



(12) 发明专利

(10) 授权公告号 CN 114067197 B

(45) 授权公告日 2022. 11. 18

(21) 申请号 202111360831.6

(22) 申请日 2021.11.17

(65) 同一申请的已公布的文献号

申请公布号 CN 114067197 A

(43) 申请公布日 2022.02.18

(73) 专利权人 河南大学

地址 475004 河南省开封市顺河区明伦街
85号

(72) 发明人 何洪权 葛昱彤 申安慧 张琳
朱一沁

(74) 专利代理机构 郑州优盾知识产权代理有限公司 41125

专利代理师 张真真

(51) Int.Cl.

G06V 20/00 (2022.01)

G06V 10/764 (2022.01)

G06V 10/80 (2022.01)

G06V 10/44 (2022.01)

G06V 10/82 (2022.01)

G06T 7/00 (2017.01)

G06T 7/73 (2017.01)

G06T 7/80 (2017.01)

G06T 7/55 (2017.01)

G06T 5/00 (2006.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

(56) 对比文件

CN 113393439 A, 2021.09.14

CN 113393439 A, 2021.09.14

CN 112092929 A, 2020.12.18

CN 113524194 A, 2021.10.22

CN 107248159 A, 2017.10.13

Heiko Hirschmuller. Stereo Processing
by Semiglobal Matching and Mutual
Information.《IEEE Transactions on Pattern
Analysis and Machine Intelligence》.2007,
第30卷(第2期),第328-341页.

审查员 谢婉婉

权利要求书5页 说明书13页 附图7页

(54) 发明名称

一种基于目标检测及双目视觉的管道缺陷
识别与定位方法

(57) 摘要

本发明提出了一种基于目标检测及双目视觉的管道缺陷识别与定位方法,其步骤包括:对双目相机进行标定与立体匹配,得到畸变相对很小的图像,通过立体匹配与深度计算确定图像中某一点在世界坐标系下的具体位置与其像素点位置的具体函数关系;通过目标检测算法训练目标检测模型并对视频图像进行识别;当识别到图像存在某一类缺陷时,该模型将框定缺陷目标,将其中心位置的世界坐标显示在屏幕上。本发明的管道缺陷检测与定位方法具有监测范围广,实时性好,精确度高,并能够准确定位等优点。同时,该方法克服了漏磁检测法等传统管道检测方法只能检测铁磁性材料管道的局限性,且不会产

生虚假信号,对漏报、误报的情况有一定程度的改善。



1. 一种基于目标检测及双目视觉的管道缺陷识别与定位方法,其特征在于,包括相机标定、立体校正、图像捕获与目标检测、立体匹配与深度计算模块四部分;其步骤如下:

步骤一:相机标定:建立相机成像的几何模型,确定空间物体表面上一点的三维几何位置与其在图像中对应点之间的相互关系,由标定板上特征点的图像坐标和世界坐标求解双目摄像机的标定参数,其中,标定参数包括左摄像头和右摄像头的内参、外参和畸变参数;

步骤二:立体校正:通过极线约束进行立体校正,使两幅图像中的对应点在同一条水平极线上,获取校正后的双目摄像机的标定参数;

步骤三:图像捕获与目标检测:自建数据集,进行目标检测模型训练,利用双目摄像机捕获图像,预测左摄像头显示的实时画面信息,并输出物体识别信息;

所述图像捕获与目标检测的实现方法为:

步骤3.1:利用双目相机与网络爬虫捕获并筛选管道缺陷图像,并对管道缺陷图像进行标记与分类,创建数据集;

缺陷的分类如下,除ND类为无注释类,其余9个语义标签为:

OB类:表面损伤类;FO类:障碍物类别;RB类:包含裂纹、断裂和塌陷的类;RO类:植物的根入侵的类;DE类:变形;FS类:错位;GR类:支管;BE类:附着沉积物的类;AF类:固定沉积物的类;

步骤3.2:对管道缺陷图像和标记进行组织目录上的整理,通过getitem方法加载数据集中的管道缺陷图像和标记,使用COCO Evaluator作为评估器,将深度学习模型中涉及的所有内容都放在一个单独的COCO Exp文件中,包括模型设置、训练设置和测试设置;

步骤3.3:设置完毕后,使用COCO开源数据集训练完成的预训练权重初始化模型,在命令行输入训练命令以使用GPU开始训练,训练完成后,得到权重文件以及深度学习目标检测模型;

深度学习卷积神经网络结构,由特征提取、特征融合、预测类别并标定检测框共计三部分构成;

特征提取部分,包括Focus隔行采样拼接结构、四个DarkNet结构;Focus作为整个网络的输入部分,由默认的 $3 \times 640 \times 640$ 输入开始,复制四份,通过切片操作将四个图片切成了四个 $3 \times 320 \times 320$ 的切片,接下来使用concat从深度上拼接这四个切片,生成 $12 \times 320 \times 320$ 的输出,最后经过batch_norm和leaky_relu将结果输入到下一个卷积层;DarkNet结构由CSP与SPP组成,将原输入分成两个分支,分别进行卷积操作使得通道数减半,分支一抛弃了YOLOV5的Bottleneck操作,采用ResUnit结构,随后concat分支一和分支二,从而使得CSP的输入与输出是一样的大小,CSP的目的是为了让模型学习到更多的特征;SPP的输入为 $512 \times 20 \times 20$,经过 1×1 的卷积层后输出 $256 \times 20 \times 20$,然后经过并列的三个MaxPool进行采样,将结果与其初始特征相加,输出 $1024 \times 20 \times 20$,最后用512的卷积核将其恢复到 $512 \times 20 \times 20$;

特征融合部分的作用是混合和组合图像特征,并将图像特征传递到预测层,采用PAFPN,基于Mask R-CNN和FPN框架,加强了信息传播,具有准确保留空间信息的能力,有助于对像素进行适当的定位以形成掩模;

最后一部分网络结构对图像特征进行预测,生成边界框并预测类别;首先采用Decoupled Head提取三个特征层进行目标检测,同时加入了无锚点的检测器提高检测速

度,在输出部分,三个输出分别是 $85 \times 20 \times 20$ 、 $85 \times 40 \times 40$ 、 $85 \times 80 \times 80$;通过对预测结果解码,即得到缺陷预测信息;

深度学习卷积神经网络结构参考YOLOX网络搭建,相比YOLOX网络在特征提取部分减少了部分CBS操作和卷积次数,简化了该网络的特征提取部分;同时,修改了损失函数的计算方式,使用标准二进制交叉熵损失进行训练:

$$L(x, y) = \frac{1}{C} \sum_c^C -[w_c y_c \log(\sigma(x_c)) + (1 - y_c) \log(1 - \sigma(x_c))];$$

其中,C为语义标签数 $\text{num}_{\text{classes}}=9$, $y_c = \begin{cases} 0, & \text{class } c \text{ exists} \\ 1, & \text{no class } c \text{ exists} \end{cases}$ 表示当前图像中是否存在c类, x_c 是c类模型的原始输出, σ 为Sigmoid激活函数, $w_c = \frac{N-N_c}{N_c}$ 为c类的权重,N表示训练中的图像数, N_c 表示包含c类的图像数;

步骤3.4:通过双目相机捕获管道实时图像,并将管道实时图像输入深度学习目标检测模型中,输出物体识别信息;

步骤四:立体匹配与深度计算模块:将经过立体校正后的左摄像头和右摄像头图像与步骤三中的物体识别信息作为输入量传输至立体匹配与深度计算模块,得到双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标;

所述立体匹配与深度计算模块的实现方法为:

步骤4.1:根据立体校正后的双目摄像头左摄像头和右摄像头图像计算得到原图像在预设视差范围内各像素点的匹配代价;

步骤4.2:根据预设视差范围内各像素点的匹配代价计算得到各像素在预设视差范围内的多路径代价聚合值;

步骤4.3:根据各像素在预设视差范围内的多路径代价聚合值计算得到各像素在代价聚合后的视差;

步骤4.4:根据各像素在代价聚合后的视差进行视差优化,得到立体校正后的双目摄像头左摄像头视差图;

视差优化包括剔除错误匹配、提高视差精度和抑制噪声;剔除错误匹配采用左右一致性检查法,将左右影像互换位置,再做一次立体匹配,得到另一张视差图,依据视差的唯一性约束,通过左影像的视差图,找到每个像素在右影像的同名点像素及同名点像素对应的视差值、与同名点像素对应的像素在左影像的视差图中的视差值,这两个视差值之差的绝对值若小于1,则满足唯一性约束被保留,反之则不满足唯一性约束而被剔除;同时,采用连通域检测的方法来剔除孤立异常点,去除视差图中由于误匹配造成的小团块,过滤小的孤立的散斑;

$$D_p = \begin{cases} D_{bp}, & \text{if } |D_{bp} - D_{mp}| \leq 1 \\ D_{invalid}, & \text{otherwise} \end{cases};$$

提高视差精度采用子像素优化技术,用二次曲线内插的方法获得子像素精度,对最优视差的代价值以及前后两个视差的代价值进行二次曲线拟合,曲线的极值点所对应的视差值即为新的子像素视差值;

抑制噪声采用卡尔曼滤波,利用卡尔曼增益最优化地修正视差; $\hat{x}(k+1|k)$ 为状态估计

方程,表示在第k时刻,对 $x(k+1)$ 的估计; $W(k+1)$ 为卡尔曼增益, $\hat{e}(k+1)$ 为观测估计方程;

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + W(k+1)\hat{e}(k+1);$$

步骤4.5:对立体校正后的双目摄像头左摄像头视差图进行深度计算,得到立体校正后的双目摄像头左摄像头图像的深度图,并结合步骤3.4中得到的物体识别信息,最终得到经立体校正后的双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标。

2.根据权利要求1所述的基于目标检测及双目视觉的管道缺陷识别与定位方法,其特征在于,所述相机标定的实现方法为:

步骤1.1:制作一块由黑白方块间隔组成的棋盘标定板,用双目摄像头在不同方向下拍摄棋盘标定板的棋盘图像,提取每一张棋盘图像的角点信息,获得棋盘图像上所有内角点的图像坐标及其空间三维坐标;

步骤1.2:建立相机成像的几何模型,确定空间物体表面一点的三维几何位置与其在图像中对应点之间的相互关系;

步骤1.3:将步骤1.1中获得的棋盘图像上所有内角点的图像坐标及其空间三维坐标作为输入,依据相机成像的几何模型,通过实验与计算求解并分别输出双目摄像机左摄像头和右摄像头的内参数矩阵和外参数矩阵;

步骤1.4:根据双目摄像机左摄像头和右摄像头的内参数矩阵和外参数矩阵,利用畸变校正前后的坐标关系,求解5个畸变参数 k_1 、 k_2 、 k_3 、 p_1 、 p_2 以进行畸变矫正。

3.根据权利要求2所述的基于目标检测及双目视觉的管道缺陷识别与定位方法,其特征在于,在步骤1.2中,外参数矩阵 $W = \begin{bmatrix} R & t \\ 0_3^T & 1 \end{bmatrix}$ 反应相机坐标系与世界坐标系的转换, R 为双目摄像机右摄像头相对左摄像头的旋转矩阵, t 为双目摄像机右摄像头相对左摄像头的

平移向量;内参数矩阵 $M = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ 反应像素坐标系与

相机坐标系的转换, f 为透镜焦距, (u_0, v_0) 为图像坐标系原点在像素坐标系下的坐标, d_x 、 d_y 为每个像素点在图像坐标系 x 轴、 y 轴方向的尺寸。

4.根据权利要求2所述的基于目标检测及双目视觉的管道缺陷识别与定位方法,其特征在于,在步骤1.4中,畸变矫正的方法为:

$$\begin{cases} x_{tcorr} = x_p(1 + k_1r^2 + k_2r^4 + k_3r^6) + [2p_1x_p y_p + p_2(r^2 + 2x_p^2)] \\ y_{tcorr} = y_p(1 + k_1r^2 + k_2r^4 + k_3r^6) + [p_1(r^2 + 2y_p^2) + 2p_2x_p y_p] \end{cases};$$

其中, (x_p, y_p) 是图像的原坐标, (x_{tcorr}, y_{tcorr}) 是图像修正后的坐标, r 为变量。

5.根据权利要求3所述的基于目标检测及双目视觉的管道缺陷识别与定位方法,其特征在于,所述立体校正的实现方法为:

步骤2.1:将双目摄像机右摄像头相对左摄像头的旋转矩阵 R 划分为左摄像头的合成旋转矩阵 r_1 和右摄像头的合成旋转矩阵 r_2 ,左摄像头和右摄像头各旋转一半,使得左摄像头和右摄像头的光轴平行,实现左摄像头和右摄像头的成像面共面;

步骤2.2:将左摄像头的合成旋转矩阵 r_1 、右摄像头的合成旋转矩阵 r_2 、原始的左摄像头

和右摄像头内参矩阵、平移向量 t 以及棋盘图像的大小输入OpenCV中,通过cvStereoRectify函数输出左摄像头的行对准校正旋转矩阵 R_1 、右摄像头行对准校正旋转矩阵 R_2 、校正后的左摄像头的内参矩阵 M_1 、校正后的右摄像头的内参矩阵 M_r 、校正后的左摄像头投影矩阵 P_1 、校正后的右摄像头投影矩阵 P_r 以及重投影矩阵 Q ;

步骤2.3:将步骤2.2中cvStereoRectify函数的输出结果作为已知常量,通过左右视图的校正查找映射表,采用逆向映射,对目标图像上的每一个整形的像素位置,查找出其对应源图像上的浮点位置,并对周围源像素的每个整形值进行插值,在校正图像都被赋值之后,对图像进行裁剪,保存校正结果。

6. 根据权利要求1所述的基于目标检测及双目视觉的管道缺陷识别与定位方法,其特征在于,所述匹配代价的计算方法为:

$$e(x_R, y, d) = \min \left\{ \min_{x_R - \frac{1}{2} \leq x \leq x_R + \frac{1}{2}} |I_R(x_R, y) - \tilde{I}_T(x + d, y)|, \min_{x_R - \frac{1}{2} \leq x \leq x_R + \frac{1}{2}} |I_T(x_R + d, y) - \tilde{I}_R(x, y)| \right\};$$

其中, $e(x_R, y, d)$ 表示像素灰度值差值的绝对值, $I_R(x_R, y)$ 表示候选图像像素点 $(x_R - 0.5, y)$ 与 $(x_R + 0.5, y)$ 之间亚像素位置的灰度值, $\tilde{I}_T(x + d, y)$ 表示匹配图像像素点 $(x + d - 0.5, y)$ 与 $(x + d + 0.5, y)$ 之间亚像素位置的灰度值, $I_T(x_R + d, y)$ 表示候选图像像素点 $(x_R + d - 0.5, y)$ 与 $(x_R + d + 0.5, y)$ 之间亚像素位置的灰度值, $\tilde{I}_R(x, y)$ 表示匹配图像像素点 $(x - 0.5, y)$ 与 $(x + 0.5, y)$ 之间亚像素位置的灰度值, x_R 表示候选图像像素的横坐标, y 表示图像像素的纵坐标, d 表示视差, x 表示匹配图像像素的横坐标。

7. 根据权利要求1所述的基于目标检测及双目视觉的管道缺陷识别与定位方法,其特征在于,所述各像素在预设视差范围内的多路径代价聚合值的计算方法为:

计算像素 p 沿着路径 r 的路径代价:

$$L_r(p, d) = C(p, d) + \min \left\{ \begin{array}{l} L_r(p - r, d) \\ L_r(p - r, d - 1) + P_1 \\ L_r(p - r, d + 1) + P_1 \\ \min_i L_r(p - r, i) + P_2 \end{array} \right\} - \min_i L_r(p - r, i);$$

其中, p 代表像素, r 代表路径, d 代表视差, $p - r$ 代表像素 p 路径内的像素, L 表示某一路径聚合代价值; $L_r(p - r, d)$ 表示路径内上一个像素视差为 d 时的代价值, $L_r(p - r, d - 1)$ 表示路径内上一个像素视差为 $d - 1$ 时的代价值, $L_r(p - r, d + 1)$ 表示路径内上一个像素视差为 $d + 1$ 时的代价值, $\min_i L_r(p - r, i)$ 表示路径内上一个像素所有代价值的最小值; $C(p, d)$ 表示初始代价值;

由像素 p 沿着路径 r 的路径代价可得到总路径代价值 S :

$$S(p, d) = \sum_r L_r(p, d)。$$

8. 根据权利要求1所述的基于目标检测及双目视觉的管道缺陷识别与定位方法,其特征在于,所述深度计算的计算方法为:

$$depth = \frac{f \times b}{d + (c_{xr} - c_{xl})};$$

其中, f 为透镜焦距, b 为基线长度, d 为视差, c_{xr} 为右摄像头的主点的列坐标, c_{xl} 为左摄

像头的主点的列坐标。

一种基于目标检测及双目视觉的管道缺陷识别与定位方法

技术领域

[0001] 本发明涉及计算机视觉中的目标检测及多目视觉定位技术领域,特别是指一种基于目标检测及双目视觉的管道缺陷识别与定位方法。

背景技术

[0002] 目标检测是计算机视觉中的重要部分,用于检测数字图像和视频中某一类语义对象,研究重点是复杂环境下的类别检测,其中YOLO在速度方面表现突出,该方法采用单个神经网络直接预测对象边界和类别概率,实现了端到端的类别检测。双目视觉是计算机被动测距方法中重要的距离感知技术,在非接触式测量的前提下,通过对摄像头进行校准、立体匹配等操作后,计算左右图像的视差并得到深度图,以此反映物体在真实世界的坐标位置。

[0003] YOLOX是YOLO算法的最新成果,相比于YOLOV3-V5系列,在识别精度上有提升,速度上也具有一定的竞争优势。在整合各个版本的优点后,通过改进数据增强策略、增加训练的正例样本,来提高目标检测精度,同时加入了无锚点的检测器提高检测速度。将YOLOX网络框架下训练出的模型代入识别网络中,可以实现对多目标、多类别的目标检测,具有检测类别范围广、准确率高、速度快等特点。

[0004] 然而,目前主流的对管道缺陷检测方法,即光电检测方法,包含内窥镜法、激光投射法、CCD摄像机图像采集法等。尽管上述方法采用了图像视觉领域的检测技术,但检测精度较低且未涉及缺陷尺寸的进一步测量,需要耗费一定的人力,同时无法提取管道内表面环的数据,不能在视觉上进一步表示管道内表面的缺陷情况。而计算机视觉方法在管道缺陷分类领域的限制主要是检测精确度较低与数据集的缺乏。

发明内容

[0005] 针对现有管道识别与定位技术在实际检测缺陷目标时,无法高精度检测与定位管道内部缺陷的技术问题,本发明提出了一种基于目标检测及双目视觉的管道缺陷识别与定位方法,将端到端的深度学习算法与双目视觉的位置测量结合,通过自建管道缺陷数据集,训练轻量化深度学习网络,预测缺陷类别,兼顾定位的速度与精度,能够在不同的管道环境下均呈现良好的检测效果,具有全自动、非接触式测量、定位精准且实时性高的特点。

[0006] 本发明的技术方案是这样实现的:

[0007] 一种基于目标检测及双目视觉的管道缺陷识别与定位方法,包括相机标定、立体校正、图像捕获与目标检测、立体匹配与深度计算模块四部分;其步骤如下:

[0008] 步骤一:相机标定:建立相机成像的几何模型,确定空间物体表面上一点的三维几何位置与其在图像中对对应点之间的相互关系,由标定板上特征点的图像坐标和世界坐标求解双目摄像机的标定参数,其中,标定参数包括左摄像头和右摄像头的内参、外参和畸变参数;

[0009] 步骤二:立体校正:通过极线约束进行立体校正,使两幅图像中的对应点在同一条水平极线上,获取校正后的双目摄像机的标定参数;

[0010] 步骤三:图像捕获与目标检测:自建数据集,进行目标检测模型训练,利用双目摄像机捕获图像,预测左摄像头显示的实时画面信息,并输出物体识别信息;

[0011] 步骤四:立体匹配与深度计算模块:将经过立体校正后的左摄像头和右摄像头图像与步骤三中的物体识别信息作为输入量传输至立体匹配与深度计算模块,得到双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标。

[0012] 优选地,所述相机标定的实现方法为:

[0013] 步骤1.1:制作一块由黑白方块间隔组成的棋盘标定板,用双目摄像头在不同方向下拍摄棋盘标定板的棋盘图像,提取每一张棋盘图像的角点信息,获得棋盘图像上所有内角点的图像坐标及其空间三维坐标;

[0014] 步骤1.2:建立相机成像的几何模型,确定空间物体表面一点的三维几何位置与其在图像中对对应点之间的相互关系;

[0015] 步骤1.3:将步骤1.1中获得的棋盘图像上所有内角点的图像坐标及其空间三维坐标作为输入,依据相机成像的几何模型,通过实验与计算求解并分别输出双目摄像机左摄像头和右摄像头的内参数矩阵和外参数矩阵;

[0016] 步骤1.4:根据双目摄像机左摄像头和右摄像头的内参数矩阵和外参数矩阵,利用畸变校正前后的坐标关系,求解5个畸变参数 k_1 、 k_2 、 k_3 、 p_1 、 p_2 以进行畸变矫正。

[0017] 优选地,在步骤1.2中,外参数矩阵 $W = \begin{bmatrix} R & t \\ 0_3^T & 1 \end{bmatrix}$ 反应相机坐标系与世界坐标系的转换, R 为双目摄像机右摄像头相对左摄像头的旋转矩阵, t 为双目摄像机右摄像头相对左

摄像头的平移向量;内参数矩阵 $M = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ 反应像素

坐标系与相机坐标系的转换, f 为透镜焦距, (u_0, v_0) 为图像坐标系原点在像素坐标系下的坐标, d_x 、 d_y 为每个像素点在图像坐标系 x 轴、 y 轴方向的尺寸。

[0018] 优选地,在步骤1.4中,畸变矫正的方法为:

$$[0019] \quad \begin{cases} x_{tcorr} = x_p(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + [2p_1 x_p y_p + p_2(r^2 + 2x_p^2)] \\ y_{tcorr} = y_p(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + [p_1(r^2 + 2y_p^2) + 2p_2 x_p y_p] \end{cases};$$

[0020] 其中, (x_p, y_p) 是图像的原坐标, (x_{tcorr}, y_{tcorr}) 是图像修正后的坐标, r 为变量。

[0021] 优选地,所述立体校正的实现方法为:

[0022] 步骤2.1:将双目摄像机右摄像头相对左摄像头的旋转矩阵 R 划分为左摄像头的合成旋转矩阵 r_1 和右摄像头的合成旋转矩阵 r_2 ,左摄像头和右摄像头各旋转一半,使得左摄像头和右摄像头的光轴平行,实现左摄像头和右摄像头的成像面共面;

[0023] 步骤2.2:将左摄像头的合成旋转矩阵 r_1 、右摄像头的合成旋转矩阵 r_2 、原始的左摄像头和右摄像头内参矩阵、平移向量 t 以及棋盘图像的大小输入OpenCV中,通过cvStereoRectify函数输出左摄像头的行对准校正旋转矩阵 R_1 、右摄像头行对准校正旋转矩阵 R_2 、校正后的左摄像头的内参矩阵 M_1 、校正后的右摄像头的内参矩阵 M_r 、校正后的左摄像头投影矩阵 P_1 、校正后的右摄像头投影矩阵 P_r 以及重投影矩阵 Q ;

[0024] 步骤2.3:将步骤2.2中cvStereoRectify函数的输出结果作为已知常量,通过左右

视图的校正查找映射表,采用逆向映射,对目标图像上的每一个整形的像素位置,查找出其对源图像上的浮点位置,并对周围源像素的每个整形值进行插值,在校正图像都被赋值之后,对图像进行裁剪,保存校正结果。

[0025] 优选地,所述图像捕获与目标检测的实现方法为:

[0026] 步骤3.1:利用双目相机与网络爬虫捕获并筛选管道缺陷图像,并对管道缺陷图像进行标记与分类,创建适应该环境下的数据集;

[0027] 步骤3.2:对管道缺陷图像和标记进行组织目录上的整理,通过getitem方法加载数据集中的管道缺陷图像和标记,使用COCO Evaluator作为评估器,将深度学习模型中涉及的所有内容都放在一个单独的COCO Exp文件中,包括模型设置、训练设置和测试设置;

[0028] 步骤3.3:设置完毕后,使用COCO开源数据集训练完成的预训练权重初始化模型,在命令行输入训练命令以使用GPU开始训练,训练完成后,得到权重文件以及深度学习目标检测模型;

[0029] 步骤3.4:通过双目相机捕获管道实时图像,并将管道实时图像输入深度学习目标检测模型中,输出物体识别信息。

[0030] 优选地,所述图像捕获与目标检测的实现方法为:

[0031] 步骤4.1:根据立体校正后的双目摄像头左摄像头和右摄像头图像计算得到原图像在预设视差范围内各像素点的匹配代价;

[0032] 步骤4.2:根据预设视差范围内各像素点的匹配代价计算得到各像素在预设视差范围内的多路径代价聚合值;

[0033] 步骤4.3:根据各像素在预设视差范围内的多路径代价聚合值计算得到各像素在代价聚合后的视差;

[0034] 步骤4.4:根据各像素在代价聚合后的视差进行视差优化,得到立体校正后的双目摄像头左摄像头视差图;

[0035] 步骤4.5:对立体校正后的双目摄像头左摄像头视差图进行深度计算,得到立体校正后的双目摄像头左摄像头图像的深度图,并结合步骤3.4中得到的物体识别信息,最终得到经立体校正后的双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标。

[0036] 优选地,所述匹配代价的计算方法为:

$$e(x_R, y, d) = \min \left\{ \min_{x_R - \frac{1}{2} \leq x \leq x_R + \frac{1}{2}} |I_R(x, y) - \tilde{I}_T(x + d, y)|, \min_{x_R - \frac{1}{2} \leq x \leq x_R + \frac{1}{2}} |I_T(x_R + d, y) - \tilde{I}_R(x, y)| \right\};$$

[0038] 其中, $e(x_R, y, d)$ 表示像素灰度值差值的绝对值, $I_R(x_R, y)$ 表示候选图像像素点 $(x_R - 0.5, y)$ 与 $(x_R + 0.5, y)$ 之间亚像素位置的灰度值, $\tilde{I}_T(x + d, y)$ 表示匹配图像像素点 $(x + d - 0.5, y)$ 与 $(x + d + 0.5, y)$ 之间亚像素位置的灰度值, $I_T(x_R + d, y)$ 表示候选图像像素点 $(x_R + d - 0.5, y)$ 与 $(x_R + d + 0.5, y)$ 之间亚像素位置的灰度值, $\tilde{I}_R(x, y)$ 表示匹配图像像素点 $(x - 0.5, y)$ 与 $(x + 0.5, y)$ 之间亚像素位置的灰度值, x_R 表示候选图像像素的横坐标, y 表示图像像素的纵坐标, d 表示视差, x 表示匹配图像像素的横坐标。

[0039] 优选地,所述各像素在预设视差范围内的多路径代价聚合值的计算方法为:

[0040] 计算像素p沿着路径r的路径代价:

$$[0041] \quad L_r(p, d) = C(p, d) + \min \left\{ \begin{array}{l} L_r(p-r, d) \\ L_r(p-r, d-1) + P_1 \\ L_r(p-r, d+1) + P_1 \\ \min_i L_r(p-r, i) + P_2 \end{array} \right\} - \min_i L_r(p-r, i);$$

[0042] 其中,p代表像素,r代表路径,d代表视差,p-r代表像素p路径内的像素,L表示某一路径聚合代价值。 $L_r(p-r, d)$ 表示路径内上一个像素视差为d时的代价值, $L_r(p-r, d-1)$ 表示路径内上一个像素视差为d-1时的代价值, $L_r(p-r, d+1)$ 表示路径内上一个像素视差为d+1时的代价值, $\min_i L_r(p-r, i)$ 表示路径内上一个像素所有代价值的最小值; $C(p, d)$ 表示初始代价值;

[0043] 由像素p沿着路径r的路径代价可得到总路径代价值S:

$$[0044] \quad S(p, d) = \sum_r L_r(p, d)。$$

[0045] 优选地,所述深度计算的计算方法为:

$$[0046] \quad depth = \frac{f \times b}{d + (c_{xr} - c_{xl})};$$

[0047] 其中,f为焦距长度,b为基线长度,d为视差, c_{xr} 为右摄像头的主点的列坐标, c_{xl} 为左摄像头的主点的列坐标。

[0048] 与现有技术相比,本发明产生的有益效果为:

[0049] 1) 本发明采用双目RGBD相机对管道缺陷的类别进行预测与定位,相比现有的方法提升了检测的速度与精度,并提出了一种管道缺陷数据集的分类方法,将深度学习应用到双目视觉定位实现目标检测,参考YOLOX网络搭建深度学习卷积神经网络,并做适当调整,简化了特征提取部分并精简网络结构,保证信息传播的同时,对像素进行适当的定位提高预测率。

[0050] 2) 本发明减小了检测的成本,能够对各类管道进行多类别多数量的预测与定位,具有监测范围广,实时性好,精确度高,并能够准确定位等优点;同时,本发明克服了漏磁检测法等传统管道检测方法只能检测铁磁性材料管道的局限性,且不会产生虚假信号,对漏报、误报的情况有一定程度的改善。

附图说明

[0051] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术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。

[0052] 图1为本发明的流程图。

[0053] 图2为本发明提出的深度学习卷积神经网络结构示意图。

[0054] 图3为图2中所用网络的子结构图。

[0055] 图4为本发明提出的管道缺陷分类与标注示例。

[0056] 图5为本发明提出的相机标定流程图。

[0057] 图6为本发明提出的立体校正流程图。

[0058] 图7为本发明提出的立体匹配与深度计算流程图。

具体实施方式

[0059] 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,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付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。

[0060] 如图1所示,本发明实施例公开了一种基于目标检测及双目视觉的管道缺陷识别与定位方法,包括相机标定、立体校正、图像捕获与目标检测、立体匹配与深度计算模块四部分;摄像机标定模块:包括内参、外参系数和畸变系数,可以对之后摄像机拍摄的图像进行校正,得到畸变相对很小的图像。输入为标定板已知特征点的图像坐标和世界坐标,建立摄像机成像的几何模型,确定摄像机图像像素坐标与场景点三维坐标之间的关系,输出为摄像机的内外参数及畸变系数。立体校正模块:采用相交光轴结构的视觉系统,通过分解旋转矩阵和行对准校正旋转矩阵改变两摄像头的相对位置,使两幅图像中的对应点在同一条水平极线上。在摄像头原始数据基础上,利用数学物理方法获取校正后相机参数。将二维搜索变为一维搜索,缩小匹配搜索空间,提高立体匹配的搜索速率。最后通过校正映射查找表完成图像校正,将图像进行裁剪并保存。目标检测模块:输入为立体校正后左摄像头拍摄的RGB三通道图像,通过自建管道缺陷数据集,训练基于目标检测算法的深度学习网络,该网络将图像进行推断、预测,输出为左摄像头图像中识别的缺陷语义标签class、识别框中心坐标(x,y)以及宽高(w,h)。所述语义标签、识别框中心坐标以及宽高数据作为物体识别信息。立体匹配与深度计算模块:以经过立体校正后的双目摄像头左右摄像头图像与经过深度学习物体识别得到的物体识别信息为输入,经立体匹配算法处理后,得到经过立体校正后的双目摄像头左摄像头视差图,再经过深度计算将其转化为深度图,最后结合物体识别信息,输出为经过立体校正后的双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标。具体步骤如下:

[0061] 步骤一:相机标定:对摄像机进行标定,建立相机成像的几何模型,确定空间物体表面上一点的三维几何位置与其在图像中对应点之间的相互关系,由标定板上特征点的图像坐标和世界坐标求解双目摄像机的标定参数,其中,标定参数包括左摄像头和右摄像头的内参、外参和畸变参数;所述双目摄像机摄像头和右摄像头的内参、外参和畸变参数作为摄像机标定参数。内参和外参系数、畸变参数可以对之后摄像机拍摄的图像进行校正,得到畸变相对很小的图像。

[0062] 所述相机标定的实现方法为:

[0063] 步骤1.1:制作一块由黑白方块间隔组成的棋盘标定板,用双目摄像头在不同方向下拍摄棋盘标定板的棋盘图像,使得单平面棋盘格在左、右摄像头中均清晰成像。提取每一张棋盘图像的角点信息,获得棋盘图像上所有内角点的图像坐标及其空间三维坐标;

[0064] 步骤1.2:建立相机成像的几何模型,确定空间物体表面一点的三维几何位置与其在图像中对应点之间的相互关系;

[0065] 在步骤1.2中,外参数矩阵 $\mathbf{W} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix}$ 反应相机坐标系与世界坐标系的转换, \mathbf{R} 为

双目摄像机右摄像头相对左摄像头的旋转矩阵, t 为双目摄像机右摄像头相对左摄像头的

平移向量; 内参数矩阵 $M = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ 反应像素坐标系与

相机坐标系的转换, f 为透镜焦距, (u_0, v_0) 为图像坐标系原点在像素坐标系下的坐标, d_x 、 d_y 为每个像素点在图像坐标系 x 轴、 y 轴方向的尺寸。

[0066] 步骤1.3: 将步骤1.1中获得的棋盘图像上所有内角点的图像坐标及其空间三维坐标作为输入, 依据相机成像的几何模型, 通过实验与计算求解并分别输出双目摄像机左摄像头和右摄像头的内参数矩阵和外参数矩阵;

[0067] 步骤1.4: 根据双目摄像机左摄像头和右摄像头的内参数矩阵和外参数矩阵, 利用畸变校正前后的坐标关系, 求解5个畸变参数 k_1 、 k_2 、 k_3 、 p_1 、 p_2 以进行畸变矫正。

[0068] 在步骤1.4中, 畸变矫正的方法为:

$$[0069] \quad \begin{cases} x_{tcorr} = x_p(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + [2p_1 x_p y_p + p_2(r^2 + 2x_p^2)] \\ y_{tcorr} = y_p(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + [p_1(r^2 + 2y_p^2) + 2p_2 x_p y_p] \end{cases};$$

[0070] 其中, (x_p, y_p) 是图像的原坐标, (x_{tcorr}, y_{tcorr}) 是图像修正后的坐标, r 为变量; (x_{tcorr}, y_{tcorr}) 用 $r=0$ 处的泰勒级数展开近似描述。

[0071] 步骤二: 立体校正: 通过极线约束进行立体校正, 使两幅图像中的对应点在同一条水平极线上, 获取校正后的双目摄像机的标定参数; 调用OpenCV获取校正后左、右摄像头的参数, 完成矫正, 最后通过校正映射获取校正后的图像。

[0072] 所述立体校正的实现方法为:

[0073] 步骤2.1: 将双目摄像机右摄像头相对左摄像头的旋转矩阵 R 划分为左摄像头的合成旋转矩阵 r_1 和右摄像头的合成旋转矩阵 r_2 , 左摄像头和右摄像头各旋转一半, 使得左摄像头和右摄像头的光轴平行, 实现左摄像头和右摄像头的成像面共面;

[0074] 步骤2.2: 将左摄像头的合成旋转矩阵 r_1 、右摄像头的合成旋转矩阵 r_2 、原始的左摄像头和右摄像头内参矩阵、平移向量 t 以及棋盘图像的大小输入OpenCV中, 通过 `cvStereoRectify` 函数输出左摄像头的行对准校正旋转矩阵 R_1 、右摄像头行对准校正旋转矩阵 R_2 、校正后的左摄像头的内参矩阵 M_1 、校正后的右摄像头的内参矩阵 M_r 、校正后的左摄像头投影矩阵 P_1 、校正后的右摄像头投影矩阵 P_r 以及重投影矩阵 Q ;

[0075] 步骤2.3: 将步骤2.2中 `cvStereoRectify` 函数的输出结果作为已知常量, 通过左右视图的校正查找映射表, 采用逆向映射, 对目标图像上的每一个整形的像素位置, 查找出其对应源图像上的浮点位置, 并对周围源像素的每个整形值进行插值, 在校正图像都被赋值之后, 对图像进行裁剪, 保存校正结果。

[0076] 步骤三: 图像捕获与目标检测: 自建数据集, 进行目标检测模型训练, 利用双目摄像机捕获图像, 预测左摄像头显示的实时画面信息, 并输出物体识别信息; 将步骤二校正得到的左相机图像输入目标检测模块, 利用双目相机捕获图像进行标注与分类, 自建管道缺陷数据集。在不同图像细粒度上聚合并形成图像特征的卷积神经网络, 混合和组合图像特征, 并将图像特征传递到预测层, 对图像特征进行预测, 生成边界框并预测类别。输出左相机图像中识别的缺陷语义标签 `class`、识别框中心坐标 (x, y) 以及宽高 (w, h) 并显示目标检

测结果。

[0077] 所述图像捕获与目标检测的实现方法为:

[0078] 步骤3.1:利用双目相机与网络爬虫捕获并筛选管道缺陷图像,并对管道缺陷图像进行标记与分类,创建适应该环境下的数据集;对于标记图像,采用Labelling或者CVAT。缺陷的分类如下,由于ND类为无注释类,将其剔除后,共计9个语义标签。

[0079] (1).ND类(无注释类):无缺陷。(在每个注释类周围有一米的缓冲区,如果注释类之间存在其他任何类则不被标注ND);

[0080] (2).OB类:表面损伤类;

[0081] (3).FO类:障碍物类别;

[0082] (4).RB类:包含裂纹、断裂和塌陷的类;

[0083] (5).RO类:植物的根入侵的类;

[0084] (6).DE类:变形;

[0085] (7).FS类:错位;

[0086] (8).GR类:支管;

[0087] (9).BE类:附着沉积物的类;

[0088] (10).AF类:固定沉积物的类。

[0089] 步骤3.2:对管道缺陷图像和标记进行组织目录上的整理,通过getitem方法加载数据集中的管道缺陷图像和标记,使用COCO Evaluator作为评估器,将深度学习模型中涉及的所有内容都放在一个单独的COCO Exp文件中,包括模型设置、训练设置和测试设置;训练轮数epos=300,在步骤3.1分类的基础上定义缺陷类别数量num_classes=9。

[0090] 步骤3.3:设置完毕后,使用COCO开源数据集训练完成的预训练权重初始化模型,在命令行输入训练命令以使用GPU开始训练,训练完成后,得到权重文件以及深度学习目标检测模型;

[0091] 步骤3.4:通过双目相机捕获管道实时图像,并将管道实时图像输入深度学习目标检测模型中,输出物体识别信息。在深度学习目标检测模型中,在不同图像细粒度上聚合并形成图像特征的卷积神经网络,混合和组合图像特征,并将图像特征传递到预测层,对图像特征进行预测,生成边界框并预测类别,输出物体识别信息。

[0092] 步骤四:立体匹配与深度计算模块:将经过立体校正后的左摄像头和右摄像头图像与步骤三中的物体识别信息作为输入量传输至立体匹配与深度计算模块,以物体识别信息中的识别框中心坐标为参考量,经立体匹配与深度计算模块对图像处理后,输出得到双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标。

[0093] 所述图像捕获与目标检测的实现方法为:

[0094] 步骤4.1:根据立体校正后的双目摄像头左摄像头和右摄像头图像计算得到原图像在预设视差范围内各像素点的匹配代价;匹配代价计算的目的是衡量待匹配像素与候选像素之间的相关性。两个像素无论是否为同名点,都可以通过匹配代价函数计算匹配代价,代价越小则说明相关性越大,是同名点的概率也越大。

[0095] 使用BT(Birchfield and Tomasi)算法作为匹配代价的计算方法,表示为:

$$[0096] \quad e(x_R, y, d) = \min \left\{ \min_{x_R - \frac{1}{2} \leq x \leq x_R + \frac{1}{2}} |I_R(x_R, y) - \tilde{I}_T(x + d, y)|, \min_{x_R - \frac{1}{2} \leq x \leq x_R + \frac{1}{2}} |I_T(x_R + d, y) - \tilde{I}_R(x, y)| \right\};$$

[0097] 其中, $e(x_R, y, d)$ 表示BT代价, 即像素灰度值差值的绝对值, $I_R(x_R, y)$ 表示候选图像像素点 $(x_R - 0.5, y)$ 与 $(x_R + 0.5, y)$ 之间亚像素位置的灰度值, $\tilde{I}_T(x + d, y)$ 表示匹配图像像素点 $(x + d - 0.5, y)$ 与 $(x + d + 0.5, y)$ 之间亚像素位置的灰度值, $I_T(x_R + d, y)$ 表示候选图像像素点 $(x_R + d - 0.5, y)$ 与 $(x_R + d + 0.5, y)$ 之间亚像素位置的灰度值, $\tilde{I}_R(x, y)$ 表示匹配图像像素点 $(x - 0.5, y)$ 与 $(x + 0.5, y)$ 之间亚像素位置的灰度值, x_R 表示候选图像 (左图像) 像素的横坐标, y 表示图像像素的纵坐标, d 表示视差, x 表示匹配图像 (右图像) 像素的横坐标。通过上述公式可实现对经过立体校正后的双目摄像头左右摄像头图像在预设视差范围内进行匹配代价计算, 得到原图像在预设视差范围内各像素点的匹配代价。

[0098] 步骤4.2: 根据预设视差范围内各像素点的匹配代价计算得到各像素在预设视差范围内的多路径代价聚合值; 采用全局立体匹配算法的思路, 即全局能量最优化策略, 简单来说就是寻找每个像素的最优视差使得整张影像的全局能量函数最小, 全局能量函数的定义如下式:

$$[0099] \quad E(d) = E_{\text{data}}(d) + E_{\text{smooth}}(d);$$

[0100] 采用路径代价聚合的方法, 即将像素所有视差下的匹配代价进行像素周围所有路径上的一维聚合得到路径下的路径代价值, 然后将所有路径代价值相加得到该像素聚合后的匹配代价值。计算像素 p 沿着路径 r 的路径代价:

$$[0101] \quad L_r(p, d) = C(p, d) + \min \left\{ \begin{array}{l} L_r(p - r, d) \\ L_r(p - r, d - 1) + P_1 \\ L_r(p - r, d + 1) + P_1 \\ \min_i L_r(p - r, i) + P_2 \end{array} \right\} - \min_i L_r(p - r, i);$$

[0102] 其中, p 代表像素, r 代表路径, d 代表视差, $p - r$ 代表像素 p 路径内的像素, L 表示某一路径聚合代价值。 $L_r(p - r, d)$ 表示路径内上一个像素视差为 d 时的代价值, $L_r(p - r, d - 1)$ 表示路径内上一个像素视差为 $d - 1$ 时的代价值, $L_r(p - r, d + 1)$ 表示路径内上一个像素视差为 $d + 1$ 时的代价值, $\min_i L_r(p - r, i)$ 表示路径内上一个像素所有代价值的最小值; $C(p, d)$ 表示初始代价值。

[0103] 第一项为匹配代价值 C , 属于数据项。

[0104] 第二项是平滑项, 累加到路径代价上的值取, 不做惩罚、做 P_1 惩罚和做 P_2 惩罚三种情况中代价最小的值; P_1 是为了适应倾斜或弯曲的表面, P_2 则是为了保留不连续性。 P_2 往往是根据相邻像素的灰度差来动态调整, 如下式:

$$[0105] \quad P_2 = \frac{P_2'}{|I_{bp} - I_{bq}|} \text{ as } P_2 > P_1;$$

[0106] P_2' 为 P_2 的初始值, 一般设置为远大于 P_1 的数, I_{bp} 和 I_{bq} 分别表示像素 p 和 q 的灰度值。

[0107] 第三项是为了保证新的路径代价值 L_r 不超过一定数值上限。

[0108] 由像素 p 沿着路径 r 的路径代价可得到总路径代价值 S :

[0109] $S(p, d) = \sum_r L_r(p, d)$;

[0110] 通过步骤4.2中各式计算可实现多路径的代价聚合,得到各像素在预设视差范围内的多路径代价聚合值。

[0111] 步骤4.3:根据各像素在预设视差范围内的多路径代价聚合值计算得到各像素在代价聚合后的视差;视差计算即通过代价聚合之后的代价矩阵S来确定每个像素的最优视差值,采用WTA(Winner-take-all)算法,即在某个像素的所有视差下的代价值中,选择最小代价值所对应的视差作为最优视差,最终得到各像素在代价聚合后的视差。

[0112] 步骤4.4:根据各像素在代价聚合后的视差进行视差优化,得到立体校正后的双目摄像头左摄像头视差图;视差优化的目的是对上一步得到的视差图进行进一步优化,改善视差图的质量,包括剔除错误匹配、提高视差精度和抑制噪声。

[0113] 剔除错误匹配采用左右一致性检查法,它基于视差的唯一性约束,即每个像素最多只存在一个正确视差。具体步骤是将左右影像互换位置,即左影像成为右影像,右影像成为左影像,再做一次立体匹配,得到另一张视差图,因为视差图中每个值所反映的是两个像素之间的对应关系,所以依据视差的唯一性约束,通过左影像的视差图,找到每个像素在右影像的同名点像素及该像素对应的视差值,这两个视差值之差的绝对值若小于1,则满足唯一性约束被保留,反之则不满足唯一性约束而被剔除。同时,采用连通域检测的方法来剔除孤立异常点,去除视差图中由于误匹配造成的小团块,过滤小的孤立的散斑。一致性检查的公式如下式:

$$[0114] \quad D_p = \begin{cases} D_{bp}, & \text{if } |D_{bp} - D_{mp}| \leq 1 \\ D_{invalid}, & \text{otherwise} \end{cases};$$

[0115] 提高视差精度采用子像素优化技术,用二次曲线内插的方法获得子像素精度,对最优视差的代价值以及前后两个视差的代价值进行二次曲线拟合,曲线的极值点所对应的视差值即为新的子像素视差值。

[0116] 抑制噪声采用卡尔曼滤波,使视差结果更加平滑,一定程度上消除视差图中的噪声,同时起到视差填充的作用。通过步骤4.4可实现视差优化,最终得到经过立体校正后的双目摄像头左摄像头视差图。

[0117] 步骤4.5:对立体校正后的双目摄像头左摄像头视差图进行深度计算,得到立体校正后的双目摄像头左摄像头图像的深度图,并结合步骤3.4中得到的物体识别信息,最终得到经立体校正后的双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标。

[0118] 所述深度计算的计算方法为:

$$[0119] \quad depth = \frac{f \times b}{d + (c_{xr} - c_{xl})};$$

[0120] 其中,f为焦距长度,b为基线长度,d为视差, c_{xr} 为右摄像头的主点的列坐标, c_{xl} 为左摄像头的主点的列坐标。经过深度计算后,可得到经过立体校正后的双目摄像头左摄像头图像的深度图,结合经过深度学习目标检测得到的物体识别信息,可最终得到经过立体校正后的双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标。

[0121] 如图2、图3所示,图2为本发明提出的深度学习卷积神经网络结构,由特征提取、特征融合、预测类别并标定检测框共计三部分构成。图3为该网络中包含的重要子结构。

[0122] 特征提取部分,其作用是在不同图像细粒度上聚合并形成图像特征,包括Focus隔行采样拼接结构、四个DarkNet结构。Focus作为整个网络的输入部分,由默认的 $3 \times 640 \times 640$ 输入开始,复制四份,通过切片操作将四个图片切成了四个 $3 \times 320 \times 320$ 的切片,接下来使用concat从深度上拼接这四个切片,生成 $12 \times 320 \times 320$ 的输出,最后经过batch_born和leaky_relu将结果输入到下一个卷积层。DarkNet结构的基本架构由CSP与SPP组成,将原输入分成两个分支,分别进行卷积操作使得通道数减半,分支一抛弃了YOLOV5的Bottleneck操作,采用ResUnit结构,随后concat分支一和分支二,从而使得CSP的输入与输出是一样的大小,CSP的目的是为了让模型学习到更多的特征;SPP的输入为 $512 \times 20 \times 20$,经过 1×1 的卷积层后输出 $256 \times 20 \times 20$,然后经过并列的三个MaxPool进行采样,将结果与其初始特征相加,输出 $1024 \times 20 \times 20$,最后用512的卷积核将其恢复到 $512 \times 20 \times 20$ 。

[0123] 特征融合部分的作用是混合和组合图像特征,并将图像特征传递到预测层,采用PAFPN,该方法借鉴于YOLOV3,基于Mask R-CNN和FPN框架,加强了信息传播,具有准确保留空间信息的能力,这有助于对像素进行适当的定位以形成掩模。

[0124] 最后一部分网络结构对图像特征进行预测,生成边界框并预测类别。首先采用Decoupled Head提取三个特征层进行目标检测,同时加入了无锚点的检测器提高检测速度,在输出部分,三个输出分别是 $85 \times 20 \times 20$ 、 $85 \times 40 \times 40$ 、 $85 \times 80 \times 80$ 。通过对预测结果解码,即得到缺陷预测信息。

[0125] 本发明提出的深度学习卷积神经网络结构参考YOLOX网络搭建,相比YOLOX网络在特征提取部分减少了部分CBS操作和卷积次数,简化了该网络的特征提取部分。同时,修改了损失函数的计算方式,使用标准二进制交叉熵损失进行训练。

$$[0126] \quad L(x, y) = \frac{1}{C} \sum_c -[w_c y_c \log(\sigma(x_c)) + (1 - y_c) \log(1 - \sigma(x_c))]$$

[0127] 其中,C为上文提到的语义标签数 $\text{num}_{\text{classes}} = 9$, $y_c = \begin{cases} 0, & \text{class } c \text{ exists} \\ 1, & \text{no class } c \text{ exists} \end{cases}$ 表示当前图像中是否存在c类, x_c 是c类模型的原始输出, σ 为Sigmoid激活函数, $w_c = \frac{N - N_c}{N_c}$ 为c类的权重,N表示训练中的图像数, N_c 表示包含c类的图像数。

[0128] 如图4所示,本发明提出了一种管道缺陷分类方法,将缺陷类型分为10类,排除ND类,即无缺陷类后,共计9个语义标签,分别是表面损伤类、障碍物类、裂纹类、根侵入类、变形类、错位类、支管类、附着沉积物类和固定沉积物类。同时,图中给出了9类缺陷的标注示例,可作为自建管道缺陷数据集时的参考。通过管道缺陷分类方法解决了计算机视觉方法在管道缺陷分类领域数据分类标准模糊的问题。

[0129] 如图5所示,本发明公开的相机标定方法,使用了摄像机标定模块。摄像机标定模块的输入为标定板已知特征点的图像坐标和世界坐标,由此建立摄像机成像的几何模型,确定摄像机图像像素坐标与场景点三维坐标之间的关系,输出为摄像机的内外参数矩阵及畸变系数。

[0130] 外参数矩阵W,反应相机坐标系与世界坐标系的转换;内参数矩阵M,反应像素坐标系与相机坐标系的转换。

[0131] 外参数矩阵W、内参数矩阵M以及5个畸变系数 k_1 、 k_2 、 k_3 、 p_1 、 p_2 如下所示。

$$[0132] \quad W = \begin{bmatrix} R & t \\ 0_3^T & 1 \end{bmatrix}$$

$$[0133] \quad M = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$[0134] \quad \begin{cases} x_{tcorr} = x_p(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + [2p_1 x_p y_p + p_2(r^2 + 2x_p^2)] \\ y_{tcorr} = y_p(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + [p_1(r^2 + 2y_p^2) + 2p_2 x_p y_p] \end{cases}$$

[0135] 其中,R为双目摄像机右摄像头相对做摄像头的旋转矩阵,t为双目摄像机右摄像头相对做摄像头的平移向量。f为透镜焦距,(u_0, v_0)为图像坐标系原点在像素坐标系下的坐标, d_x, d_y 为每个像素点在图像坐标系x轴、y轴方向的尺寸。(x_p, y_p)是图像的原坐标, (x_{tcorr}, y_{tcorr})是图像修正后的坐标,用 $r=0$ 处的Taylor级数展开近似描述。

[0136] 相机标定方法通过Matlab实现,应尽可能采集各种不同的棋盘图像。镜头的畸变从图像中心径向增加,并且有时在图像各帧上表现出不均匀现象,为了获取图像的畸变信息,棋盘应当处在图像的各种不同边缘处。通过Matlab标定后可得到旋转矩阵R和内参矩阵M,计算后即得到余下的参数值。为后续导入OpenCV使用,应将旋转矩阵R和内参矩阵M转置。畸变向量中畸变系数的次序为 $[k_1, k_2, p_1, p_2, k_3]$ 。

[0137] 如图6所示,本发明公开的立体校正方法,使用了立体校正模块。通过极线约束进行立体校正,将标定模块输出的旋转矩阵R划分为左右摄像头的合成旋转矩阵 r_1 和 r_2 两部分,使两幅图像中的对应点在同一条水平极线上。获取校正后的相机参数,包括合成旋转矩阵 r_1 和 r_2 、原始的左右摄像头内参矩阵、平移向量t以及棋盘图像的大小。调用OpenCV中StereoRectify与initUndistortRectifyMap函数输出左右摄像头行对准校正旋转矩阵 R_1 和 R_2 、校正后左右摄像头内参矩阵 M_1 和 M_r 、校正后左右摄像头投影矩阵 P_1 和 P_r 以及重投影矩阵Q。

[0138] 重投影矩阵Q实现了世界坐标系和像素坐标系之间的转换。具体如下:

$$[0139] \quad Q \begin{bmatrix} x \\ y \\ d \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$$

[0140] 校正后相机坐标系的原点在光轴上。其中d表示视差,三维坐标为 $(\frac{X}{W}, \frac{Y}{W}, \frac{Z}{W})$, c'_x 表示右图像的主点,当校正完成时, $c'_x = c_x$ 。在立体校正正确的前提下展开上式,如下所示:

$$[0141] \quad \left(\frac{X}{Y}, \frac{Y}{W}, \frac{Z}{W} \right) = \left(-\frac{x - c_x}{d} T_x, -\frac{y - c_y}{d} T_x, -\frac{f}{d} T_x \right)$$

[0142] 实时获取校正后左右摄像头的参数,通过校正映射获取校正后的图像。对目标图像上的每一个整形的像素位置,查找出其对对应源图像上的浮点位置,并对周围源像素的每个整形值进行插值。在校正图像都被赋值之后,对图像进行裁剪,保存校正结果。最终,双目摄像头两图像平面平行,光轴和图像平面垂直,极点处于无限远处,此时,图像中任意一点(x_0, y_0)对应的极线为 $y = y_0$ 。

[0143] 如图7所示,本发明公开的立体匹配与深度计算方法,使用了立体匹配模块。包括基于BT算法进行匹配代价计算、基于全局能量最优化策略进行代价聚合、基于WTA算法进行视差计算、视差优化以及深度计算。

[0144] 其中,匹配代价计算如下所示。BT的代价即像素灰度值差值的绝对值,同时,BT利用了亚像素的灰度信息。首先计算左图像中像素 $(x_R-0.5, y)$ 和 $(x_R+0.5, y)$ 之间亚像素位置 (x_R+x, y) 的灰度值 $\tilde{I}_R(x_R, y)$,同理可计算右图像中像素 $(x_R+d-0.5, y)$ 和 $(x_R+d+0.5, y)$ 之间亚像素位置 (x_R+d+x, y) 的灰度值 $\tilde{I}_T(x+d, y)$ 。分别计算两个代价,最终的代价为两个代价的最小值。

$$[0145] \quad e(x_R, y, d) = \min \left\{ \min_{x_R-\frac{1}{2} \leq x \leq x_R+\frac{1}{2}} |I_R(x_R, y) - \tilde{I}_T(x+d, y)|, \min_{x_R-\frac{1}{2} \leq x \leq x_R+\frac{1}{2}} |I_T(x_R+d, y) - \tilde{I}_R(x, y)| \right\}$$

[0146] 两图像的代价如下:

$$[0147] \quad \cos_1 = \min_{x_R-\frac{1}{2} \leq x \leq x_R+\frac{1}{2}} |I_R(x_R, y) - \tilde{I}_T(x+d, y)|$$

$$[0148] \quad \cos_2 = \min_{x_R-\frac{1}{2} \leq x \leq x_R+\frac{1}{2}} |I_T(x_R+d, y) - \tilde{I}_R(x, y)|$$

[0149] 代价聚合采用全局立体匹配算法的思路,即全局能量最优化策略,寻找每个像素的最优视差使得整张影像的全局能量函数最小。将像素所有视差下的匹配代价进行像素周围所有路径上的一维聚合得到路径下的路径代价值,然后将所有路径代价值相加得到该像素聚合后的匹配代价值。

[0150] 匹配代价值共分为三项,第一项为匹配代价值C,属于数据项;第二项为平滑项,累加到路径代价上的值取,不做惩罚、做 P_1 惩罚和做 P_2 惩罚三种情况中代价最小的值;第三项为缩小项,保证了新的路径代价值 L_r 不超过一定数值上限。

$$[0151] \quad L_r(p, d) = C(p, d) + \min \left\{ \begin{array}{l} L_r(p-r, d) \\ L_r(p-r, d-1) + P_1 \\ L_r(p-r, d+1) + P_1 \\ \min_i L_r(p-r, i) + P_2 \end{array} \right\} - \min_i L_r(p-r, i)$$

$$[0152] \quad P_2 = \frac{P_2'}{|I_{bp} - I_{bq}|} \text{ as } P_2 > P_1$$

[0153] 得到 $L_r(p, d)$ 后,总路径代价值S如下:

$$[0154] \quad S(p, d) = \sum_r L_r(p, d)$$

[0155] 视差计算采用基于WTA的贪心算法,即在某个像素的所有视差下的代价值中,选择最小代价值所对应的视差作为最优视差,最终得到各像素在代价聚合后的视差。

$$[0156] \quad \text{Winner}(p) = \min_d \{S(p, d)\}$$

[0157] 视差优化包括剔除错误匹配、提高视差精度和抑制噪声。剔除错误匹配采用左右一致性检查法,将左右影像互换位置,再做一次立体匹配,得到另一张视差图,依据视差的唯一性约束,通过左影像的视差图,找到每个像素在右影像的同名点像素及该像素对应的

视差值,这两个视差值之差的绝对值若小于1,则满足唯一性约束被保留,反之则不满足唯一性约束而被剔除。同时,采用连通域检测的方法来剔除孤立异常点,去除视差图中由于误匹配造成的小团块,过滤小的孤立的散斑。

$$[0158] \quad D_p = \begin{cases} D_{bp}, & \text{if } |D_{bp} - D_{mp}| \leq 1 \\ D_{invalid}, & \text{otherwise} \end{cases}$$

[0159] 提高视差精度采用子像素优化技术,用二次曲线内插的方法获得子像素精度,对最优视差的代价值以及前后两个视差的代价值进行二次曲线拟合,曲线的极值点所对应的视差值即为新的子像素视差值。

[0160] 抑制噪声采用卡尔曼滤波,利用卡尔曼增益最优化地修正视差,使视差结果更加平滑,一定程度上消除视差图中的噪声,同时起到视差填充的作用。 $\hat{x}(k+1|k)$ 为状态估计方程,表示在第k时刻,对x(k+1)的估计;W(k+1)为卡尔曼增益, $\hat{e}(k+1)$ 为观测估计方程。

$$[0161] \quad \hat{x}(k+1|k+1) = \hat{x}(k+1|k) + W(k+1)\hat{e}(k+1)$$

[0162] 像素深度计算公式如下。其中f为焦距长度,b为基线长度,d为视差, c_{xr} 与 c_{xl} 为两个相机主点的列坐标。

$$[0163] \quad depth = \frac{f \times b}{d + (c_{xr} - c_{xl})}$$

[0164] 综上所述,通过本发明公开的一种基于目标检测及双目视觉的管道缺陷识别与定位方即可最终得到经过立体校正后的双目摄像头左摄像头图像中被识别物体映射到实际三维空间中的空间三维坐标,完成了目标检测与定位。

[0165] 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,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。



图1

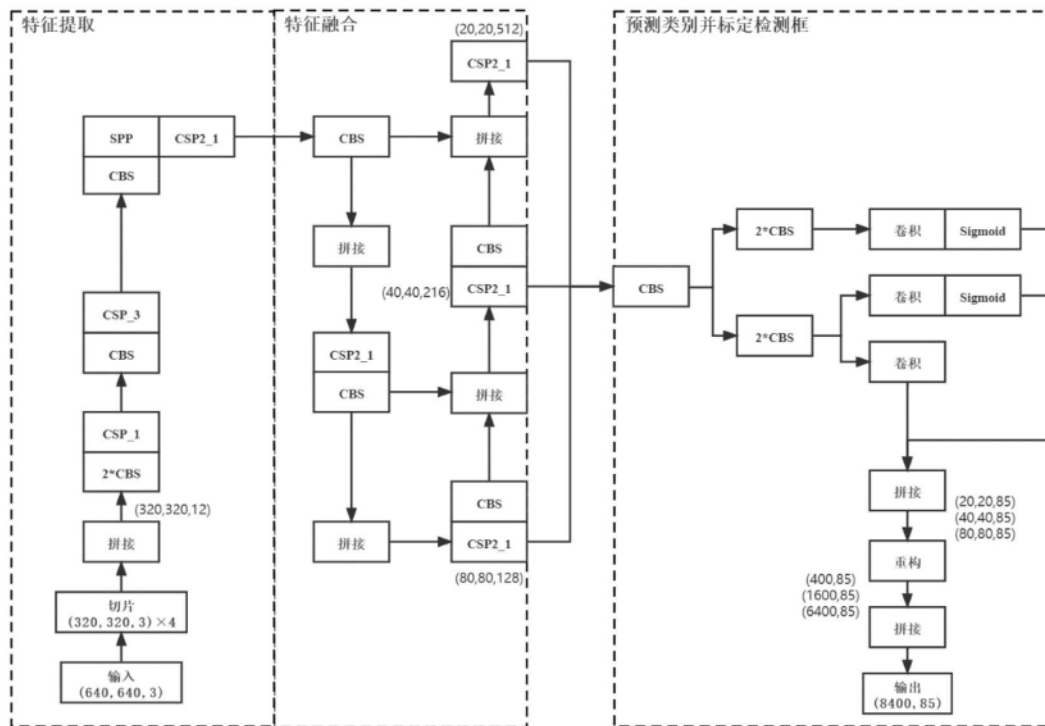


图2

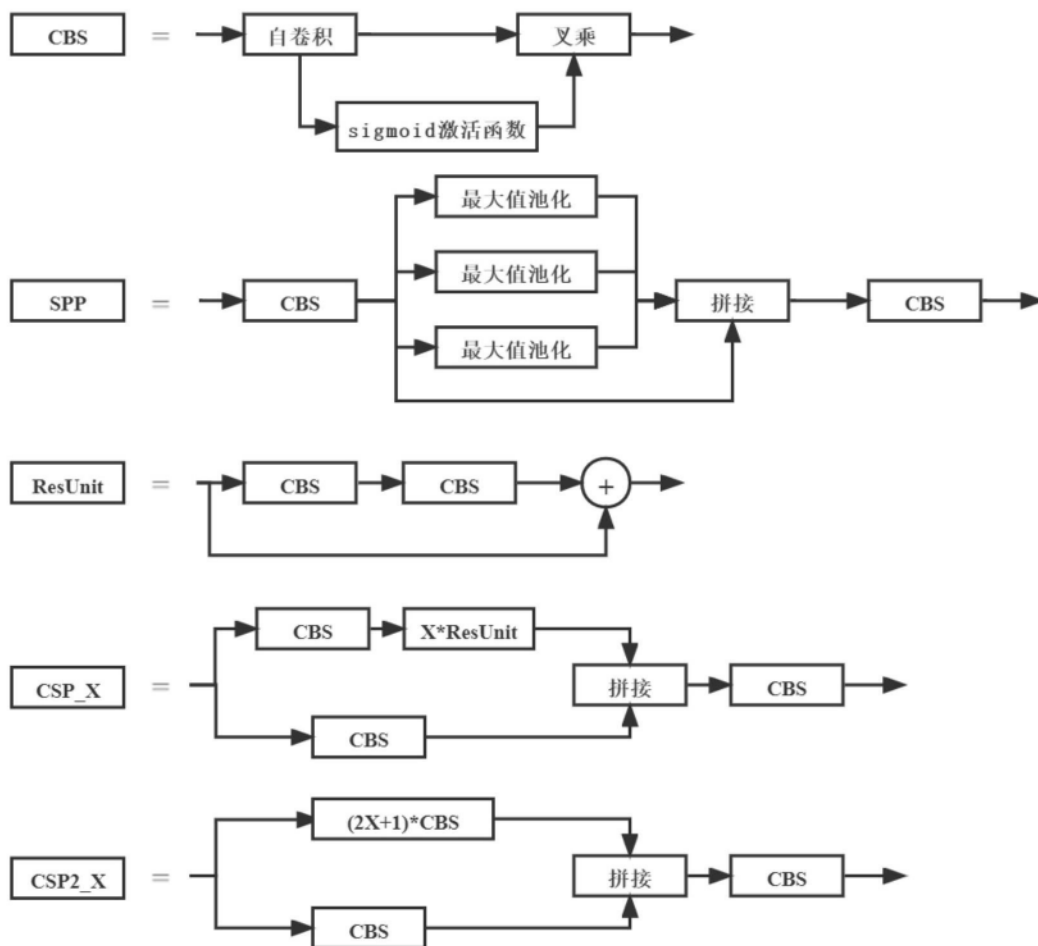


图3

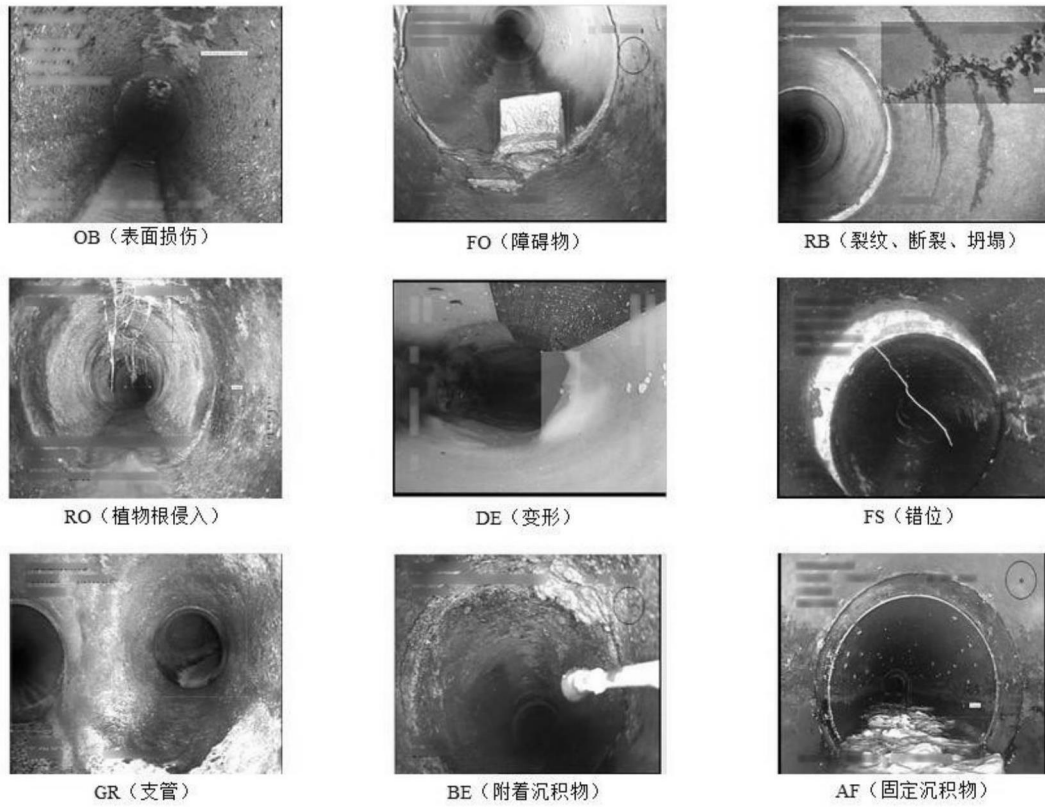


图4

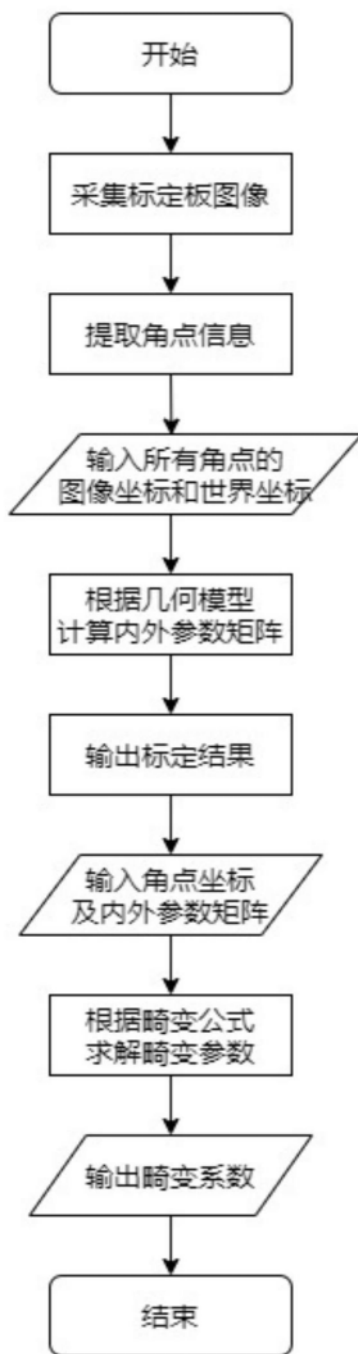


图5

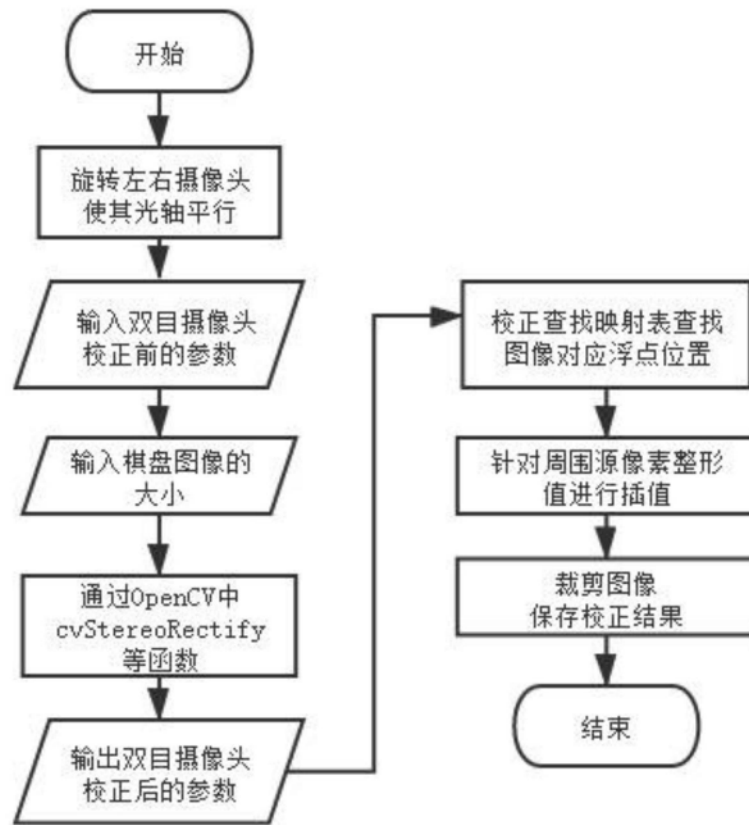


图6

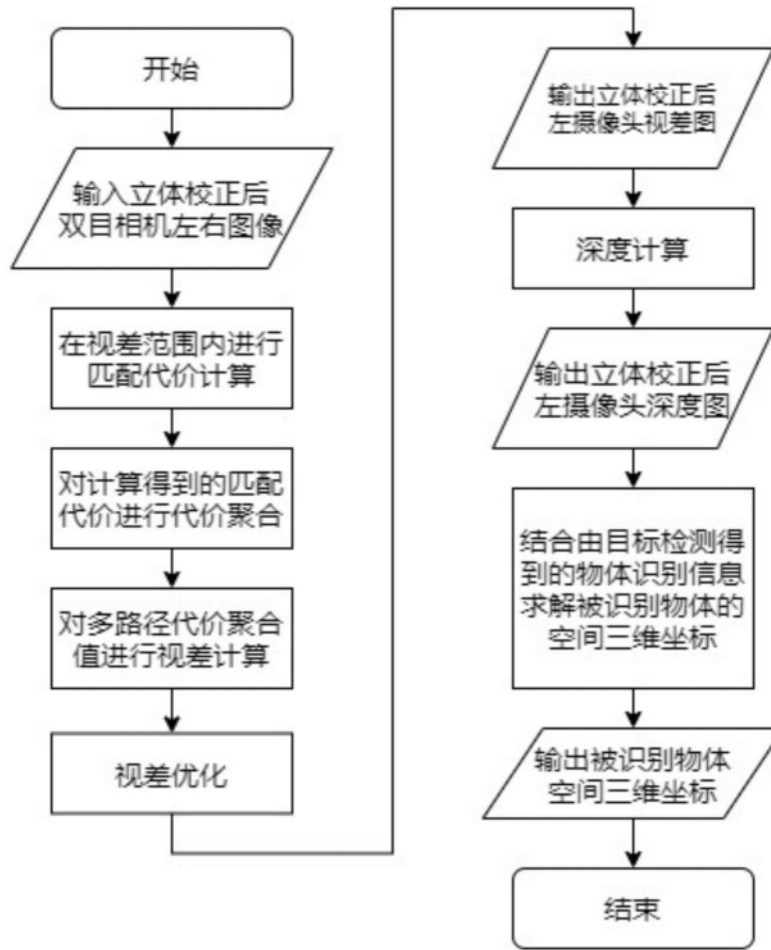


图7