上述实施例1中的基于文本角点检测的学习陪护方法,本发明实施例4还可提供一种计算机可读存储介质来实现。该计算机可读存储介质上存储有计算机程序指令;该计算机程序指令被处理器执行时实现上述实施例中的任意一种基于文本角点检测的学习陪护方法。

[0185] 综上所述,本发明实施例提供了一种基于文本角点检测的学习陪护方法、装置、机器人及介质。

[0186] 需要明确的是,本发明并不局限于上文所描述并在图中示出的特定配置和处理。为了简明起见,这里省略了对已知方法的详细描述。在上述实施例中,描述和示出了若干具体的步骤作为示例。但是,本发明的方法过程并不限于所描述和示出的具体步骤,本领域的技术人员可以在领会本发明的精神后,作出各种改变、修改和添加,或者改变步骤之间的顺序。

[0187] 以上所述的结构框图中所示的功能块可以实现为硬件、软件、固件或者它们的组合。当以硬件方式实现时,其可以例如是电子电路、专用集成电路(ASIC)、适当的固件、插件、功能卡等等。当以软件方式实现时,本发明的元素是被用于执行所需任务的程序或者代码段。程序或者代码段可以存储在机器可读介质中,或者通过载波中携带的数据信号在传输介质或者通信链路上传送。“机器可读介质”可以包括能够存储或传输信息的任何介质。机器可读介质的例子包括电子电路、半导体存储器设备、ROM、闪存、可擦除ROM(EROM)、软盘、CD-ROM、光盘、硬盘、光纤介质、射频(RF)链路,等等。代码段可以经由诸如因特网、内联网等的计算机网络被下载。

[0188] 还需要说明的是,本发明中提及的示例性实施例,基于一系列的步骤或者装置描述一些方法或系统。但是,本发明不局限于上述步骤的顺序,也就是说,可以按照实施例中

提及的顺序执行步骤,也可以不同于实施例中的顺序,或者若干步骤同时执行。

[0189] 以上所述,仅为本发明的具体实施方式,所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,上述描述的系统、模块和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。应理解,本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到各种等效的修改或替换,这些修改或替换都应涵盖在本发明的保护范围之内。

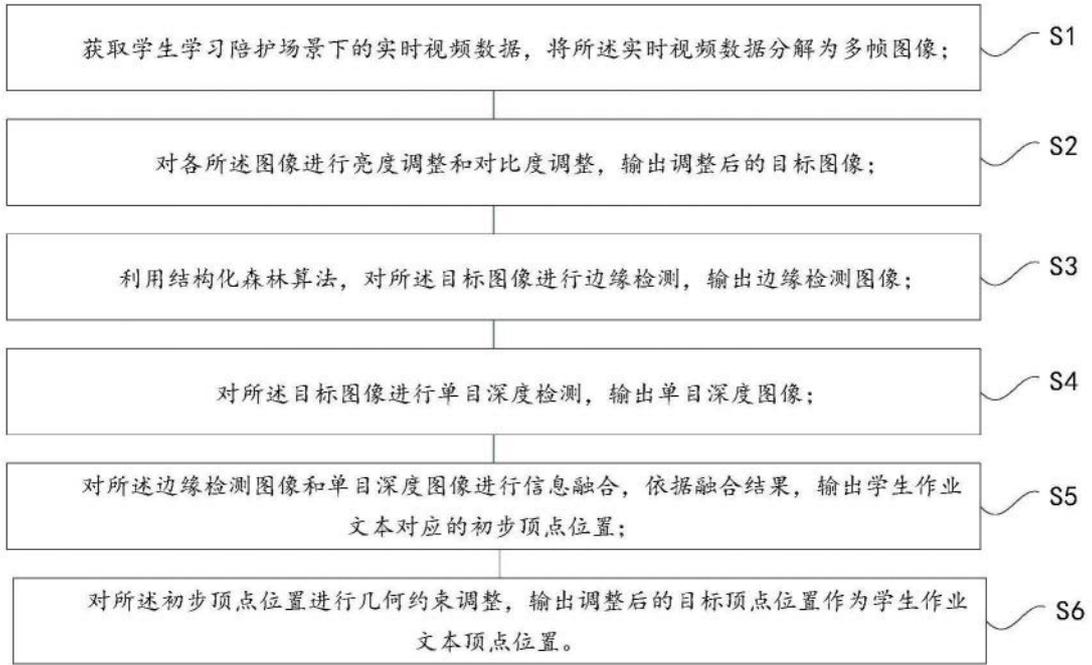


图1



图2

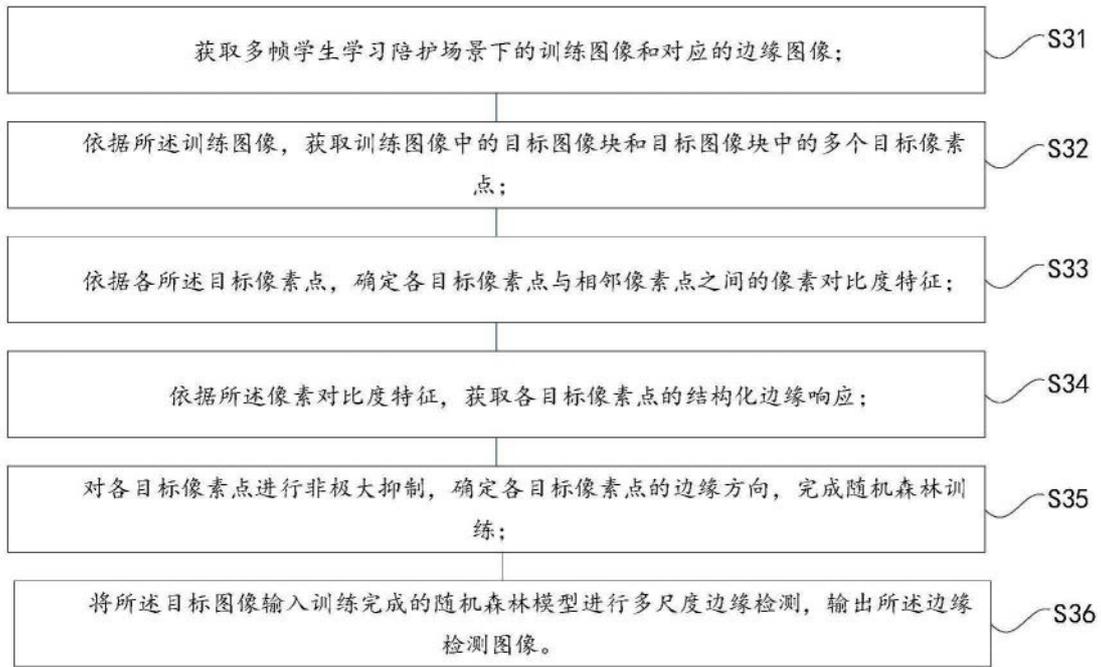


图3

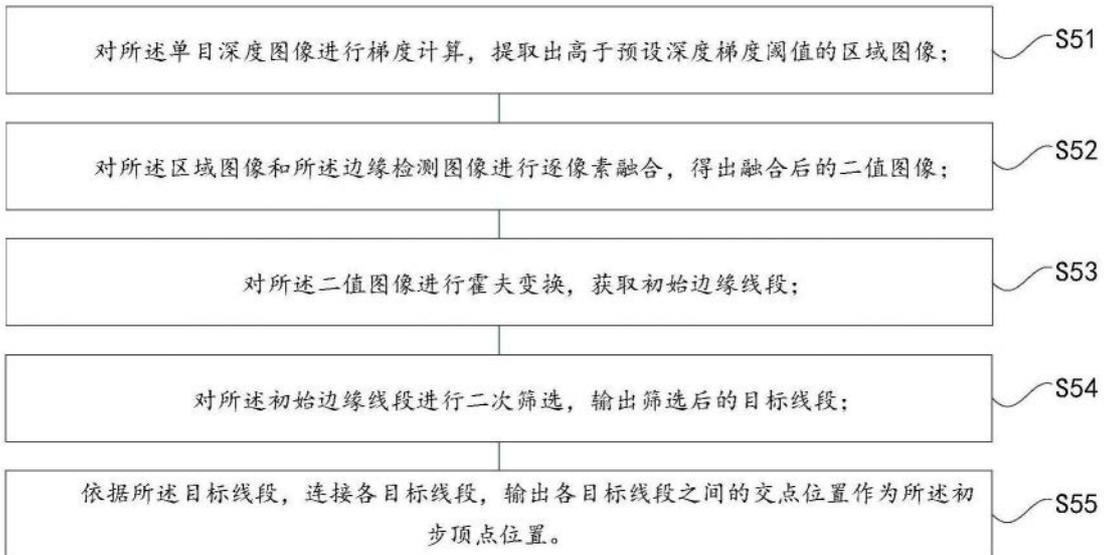


图4



图5

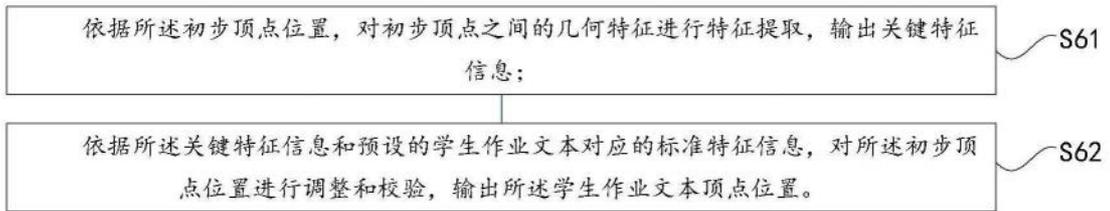


图6

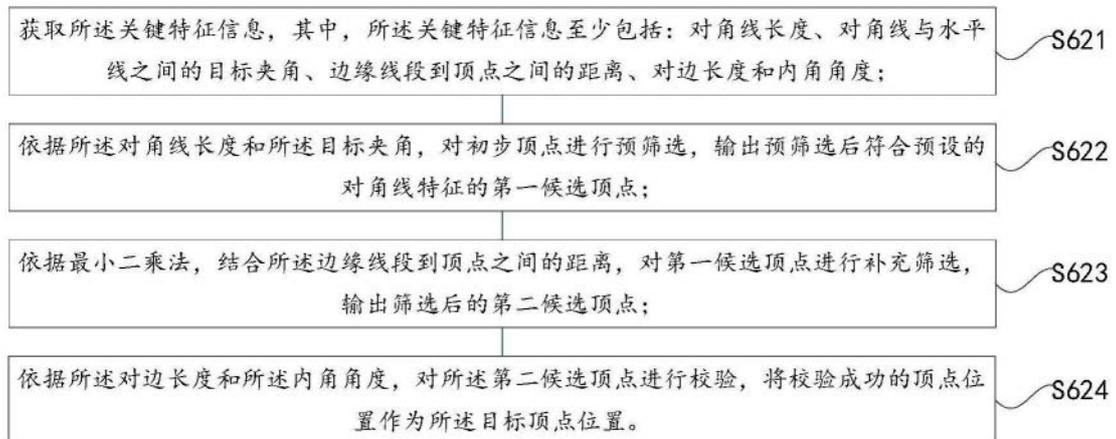


图7



图8

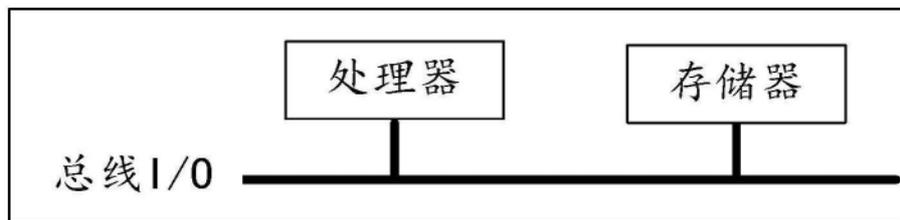


图9