



(12)发明专利申请

(10)申请公布号 CN 107844750 A

(43)申请公布日 2018.03.27

(21)申请号 201710979114.9

(22)申请日 2017.10.19

(71)申请人 华中科技大学

地址 430074 湖北省武汉市洪山区珞喻路  
1037号

(72)发明人 曹治国 杨健 肖阳 宫凯程  
朱昂帆 赵晨

(74)专利代理机构 华中科技大学专利中心  
42201

代理人 廖盈春 李智

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/32(2006.01)

G06K 9/34(2006.01)

G06K 9/62(2006.01)

权利要求书3页 说明书9页 附图2页

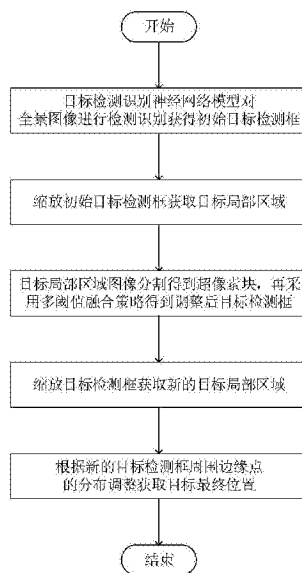
(54)发明名称

一种水面全景图像目标检测识别方法

(57)摘要

本发明公开了一种水面全景图像目标检测识别方法,属于计算机视觉识别技术领域。本发明方法首先利用目标检测识别神经网络模型对全景图像进行目标检测识别,获取目标类别和检测框初始位置;之后对目标局部区域进行图像分割,得到多个超像素块,再采用多阈值融合策略合并所述超像素块,得到调整后目标检测框;再计算新的目标局部区域的边缘图,根据目标检测框 $box^*$ 周围边缘点的分布调整 $box^*$ ,得到最终目标检测框 $box^*$ ;最后由全景图像成像平面坐标系与世界坐标系的映射关系,将最终目标检测框 $box^*$ 位置转化为目标的实际位置。本发明方法能有效克服全景图像中畸变效应,从全景图像中将目标位置准确识别出来。

CN 107844750 A



1. 一种水面全景图像目标检测识别方法,其特征在於,所述方法包括:

(1) 利用训练好的目标检测识别神经网络模型对全景图像进行目标检测识别,获取目标类别和初始目标检测框 $box$ ;

(2) 在全景图像上截取以目标检测框 $box$ 为中心,  $\gamma$  倍目标检测框 $box$ 大小的矩形区域作为目标局部区域;

(3) 对目标局部区域进行图像分割,得到多个超像素块,再采用多阈值融合策略合并所述超像素块,得到调整后目标检测框 $box^*$ ;

(4) 在全景图像上截取以目标检测框 $box^*$ 为中心,  $\gamma$  倍目标检测框 $box^*$ 大小的矩形区域作为新的目标局部区域;

(5) 计算新的目标局部区域的边缘图,根据目标检测框 $box^*$ 周围边缘点的分布调整 $box^*$ ,得到最终目标检测框 $box^*$ ;

(6) 由全景图像成像平面坐标系与世界坐标系的映射关系,将最终目标检测框 $box^*$ 位置转化为目标的实际位置。

2. 根据权利要求1所述的一种目标位置检测方法,其特征在於,所述步骤(1)具体包括:

(11) 对全景图像进行尺寸归一化,之后对归一化后图像去均值;

(12) 利用特征提取网络计算全景图像的卷积特征,以最后一层卷积特征图为输入构建目标潜在区域生成网络,由目标潜在区域生成网络生成目标潜在区域;

(13) 对目标潜在区域进行RoI Pooling归一化至固定大小,之后将其输入到预测网络中,得到多个目标检测框;

(14) 采用非极大值抑制抑制算法,剔除多余的目标检测框,得到初始目标检测框 $box$ 。

3. 根据权利要求1所述的一种目标位置检测方法,其特征在於,所述步骤(3)具体包括:

(31) 计算目标局部区域中所有像素点之间的不相似度 $e$ :

$$e = \sqrt{(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2},$$

其中, $r$ 、 $g$ 和 $b$ 分别代表像素点在RGB三通道的值;定义像素点合并阈值 $\theta$ ;若两个像素点之间的不相似度满足 $e < \theta$ ,则将两像素点合并成一个区域 $C$ ;

在区域 $C$ 中,最大的不相似度作为类内差异 $Int(C)$ :

$$Int(C) = \max e,$$

将两区域 $C_1$ 和 $C_2$ 间最小的不相似度作为类间差异 $Diff(C_1, C_2)$ :

$$Diff(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2} e(v_i, v_j),$$

其中, $v_i$ 和 $v_j$ 分别表示区域 $C_1$ 和 $C_2$ 中的像素点;

若区域 $C_1$ 和 $C_2$ 的类间差异同时小于各自的类内差异,则将区域 $C_1$ 和 $C_2$ 合并;对目标局部区域中所有区域进行迭代融合,直到不满足融合条件,最终得到的区域作为超像素块;

(32) 为每一个超像素块构建一个最小外接矩形 $s$ ,计算超像素块和初始目标检测框 $box$ 的重合率:

$$IoU(box, s) = \frac{|box \cap s|}{|box \cup s|},$$

构建两个集合:

$$S_{inner} = \{s \in S_0 \mid \text{IoU}(s, \text{box}) = 1\},$$

$$S_{inter} = \{s \in S_0 \mid 0 < \text{IoU}(s, \text{box}) < 1\},$$

其中,  $S_0$ 表示和目标检测框 $\text{box}$ 有交集的超像素块集合;  $S_{inner}$ 表示被目标检测框 $\text{box}$ 包含的超像素块集合;  $S_{inter}$ 表示和目标检测框 $\text{box}$ 部分相交的超像素块集合;

将集合 $S_{inner}$ 区域中所有超像素块融合, 获取融合后区域的最小外接矩形 $\text{box}^*$ ;

(33) 遍历集合 $S_{inter}$ , 若满足:

$$\text{IoU}(\text{box}^* \cup s_i, \text{box}^*) > \delta_1, s_i \in S_{inter},$$

则将满足上式的 $s_i$ 从集合 $S_{inter}$ 中剔除得到新的集合 $S_{inter}$ , 用 $\text{box}^* \cup s_i$ 更新 $\text{box}^*$ ; 其中 $\delta_1$ 表示第一超像素块融合阈值;

(34) 将第二超像素块融合阈值 $\delta_2$ 更换 $\delta_1$ , 代入新的 $S_{inter}$ 和 $\text{box}^*$ 重复步骤(33), 得到新的集合 $S_{inter}$ 和 $\text{box}^*$ ; 再次用第三超像素块融合阈值 $\delta_3$ 更换 $\delta_2$ , 代入新的 $S_{inter}$ 和 $\text{box}^*$ 重复步骤(33), 得到新的 $\text{box}^*$ ,  $\text{box}^*$ 既为调整后的目标检测框。

4. 根据权利要求1所述的一种目标位置检测方法, 其特征在于, 所述步骤(5)具体包括:

(51) 将新的目标局部区域转换为灰度图像, 在利用下式以滑动窗口的方式对图像进行平滑去噪处理:

$$G = \frac{1}{2\pi\delta^2} e^{-\frac{x^2+y^2}{2\delta^2}},$$

其中,  $x$ 和 $y$ 为目标局部区域图像中每一个像素的坐标值,  $\delta$ 为高斯函数模板;

(52) 对图像的 $x$ 方向和 $y$ 方向分别利用一阶有限差分来模拟偏导, 其中差分计算模板为:

$$S_x = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix},$$

$$S_y = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix},$$

利用该模板通过卷积的方式计算图像中每一个元素 $f(i, j)$ 的梯度幅值 $M(i, j)$ 和方向 $\theta(i, j)$ :

$$M(i, j) = \sqrt{(f(i, j) * S_x(i, j))^2 + (f(i, j) * S_y(i, j))^2},$$

$\theta(i, j) = \arctan((f(i, j) * S_y(i, j)) / (f(i, j) * S_x(i, j)))$ , 利用非极大值抑制算法得到局部梯度最大的点作为目标潜在边缘点, 潜在的边缘点灰度值置为非零, 其余像素点灰度值全部置零, 得到目标局部区域的二值化图像;

(53) 设置两个边缘检测阈值 $\beta_1$ 、 $\beta_2$ , 其中 $\beta_1 = 0.4\beta_2$ , 将二值化图像中梯度值小于 $\beta_1$ 的像素灰度值设为零, 得到目标边缘图像 $a$ , 同时将二值化图像中梯度值小于 $\beta_2$ 的像素灰度值设为零, 得到目标边缘图像 $b$ ; 再通过目标边缘图 $a$ 和 $b$ 间的相互补充, 得到最终目标边缘图;

(54) 计算最终目标边缘图中边缘点和目标检测框 $\text{box}^*$ 四条边框上的像素点之间的距离, 找出距离边缘点最近的像素点, 用所述像素点组成图像的最小外接矩形更新目标检测框 $\text{BOX}^*$ ;

(55) 判断 $\text{IoU}(\text{box}^*, \text{BOX}^*) > \text{TH}$ 是否成立, 若上式成立, 则 $\text{box}^* = \text{BOX}^*$ , 得到最终目标检测

框 $\text{box}^*$ ; 否则 $\text{box}^* = \text{BOX}^*$ , 返回步骤(54); 其中, TH表示停止阈值。

5. 根据权利要求1或4所述的一种目标位置检测方法, 其特征在于, 所述步骤(53)中通过边缘图a和b间的相互补充, 检测出边缘点并连接成线具体为:

(531) 对图像b进行遍历, 若检测到未访问过像素点, 且该像素点的灰度值不为零, 则记像素点为 $p(i, j)$ , 并找出以 $p(i, j)$ 为起始点的边缘线; 否则进入步骤(534);

(532) 在图像a中, 找到和图像b中点 $p(i, j)$ 对应的点, 记为 $s(i, j)$ , 若在点 $s(i, j)$ 的8邻域中发现有非零像素点, 则记该像素点为 $u(i, j)$ , 并在图像b中找到与点 $u(i, j)$ 所对应的像素点, 记为 $r(i, j)$ , 找出以像素点 $r(i, j)$ 为起始点的边缘线; 否则直接进入步骤(533);

(533) 将像素点 $p(i, j)$ 标记为已访问, 返回步骤(531);

(534) 由所有边缘线得到最终目标边缘图。

6. 根据权利要求1所述的一种目标位置检测方法, 其特征在于, 所述步骤(54)中像素点 $(x, y)$ 和边缘点 $(x', y')$ 之间的距离为 $dt(x, y)$ :

$$dt(x, y) = \min_{x', y'} (f(x', y') + (x' - x)^2 + (y' - y)^2),$$

其中,  $f(x', y')$ 表示将边缘点 $(x', y')$ 与像素点 $(x, y)$ 的灰度值交换后的灰度值。

7. 根据权利要求1所述的一种目标位置检测方法, 其特征在于, 所述步骤(1)中训练目标检测识别神经网络模型具体过程包括:

St1利用预训练的VGG16模型初始化特征提取网络, 训练目标潜在区域生成网络, 同时调整特征提取网络参数;

St2利用预训练的VGG16模型再次初始化特征提取网络, 利用上一步训练好的目标潜在区域生成网络生成目标潜在区域, 训练预测网络模型以及特征提取网络;

St3利用上一步训练好的特征提取网络和预测网络模型, 重新训练目标潜在区域生成网络, 调整目标潜在区域生成网络参数;

St4利用之前训练好的特征提取网络和目标潜在区域生成网络, 重新训练预测网络模型, 修改预测网络模型参数。

## 一种水面全景图像目标检测识别方法

### 技术领域

[0001] 本发明属于计算机视觉识别技术领域,更具体地,涉及一种水面全景图像目标检测识别方法。

### 背景技术

[0002] 随着海洋强国战略的提出,我国已经开始大力发展海洋设备。无人艇作为一种无人水面移动平台,不仅能够进入一些恶劣环境中完成任务,同时也能够提高需要长时间作业的任务的完成效率,在许多实际应用中都发挥着十分重要的作用,例如海关巡逻、浅海扫雷和水质监测等等。对于无人艇而言,环境感知技术是其完成自主航行和自主避障不可缺少的一环。特别地,在实际作业的过程当中,仅能够检测前进方向的目标或障碍物是远远不够的,危险可能出现在无人艇周围任何的方位。因此,装备全景相机来采集无人艇周围360°范围内的环境信息,并实时检测其中的目标及障碍物,是提高无人艇自主航行安全性的有效手段。

[0003] 在现有的技术中,尚没有将水面全景图像用于目标检测识别任务之中。有人曾提出一种基于三维全景环视的无人艇自动锚泊视觉辅助系统,在艇身周围布置多个相机和毫米波测距雷达,通过拼接多个相机的图像来合成全景图像辅助测距雷达进行自动锚泊,通过无人艇身上的多个相机采集图像信息实现广阔的视角监控。但是所述方法仅是利用全景图像辅助测距雷达进行自动锚泊,没有利用全景图像信息进行目标识别,目前常用的目标检测识别算法如下:

[0004] 2014年,Ross Girshick等人发表了R-CNN算法,首次提出了基于目标潜在区域结合卷积神经网络的目标检测算法,代替了使用滑动窗口和手工设计特征的传统目标检测算法,使得目标检测的准确率得到了本质的提高。2015年,Kaiming He等人提出了Faster R-CNN算法,用一个卷积神经网络模型同时完成了分类任务与回归任务,实现了目标检测与识别;同时构建了RPN网络模型,大大缩减了算法运行时间。该算法取得了当时最好的效果,在VOC 2007数据集上正确率达到了73.2%。2015年,Wei Liu等人提出SSD算法,不同于R-CNN系列算法需要提取目标潜在区域,直接利用多层卷积特征生成目标检测框。由于不需要生成目标潜在区域,SSD算法的速度得到很大的提升。SSD300网络在VOC2007数据集上获得72.1%正确率的同时,计算速度达到了58fps。2016年,Jifeng Dai等人发表了R-FCN算法。R-FCN算法为了解决卷积神经网络池化层大量丢失图像空间信息的问题,提出了位置敏感型卷积滤波器,大大提高了目标检测识别正确率。

[0005] 上述方法对当前公共数据库中的图像都有很好的效果,但是全景图像是由多个相机拍摄的照片变形拼接而形成,因而会存在拼接裂缝以及相机鱼眼模式产生的扭曲等畸变,这些畸变会导致目标产生较大的形变,上述基于深度学习的目标检测识别算法,能够克服目标一定程度的形变,但是无法解决目标大的形变,难以检测到目标准确位置,因此针对现有的全景图像,现有的基于深度学习的目标检测识别算法对目标位置识别准确率不够高。

## 发明内容

[0006] 针对现有技术的以上缺陷或改进需求,本发明提供了计算机视觉识别技术,其目的在于先利用基于深度学习的目标检测算法强大的拟合能力,找到初始目标检测框;之后针对全景图像畸变大的特点,分割目标局部区域产生若干超像素块,利用多阈值超像素块融合策略,调整初始目标检测框;接着根据目标检测框周围边缘点的分布情况,再次调整目标检测框位置,得到最终目标检测框;最后将最终目标检测框位置转化为目标的实际位置。由此解决现有技术中全景图像目标识别无法解决图像畸变影响的技术问题。

[0007] 为实现上述目的,提供了计算机视觉识别技术,所述方法包括:

[0008] (1) 利用训练好的目标检测识别神经网络模型对全景图像进行目标检测识别,获取目标类别和初始目标检测框 $box$ ;

[0009] (2) 在全景图像上截取以目标检测框 $box$ 为中心, $\gamma$ 倍目标检测框 $box$ 大小的矩形区域作为目标局部区域;其中, $1.1 \leq \gamma \leq 1.5$ ,优选 $\gamma = 1.2$ ;

[0010] (3) 对目标局部区域进行图像分割,得到多个超像素块,再采用多阈值融合策略合并所述超像素块,得到调整后目标检测框 $box^*$ ;

[0011] (4) 在全景图像上截取以目标检测框 $box^*$ 为中心, $\gamma$ 倍目标检测框 $box^*$ 大小的矩形区域作为新的目标局部区域;

[0012] (5) 计算新的目标局部区域的边缘图,根据目标检测框 $box^*$ 周围边缘点的分布调整 $box^*$ ,得到最终目标检测框 $box^*$ ;

[0013] (6) 由全景图像成像平面坐标系与世界坐标系的映射关系,将最终目标检测框 $box^*$ 位置转化为目标的实际位置。

[0014] 进一步地,所述步骤(1)具体包括:

[0015] (11) 对全景图像进行尺寸归一化,之后对归一化后图像去均值;

[0016] (12) 利用特征提取网络计算全景图像的卷积特征,以最后一层卷积特征图为输入构建目标潜在区域生成网络,由目标潜在区域生成网络生成目标潜在区域;

[0017] (13) 对目标潜在区域进行RoI Pooling归一化至固定大小,之后将其输入到预测网络中,得到多个目标检测框;

[0018] (14) 采用非极大值抑制抑制算法,剔除多余的目标检测框,得到初始目标检测框 $box$ 。

[0019] 进一步地,所述步骤(3)具体包括:

[0020] (31) 计算目标局部区域中所有像素点之间的不相似度 $e$ :

$$[0021] \quad e = \sqrt{(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2},$$

[0022] 其中, $r$ 、 $g$ 和 $b$ 分别代表像素点在RGB三通道的值;定义像素点合并阈值 $\theta$ , $8 \leq \theta \leq 10$ ,优选 $\theta = 9$ ;若两个像素点间的不相似度满足 $e < \theta$ ,则将两像素点合并成一个区域 $C$ ;

[0023] 在区域 $C$ 中,最大的不相似度作为类内差异 $Int(C)$ :

$$[0024] \quad Int(C) = \max e,$$

[0025] 将两区域 $C_1$ 和 $C_2$ 间最小的不相似度作为类间差异 $Diff(C_1, C_2)$ :

$$[0026] \quad Diff(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2} e(v_i, v_j),$$

[0027] 其中,  $v_i$ 和 $v_j$ 分别表示区域 $C_1$ 和 $C_2$ 中的像素点;

[0028] 若区域 $C_1$ 和 $C_2$ 的类间差异同时小于各自的类内差异,则将区域 $C_1$ 和 $C_2$ 合并;对目标局部区域中所有区域进行迭代融合,直到不满足融合条件,最终得到的区域作为超像素块;

[0029] (32)为每一个超像素块构建一个最小外接矩形 $s$ ,计算超像素块和初始目标检测框 $box$ 的重合率:

$$[0030] \quad IoU(box, s) = \frac{|box \cap s|}{|box \cup s|},$$

[0031] 构建两个集合:

$$[0032] \quad S_{inner} = \{s \in S_0 \mid IoU(s, box) = 1\},$$

$$[0033] \quad S_{inter} = \{s \in S_0 \mid 0 < IoU(s, box) < 1\},$$

[0034] 其中,  $S_0$ 表示和目标检测框 $box$ 有交集的超像素块集合;  $S_{inner}$ 表示被目标检测框 $box$ 包含的超像素块集合;  $S_{inter}$ 表示和目标检测框 $box$ 部分相交的超像素块集合;

[0035] 将集合 $S_{inner}$ 区域中所有超像素块融合,获取融合后区域的最小外接矩形 $box^*$ ;

[0036] (33)遍历集合 $S_{inter}$ ,若满足:

$$[0037] \quad IoU(box^* \cup s_i, box^*) > \delta_1, s_i \in S_{inter},$$

[0038] 则将满足上式的 $s_i$ 从集合 $S_{inter}$ 中剔除得到新的集合 $S_{inter}$ ,用 $box^* \cup s_i$ 更新 $box^*$ ;其中 $\delta_1$ 表示第一超像素块融合阈值,  $0.5 \leq \delta_1 \leq 0.7$ , 优选 $\delta_1 = 0.5$ ;

[0039] (34)将第二超像素块融合阈值 $\delta_2$ 更换 $\delta_1$ ,代入新的 $S_{inter}$ 和 $box^*$ 重复步骤(33),得到新的集合 $S_{inter}$ 和 $box^*$ ;再次用第三超像素块融合阈值 $\delta_3$ 更换 $\delta_2$ ,代入新的 $S_{inter}$ 和 $box^*$ 重复步骤(33),得到新的 $box^*$ ,  $box^*$ 既为调整后的目标检测框;其中,  $0.6 \leq \delta_2 \leq 0.8$ , 优选 $\delta_2 = 0.7$ ;  $0.7 \leq \delta_3 \leq 0.9$ , 优选 $\delta_3 = 0.8$ 。

[0040] 进一步地,所述步骤(5)具体包括:

[0041] (51)将新的目标局部区域转换为灰度图像,在利用下式以滑动窗口的方式对图像进行平滑去噪处理;

$$[0042] \quad G = \frac{1}{2\pi\delta^2} e^{-\frac{x^2+y^2}{2\delta^2}},$$

[0043] 其中,  $x$ 和 $y$ 为目标局部区域图像中每一个像素的坐标值,  $\delta$ 为高斯函数模板;

[0044] (52)对图像的 $x$ 方向和 $y$ 方向分别利用一阶有限差分来模拟偏导,其中差分计算模板为:

$$[0045] \quad S_x = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix},$$

$$[0046] \quad S_y = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix},$$

[0047] 利用该模板通过卷积的方式计算图像中每一个元素 $f(i, j)$ 的梯度幅值 $M(i, j)$ 和方向 $\theta(i, j)$ :

$$[0048] \quad M(i, j) = \sqrt{(f(i, j) * S_x(i, j))^2 + (f(i, j) * S_y(i, j))^2},$$

$$[0049] \quad \theta(i, j) = \arctan((f(i, j) * S_y(i, j)) / (f(i, j) * S_x(i, j))),$$

[0050] 利用非极大值抑制算法得到局部梯度最大的点作为目标潜在边缘点,潜在的边缘点灰度值置为128,其余像素点灰度值全部置零,得到目标局部区域的二值化图像;

[0051] (53) 设置两个边缘检测阈值 $\beta_1$ 、 $\beta_2$ ,其中 $\beta_1=0.4\beta_2$ ,将二值化图像中梯度值小于 $\beta_1$ 的像素灰度值设为零,得到目标边缘图像a,同时将二值化图像中梯度值小于 $\beta_2$ 的像素灰度值设为零,得到目标边缘图像b;再通过目标边缘图a和b间的相互补充,得到最终目标边缘图;

[0052] (54) 计算最终目标边缘图中边缘点和目标检测框 $box^*$ 四条边框上的像素点之间的距离,找出距离边缘点最近的像素点,用所述像素点组成图像的最小外接矩形更新目标检测框 $BOX^*$ ;

[0053] (55) 判断 $IoU(box^*,BOX^*) > TH$ 是否成立,若上式成立,则 $box^*=BOX^*$ ,得到最终目标检测框 $box^*$ ;否则 $box^*=BOX^*$ ,返回步骤(54);其中,TH表示停止阈值, $0.8 \leq TH \leq 0.9$ ,优选 $TH=0.9$ 。

[0054] 进一步地,所述步骤(53)中通过边缘图a和b间的相互补充,检测出边缘点并连接成线具体为:

[0055] (531) 对图像b进行遍历,若检测到未访问过像素点,且该像素点的灰度值不为零,则记像素点为 $p(i, j)$ ,并找出以 $p(i, j)$ 为起始点的边缘线;否则进入步骤(534);

[0056] (532) 在图像a中,找到和图像b中点 $p(i, j)$ 对应的点,记为 $s(i, j)$ ,若在点 $s(i, j)$ 的8邻域中发现有非零像素点,则记该像素点为 $u(i, j)$ ,并在图像b中找到与点 $u(i, j)$ 所对应的像素点,记为 $r(i, j)$ ,找出以像素点 $r(i, j)$ 为起始点的边缘线;否则直接进入步骤(533);

[0057] (533) 将像素点 $p(i, j)$ 标记为已访问,返回步骤(531);

[0058] (534) 由所有边缘线得到最终目标边缘图。

[0059] 进一步地,所述步骤(54)中像素点 $(x, y)$ 和边缘点 $(x', y')$ 之间的距离为 $dt(x, y)$ :

$$[0060] \quad dt(x, y) = \min_{x', y'} (f(x', y') + (x' - x)^2 + (y' - y)^2),$$

[0061] 其中, $f(x', y')$ 表示将边缘点 $(x', y')$ 与像素点 $(x, y)$ 的灰度值交换后的灰度值。

[0062] 进一步地,所述步骤(1)中训练目标检测识别神经网络模型具体过程包括:

[0063] St1利用预训练的VGG16模型初始化特征提取网络,训练目标潜在区域生成网络,同时调整特征提取网络参数;

[0064] St2利用预训练的VGG16模型再次初始化特征提取网络,利用上一步训练好的目标潜在区域生成网络生成目标潜在区域,训练预测网络模型以及特征提取网络;

[0065] St3利用上一步训练好的特征提取网络和预测网络模型,重新训练目标潜在区域生成网络,调整目标潜在区域生成网络参数;

[0066] St4利用之前训练好的特征提取网络和目标潜在区域生成网络,重新训练预测网络模型,修改预测网络模型参数。

[0067] 总体而言,通过本发明所构思的以上技术方案与现有技术相比,具有以下技术特征及有益效果:

[0068] 本发明采用基于深度学习的目标检测算法检测出全景图像中所有目标的初始位置,并识别出目标的类别;利用图像分割算法处理目标局部区域得到若干超像素块,利用多



阈值超像素块融合策略调整目标检测框位置;最后利用边缘提取算法得到的目标局部区域边缘图,根据目标检测框周围边缘点的分布情况,调整得到最终的目标检测框,并将最终目标检测框的位置转化为实际目标位置,可以有效克服全景图像中畸变效应,从全景图像中将目标位置准确识别出来。

### 附图说明

- [0069] 图1是本发明方法的流程图;  
 [0070] 图2是本发明方法实施例的实施过程示意图;  
 [0071] 图3是实施例检测识别结果。

### 具体实施方式

[0072] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本发明进行进一步详细说明。应当理解,此处所描述的具体实施例仅用以解释本发明,并不用于限定本发明。此外,下面所描述的本发明各个实施方式中所涉及到的技术特征只要彼此之间未构成冲突就可以相互组合。

[0073] 如图1所示,结合实施例对本发明方法的流程进行介绍,本发明实施例方法包括以下步骤:

[0074] 步骤1训练目标检测识别神经网络模型。目标检测识别网络由特征提取网络、目标潜在区域生成网络(RPN网络)和预测网络三部分组成,训练的过程为分步式训练

[0075] 特征提取网络用来提取图像的卷积特征,目标潜在区域生成网络利用卷积特征生成可能存在目标的目标检测框,预测网络对这些可能存在目标的检测框进行进一步的预测和调整,得到最终的目标位置和类别信息。

[0076] 步骤1具体包括以下子步骤:

[0077] (1-1)定义网络误差函数。对于目标潜在区域生成网络,误差函数如下:

$$[0078] \quad L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)$$

[0079] 其中*i*为目标潜在区域的序号, $p_i$ 为第*i*个目标潜在区域的预测的类别, $p_i^*$ 为第*i*个真实目标框的类别, $p_i^*$ 的值为1。如果第*i*个目标潜在区域与任何一个真实目标框的交集-并集比(IoU)大于0.7,则 $p_i$ 等于1否则等于0。 $t_i$ 是第*i*个目标潜在区域的4个坐标,其中 $t_i \in \{t_x, t_y, t_w, t_h\}$ , $t_i^*$ 为第*i*个真实目标框的4个坐标。 $N_{cls}$ 为训练时每一批图片的数量, $N_{reg}$ 为所有锚点框的数量, $\lambda$ 为分类损失与回归损失的平衡因子。 $L_{cls}$ 为分类损失函数,是一个标准的对数损失函数。 $L_{reg}$ 是回归损失函数,具体如下所示:

$$[0080] \quad L_{reg}(t_i, t_i^*) = \sum_{i \in \{x, y, w, h\}} smooth_{t_i}(t_i - t_i^*)$$

$$[0081] \quad smooth_{t_i}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}$$

$$[0082] \quad t_x = (x - x_a) / w_a, t_y = (y - y_a) / h_a, t_w = \log(w / w_a), t_h = \log(h / h_a)$$

$$[0083] \quad t_x^* = (x^* - x_a) / w_a, t_y^* = (y^* - y_a) / h_a, t_w^* = \log(w^* / w_a), t_h^* = \log(h^* / h_a)$$

[0084] 其中,  $x_a$ 代表锚点框左上角的横坐标,  $x$ 代表目标潜在区域左上角的横坐标,即为锚点框回归之后的坐标,  $x^*$ 代表真实目标框左上角的横坐标。  $y$ 、 $w$ 和 $h$ 依次代表目标框的左上角纵坐标、宽和高。

[0085] 对于预测网络,误差定义如下:

$$[0086] \quad L(p, u, t, t^*) = L_{cls}(p, u) + \lambda[u \geq 1]L_{reg}(t, t^*)$$

[0087] 其中 $p$ 代表目标被预测的类别,在本实施例中, $p$ 共有军舰、帆船、浮标等25种类型。 $u$ 代表目标检测框的真实类别,  $[u \geq 1]$ 表示当 $u \geq 1$ 时此式等于1,否则此式的值为0。 $t$ 表示目标被预测的位置,  $t^*$ 表示目标实际的位置,其具体含义与目前潜在区域生成网络中对应的变量相同。 $L_{cls}$ 、 $L_{reg}$ 以及 $\lambda$ 的含义均与目标潜在区域生成网络中对应变量相同。

[0088] (1-2)分步式训练。训练的过程主要分为4个部分:

[0089] 1) 利用预训练好的VGG16模型初始化特征提取网络,训练RPN网络,同时调整特征提取网络参数;

[0090] 2) 利用预训练好的VGG16模型再次初始化特征提取网络,利用上一步训练好的RPN网络生成目标潜在区域,训练预测网络模型以及特征提取网络;

[0091] 3) 利用上一步训练好的特征提取网络和预测网络模型,重新训练RPN网络,这一过程只调整RPN网络参数,不改变特征提取网络与预测网络模型的参数;

[0092] 4) 利用之前训练好的特征提取网络和RPN网络,重新训练预测网络模型,此过程只修改预测网络模型参数。通过以上四个步骤实现了RPN网络和预测网络的卷积共享,完成了算法的训练过程。

[0093] 步骤2拍摄全景图像。此步骤可以利用一体化的全景相机,也可以利用多个普通相机组合成全景相机。在本实施例中,将6个GoPro相机置于船体上方,分别拍摄6个方向的图片,然后利用全景拼接软件Autopano Giga将6张图片拼接成一副完成的全景图像。此外,在拍摄前利用相机标定板标定相机,建立成像平面坐标系与世界坐标系的映射关系,即能够将目标在全景图像上的坐标转换为真实空间的坐标。

[0094] 步骤3目标检测识别。对全景图像进行目标检测,具体包括以下子步骤:

[0095] (3-1) 图片初始化。在本发明实施例中,全景图像由多个相机拍摄结果拼接而成,因此图片尺寸非常大。首先将全景图像归一化至 $800 \times 1000$ 大小,然后对图像去均值。

[0096] (3-2) 计算目标潜在区域。利用特征提取网络计算全景图像的卷积特征,以最后一层卷积特征图为输入构建RPN网络。在RPN网络上,以每个点为中心,构建9种尺度的锚点框。并通过滑动窗口的方式,遍历RPN网络。接着将所有的锚点框进行分类和回归操作,根据分类得分排序,取得分前300的锚点框作为目标潜在区域。

[0097] (3-3) 计算目标初始位置与类别。将步骤(3-2)得到的300个目标潜在区域,经过RoI Pooling层归一化至固定大小。再将这些目标潜在区域送入预测网络中,进行分类和回归,得到目标的初始位置和类别。

[0098] (3-4) 剔除冗余目标框。步骤(3-3)会产生很多重复的目标检测框,同时框住一个目标。利用非极大值抑制算法,剔除冗余的目标检测框。

[0099] 如图2所示,进行步骤4获取目标局部区域。由步骤3可得初始目标检测框,在全景图像上截取以此框为中心,面积为目标框 $\gamma$ 倍的矩形区域,作为目标局部区域。在本实施例中, $\gamma = 1.2$ 。

[0100] 步骤5基于超像素分割调整目标检测框。首先将步骤4得到的目标局部区域进行图像分割,得到若干超像素块。然后利用本发明提出的多阈值融合策略合并相应的超像素块,从而调整目标检测框的大小和位置。步骤5具体包括以下子步骤:

[0101] (5-1) 图像分割。将图像中的像素点看作是一副图的顶点,像素点间的不相似度看作是顶点之间的边,通过将相似的点融合成一块块的区域,从而将图像分割成为若干超像素块。将像素点间的颜色距离作为不相似度 $e$ 测量方式,定义如下:

$$[0102] \quad e = \sqrt{(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2}$$

[0103] 其中 $r$ 、 $g$ 、 $b$ 分别代表像素点在RGB三通道的值。定义像素点合并阈值 $\theta$ ,若两个像素点间的不相似度 $e$ 小于合并阈值 $\theta$ ,则将像素点合并成一个区域 $C$ 。对于区域 $C$ ,定义其类内差异 $Int(C)$ 如下:

$$[0104] \quad Int(C) = \max_{e \in (MST, E)} e$$

[0105] 将区域 $C$ 中,不相似度最大的两个像素点之间的颜色距离作为类内差异。对于两个区域 $C_1$ 、 $C_2$ ,将两个区域间最小的颜色距离作为类间差异,定义如下:

$$[0106] \quad Diff(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2} e(v_i, v_j)$$

[0107] 其中 $v_i$ 和 $v_j$ 表示两个区域中的像素点。若两个区域间的类间差异同时小于各自区域的类内差异,则合并区域 $C_1$ 、 $C_2$ 。区域融合是一个迭代的过程,如果融合一次后,仍有两个区域的类间差异小于各自类内差异,则继续融合,否则停止融合。

[0108] (5-2) 基于超像素块融合调整目标检测框。对于步骤3得到的初始目标检测框 $box$ ,经过步骤(5-1)计算出对应目标局部区域的超像素分割图。为每一个超像素块构建一个最小外接矩形 $s$ 来代表超像素块,按照下式计算每一个超像素块 $s$ 与初始目标检测框 $box$ 的IoU值。

$$[0109] \quad IoU(box, s) = \frac{|box \cap s|}{|box \cup s|}$$

[0110] 对于与目标检测框 $box$ 有交集的超像素块 $s \in S_\theta$ ,按照以下公式构建两个集合 $S_{inner}$ 、 $S_{inter}$ ,分别代表完全被检测框包含的超像素块集合以及与检测框部分相交的超像素块集合。

$$[0111] \quad S_{inner} = \{s \in S_\theta | IoU(s, box) = 1\}$$

$$[0112] \quad S_{inter} = \{s \in S_\theta | 0 < IoU(s, box) < 1\}$$

[0113] 首先将集合 $S_{inner}$ 中的超像素块融合,得到最小外接矩形 $box^*$ 。然后遍历集合 $S_{inter}$ 中的每一个超像素块 $s$ ,计算 $s$ 与 $box^*$ 融合之后的矩形框与 $box^*$ 的IoU值,如果满足以下要求,则在遍历完集合 $S_{inter}$ 后,用 $box^* \cup s$ 更新矩形框 $box^*$ 。

$$[0114] \quad IoU(box^* \cup s_i) > \delta_1, s_i \in S_{inter}$$

[0115] 其中 $\delta_1$ 是人工设定的超参数,代表超像素块融合阈值。经过一轮遍历 $S_{inter}$ 后,将满足上述的超像素块从中剔除,得到新的 $S_{inter}$ 集合。按照上述方法再次遍历 $S_{inter}$ 集合,将融合阈值更改为 $\delta_2$ ,更新 $box^*$ 矩形框。同样的步骤,以融合阈值 $\delta_3$ 再次更新 $box^*$ ,从而得到调整之后的目标检测框 $box^*$ 。在本实施例中, $\delta_1$ 的值设为0.5, $\delta_2$ 的值设为0.7。

[0116] 步骤6获取目标局部区域。由步骤5得到初次调整后的目标检测框,在全景图像上

截取以此框为中心,面积为目标框  $\gamma$  倍的矩形区域,作为目标局部区域。在本实施例中,  $\gamma = 1.2$ 。

[0117] 步骤7基于边缘信息调整目标检测框。计算步骤6得到的目标局部区域的边缘图,根据目标检测框  $\text{box}^*$ 附近的边缘点分布情况,再次调整目标检测框  $\text{box}^*$ 。具体包括以下子步骤:

[0118] (7-1) 目标局部区域预处理。首先将目标局部区域从彩色图转换为灰度图像,再利用离散化的二维高斯函数以滑动窗口的方式对图像进行平滑,以消除图像中噪声的影响。高斯模糊公式如下:

$$[0119] \quad G = \frac{1}{2\pi\delta^2} e^{-\frac{x^2+y^2}{2\delta^2}}$$

[0120] 其中,  $x, y$  为目标局部区域图像中每一个像素的位置,为高斯函数模板,常见的模板有  $3 \times 3$ 、 $5 \times 5$ 、 $7 \times 7$ 。在本实施例中,利用  $7 \times 7$  大小的高斯模板。

[0121] (7-2) 计算梯度。对图像的  $x$  方向和  $y$  方向分别利用一阶有限差分,来近似两个方向上的偏导,其中差分计算模板如下:

$$[0122] \quad S_x = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \quad S_y = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

[0123] 利用如上模板通过卷积的方式,计算图像中每一个元素  $f(i, j)$  的梯度幅值  $M(i, j)$  和方向  $\theta(i, j)$ ,计算公式如下:

$$[0124] \quad M(i, j) = \sqrt{(f(i, j) * S_x(i, j))^2 + (f(i, j) * S_y(i, j))^2}$$

$$[0125] \quad \theta(i, j) = \arctan((f(i, j) * S_y(i, j)) / (f(i, j) * S_x(i, j)))$$

[0126] 此时计算的梯度为全局梯度,还不足以确定边缘,需要利用非极大值抑制算法保留局部梯度最大的点。完成非极大值抑制操作后,得到目标局部区域的二值化图像,非边缘点的灰度值为零,潜在的边缘点灰度值为128。

[0127] (7-3) 检测和连接边缘。设置两个边缘检测阈值  $\beta_1, \beta_2$ , 其中  $\beta_1 = 0.4\beta_2$ 。在步骤 (7-2) 所得的目标局部区域二值化图像中,将梯度值小于  $\beta_1$  的像素灰度值设为零,得到包含较多边缘的图像  $a$ 。同时,将目标局部区域二值化图像中梯度值小于  $\beta_2$  的像素灰度值设为零,得到图像  $b$ 。由于  $\beta_2$  的阈值较高,会过滤掉包括真实边缘信息在内的大部分边缘信息。考虑到图像  $a$  中保留了大部分边缘信息,因此利用图像  $a$  作为补充,在图像  $b$  的基础上连接边缘,从而得到最终的边缘图。

[0128] 对图像  $b$  进行遍历,若检测到像素  $p(i, j)$  的灰度值不为零,则跟踪以  $p(i, j)$  为起始点的边缘线,找到边缘线终点像素  $q(i, j)$ 。然后,在图像  $a$  中,找到与图像  $b$  中像素点  $p(i, j)$  对应的点  $s(i, j)$ 。若在  $s(i, j)$  的8邻域中发现有非零像素点  $u(i, j)$ ,则在图像  $a$  中找到与  $u(i, j)$  所对应的像素点,记为  $r(i, j)$ 。在图像  $b$  中,以像素点  $r(i, j)$  为起始点,重复第一个步骤,继续寻找边缘线。若在图像  $a$  和图像  $b$  中,无法找到非零像素点,则以  $p(i, j)$  为起点的边缘线检测完成,将其标记为已访问。继续遍历图像  $b$ ,重复上述步骤,检测出所有边缘线,得到最终的边缘检测图。

[0129] (7-4) 基于边缘信息调整目标检测框。在边缘图上,计算每一个像素点  $(x, y)$  与边缘点之间的距离,找出距离最近的边缘点  $(x', y')$ ,通过下式能够快速求出像素点与周围边

缘点之间的距离：

$$[0130] \quad dt(x, y) = \min_{x', y'} (f(x', y') + (x' - x)^2 + (y' - y)^2)$$

[0131] 其中,  $f$  表示将边缘像素点与背景像素点的灰度值交换,  $f(x', y')$  表示交换之后像素的灰度值。边缘点的灰度值此时为零, 背景点的灰度值为128, 利用上述公式能够快速找到距离像素点  $(x, y)$  最近的边缘点  $(x', y')$ 。

[0132] 首先找到目标检测框  $\text{box}^*$  四条边框上的像素点, 然后计算边缘点距离这些像素点的距离, 找出其中距离最近的边缘点, 用这些边缘点的最小外接矩形更新目标检测框  $\text{box}^*$ 。计算更新前后目标检测框的IoU, 若IoU大于停止阈值, 则停止调整, 否则继续按照之前的步骤更新  $\text{box}^*$ 。在本实施例中, 停止阈值设为0.9。如图3所示, 得到目标的在全景图像中的最终目标检测框。

[0133] 步骤8根据步骤2得到的映射关系, 将步骤7得到的目标位置坐标转换到世界坐标系中, 即在现实空间中的位置, 然后将坐标输入到航迹规划模块, 从而控制无人艇前进方向。

[0134] 以上内容本领域的技术人员容易理解, 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、等同替换和改进等, 均应包含在本发明的保护范围之内。

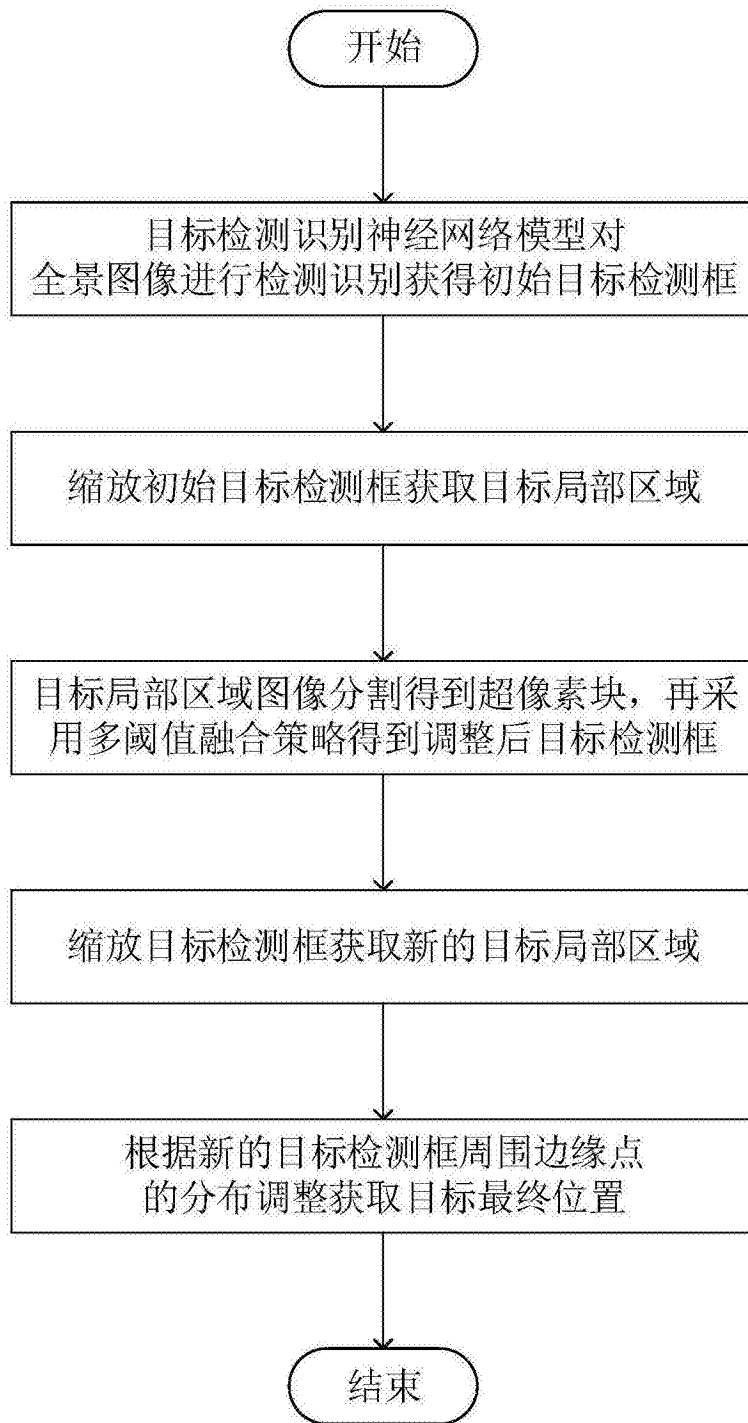


图1

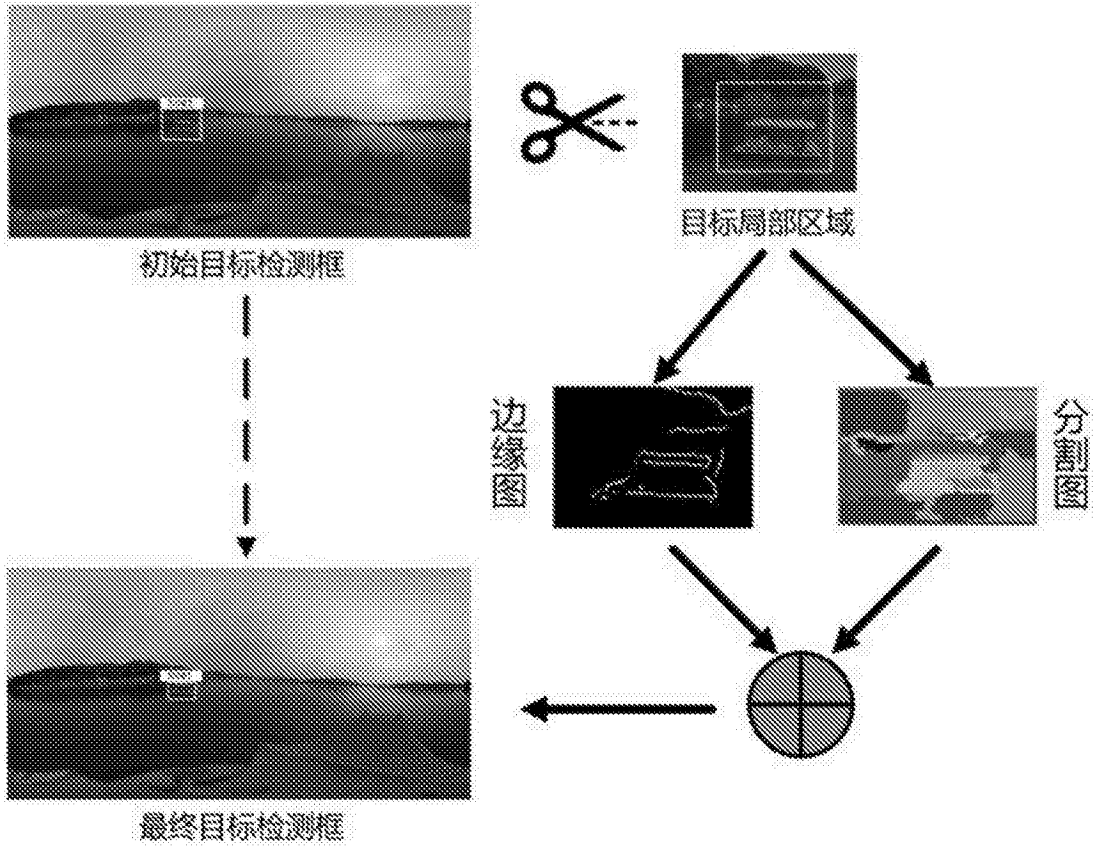


图2

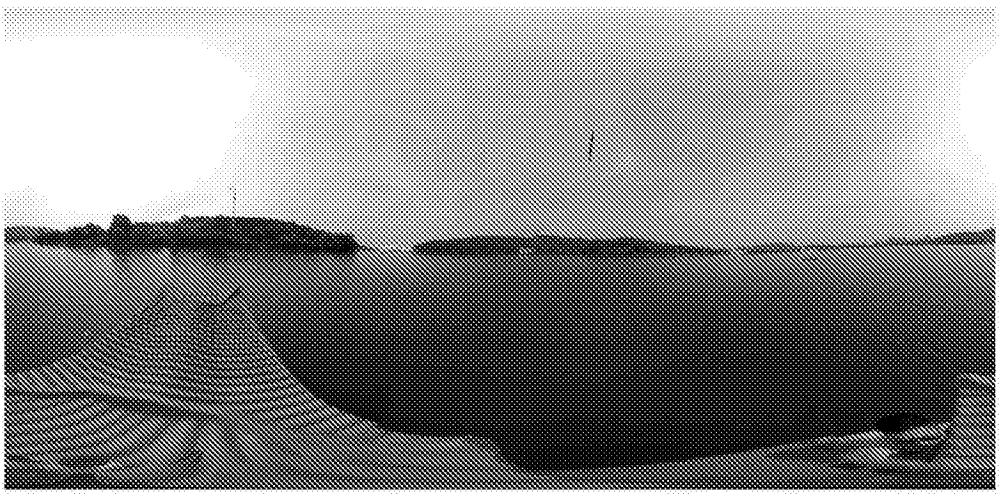


图3