

# 基于视觉导航的输电线杆塔方位确定方法

王祖武<sup>1,2,\*</sup>, 韩军<sup>1,2</sup>, 孙晓斌<sup>3</sup>, 杨波<sup>3</sup>

<sup>1</sup>上海大学通信与信息工程学院, 上海 200444;

<sup>2</sup>上海先进通信与数据科学研究院, 上海 200444;

<sup>3</sup>国网山东省电力公司, 山东 济南 250000

**摘要** 通过分析杆塔镂空的结构特征,提出了一种基于杆塔梯度方向直方图(HOG)的由远及近杆塔部件检测方法。使用不同方位下杆塔 HOG 特征训练多层感知机(MLP),得到训练后的分类模型,将航拍图像输入到分类模型中识别杆塔的方位,最终实现了局部目标的检测。相比于深度学习神经网络,该方法的分类特征更加明确,更具有代表性。实验结果表明,所提方法的检测准确率比 Faster RCNN(Regions with Convolutional Neural Networks)方法高 27.9%,运算时间比 Faster RCNN 减少 70.6%。所提方法适用于在开阔环境下利用无人机对杆塔方位及其局部部件的精确检测。

**关键词** 图像处理; 视觉导航; 梯度方向直方图特征; 多层感知机; 杆塔检测; 神经网络

中图分类号 TN911.73

文献标识码 A

doi: 10.3788/LOP56.081006

## Method for Orientation Determination of Transmission Line Tower Based on Visual Navigation

Wang Zuwu<sup>1,2,\*</sup>, Han Jun<sup>1,2</sup>, Sun Xiaobin<sup>3</sup>, Yang Bo<sup>3</sup>

<sup>1</sup> School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China;

<sup>2</sup> Shanghai Institute of Advanced Communications and Data Science, Shanghai 200444, China;

<sup>3</sup> State Grid Shandong Electric Power Company, Jinan, Shandong 250000, China

**Abstract** A method for detecting the tower components from far to near is proposed based on the histogram of the gradient (HOG) along the tower gradient direction by analyzing the structural characteristics of hollowing-out of a tower. The multi-layer perceptron (MLP) is first trained by using the HOG feature of the tower under different orientations to obtain a trained classification model, then the aerial image is input into this classification model to identify the orientation of the tower, and the detection of a local target is finally realized. Compared with that of a deep learning neural network, the classification feature of the proposed method is clear and representative. The experimental results show that the detection accuracy of the proposed method is 27.9% higher than that of the Faster RCNN (Regions with Convolutional Neural Networks) method, but the computation time is 70.9% lower than the latter. The proposed method is suitable for the accurate detection of the tower orientation and its local parts by the unmanned aerial vehicle in an open environment.

**Key words** image processing; visual navigation; histogram of the gradient (HOG) feature; multi-layer perceptron; tower detection; neural network

**OCIS codes** 100.2000; 110.2960; 100.3008; 100.4996

## 1 引言

由于电力线基础设施长期受到自然界风、雨等的

侵蚀,杆塔倾斜、绝缘子爆裂、均压环螺丝脱落等电力部件损坏的现象频繁出现,巡检人员必须对其进行周期性地检查,其中最主要的是对电力线与杆塔连接部

收稿日期: 2018-10-22; 修回日期: 2018-11-08; 录用日期: 2018-11-13

基金项目: 国家重点研发计划(2017YFC08067)

\* E-mail: wangzuwu@shu.edu.cn

件的局部进行巡检。随着多旋翼无人机(UAV)技术的快速发展,其在电力巡检领域有着越来越广泛的应用,极大地提高了高压输电线路的维护效率。但是UAV在电力巡检作业时,巡检人员不仅需操控UAV安全地飞行,还要操作UAV上安装的相机设备,对杆塔进行有效地拍摄。这使得巡检人员获取高质量的巡检图像的难度较大,在对一条线路多次巡检时不能保证获所取的巡检图像的质量是相同的。

杆塔具有丰富的直线特征,大多数研究者通过提取无人机航拍图像中的直线特征来检测杆塔<sup>[1-3]</sup>。Cheng等<sup>[1]</sup>使用一种基于graph-cut的图分割方法检测电力线杆塔区域,这种图分割方法适用于检测低压线路的杆塔。Han等<sup>[2]</sup>使用一种模板匹配的方法,将提取到的杆塔直线特征和预先构建好的杆塔模板库进行匹配,检测拍摄图像中的杆塔目标。该方法事先使用聚类 and 投影来构建杆塔的方位模板库,但是其所使用的直线段检测(LSD)方法是一种局部方法,易受到噪声的干扰,无法处理杆件交叉导致的直线段断开的情况。Cetin等<sup>[3]</sup>提出一种通过提取航拍图像中杆塔阴影区域的特征来检测杆塔的方法。由于该方法依赖于杆塔阴影,因此其实际应用中受到了较大的限制。除了直线特征之外,也有些研究者使用特征点来检测杆塔,如角点特征、ORB(Oriented Features from Accelerated Segment Test and Rotated Brief)有向旋转描述特征等。Golightly等<sup>[4]</sup>抛弃了使用线特征的检测方法,基于杆塔的角点和边缘对杆塔进行识别,但是该方法主要适用于检测小型普通低压电力线杆塔。Cerón等<sup>[5]</sup>使用基于网格的二维ORB描述符来对电力线杆塔的线性结构进行检测,将对应区域的特征输入分类器进行分类。该方法可实现对杆塔的实时检测。该方法由于使用了FAST角点特征和Brief描述子,没有对区域内部整体特征统计分析,因此该方

法无法识别杆塔的方位。

近年来,随着深度学习在目标检测中的发展,有一些研究者将深度学习引入杆塔检测中来。Hui等<sup>[6]</sup>使用深度学习方法Faster RCNN(Regions with Convolutional Neural Networks)<sup>[7]</sup>对杆塔进行目标检测,并结合KCF<sup>[8]</sup>(Kernel Correlation Filter)核相关滤波器目标跟踪算法对杆塔进行跟踪,实现了利用无人机对杆塔目标的自动巡检。然而,Faster RCNN方法实际处理速度慢,且无法避免杆塔检测失败的情况。因此,Zhao等<sup>[9]</sup>在Ren等<sup>[7]</sup>研究的基础上将检测过程和跟踪过程做了融合,实现了更稳健的电力线检测与跟踪。由于仍然使用Faster RCNN方法对杆塔进行检测,Zhao等的结果在处理速度和识别效果上并没有实质性提升。以上分析可以发现,现有的杆塔的检测方法仅仅实现了杆塔的全局检测或者跟踪,并没有实现对杆塔方位的分类和识别,也没有对电力线与杆塔连接处的局部进行检测。

在杆塔巡检过程中,人们感兴趣的其实主要是杆塔与导线的连接区域。在杆塔巡检中首先要确定当前待巡检的杆塔,而后操控无人机靠近要巡检的杆塔并完成对杆塔局部细节的拍摄。基于这种在全局拍摄的基础上再对局部进行特写拍摄的模式,本文根据杆塔内部不同长度、位置、方向的钢件及其之间的梯度信息提出一种由远及近的杆塔检测方法。该方法结合杆塔的HOG特征,使用多层感知机(MLP)对杆塔方位进行识别,进而对电力线和杆塔连接部件进行检测,最终实现了利用无人机由远及近对杆塔部件的自动跟踪与拍摄。

## 2 由远及近的杆塔检测方法

在电力线巡检任务中,无人机拍摄的杆塔图像如图1所示。电力铁塔由4个主要面组成,无人机以任意一个角度拍摄铁塔时,仅能拍摄到塔身的2个面。

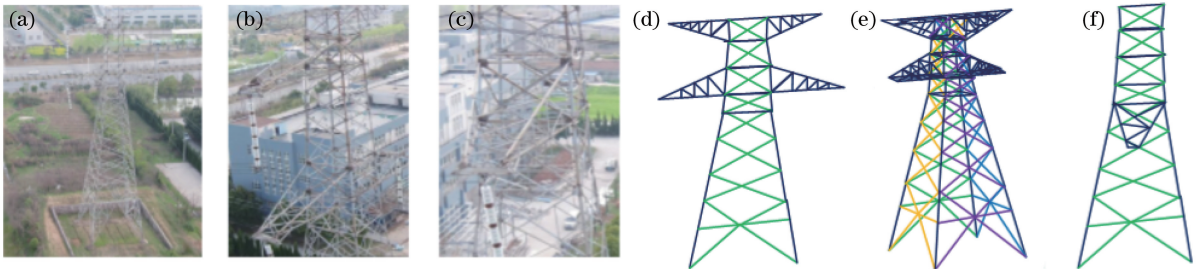


图1 杆塔的航拍图像与示意图。(a)正视图像;(b)正侧面图像;(c)侧面图像;(d)正视示意图;(e)正侧面示意图;(f)侧面示意图

Fig. 1 Aerial images and diagrams of towers. (a) Front image; (b) front side image; (c) side image; (d) front view diagram; (e) front side view diagram; (f) side view diagram

从输电线路结构分析,其主要有两种类型的杆塔:一种是直线杆塔,另一种是耐张杆塔。两种杆塔在主体结构上是相同的,在导线与杆塔的连接方式上有所不同,如图2和3所示。在感知平行直线组、平行下凸曲线组和杆塔区域的情况下,通过共连通区域的计算,研究人员可建立识别直线杆塔与耐张杆塔结构的约束条件。当导线通过直线杆塔时,合成绝缘子的两端分别与导线相连,并对导线形成一种拉力,这使得导线呈现折线(或锐角的连线)。通过计算平行直线组的距离及夹角,即可识别直线杆塔。当导线经过耐张杆塔时,向下凸的引流线会将导线两端连接。通过计算平行直线组与平行曲线组之间距离及夹角,即可识别耐张杆塔。同时,依据绝缘子的安装位置,可以推理出绝缘子的大致位置,进

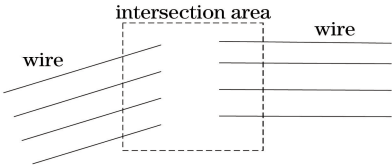


图2 导线穿过直线杆塔时所形成的折线

Fig. 2 Broken lines formed when wire passing through straight tower

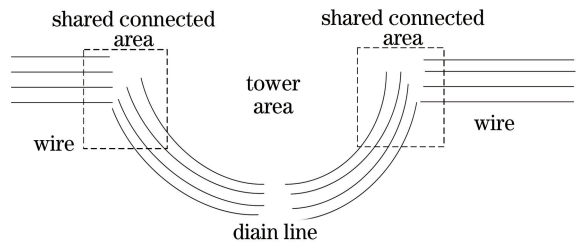


图3 耐张杆塔中引流线与导线位置的关系

Fig. 3 Location relationship between guiding lines and wires in tensile tower

一步诊断连接区域是否存在疑似缺陷。

从铁塔内部结构分析,每个面上的钢件呈规则性排列,在每个面上安装的钢件具有平行及相互交叉的特征。同时,不同面上的钢件会聚到塔架上时呈现出交叉“X”型结构、“K”型结构等不同的特征,如图4所示。在正面拍摄的图像中,塔头呈现正梯形或倒梯形的轮廓形状。在正侧面拍摄图像中,塔头呈现相互重叠的线段轮廓,且塔身内部钢件交错。侧面拍摄的图像可呈现塔身的轮廓形状。不同轮廓形状下杆塔钢件具有不同的结构特征。同时,不同面上的斜钢件会聚到塔架上时会带来差异化的梯度特征。

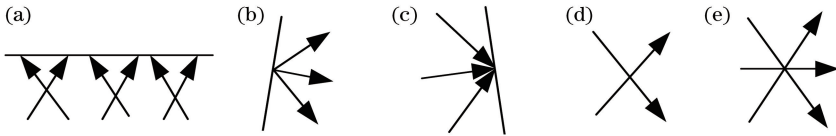


图4 杆塔结构特征。(a)塔头处的交叉斜线段;(b)(c)塔身上的斜线段会聚到主框架;(d)(e)塔内交叉特征

Fig. 4 Structural characteristics of tower. (a) Cross slanted segments of tower head; (b)(c) slanted segments converged onto main frame of tower body; (d)(e) crossing characteristics in tower

图4(a)描述了在塔头区域中交叉的斜线段会聚到横框架上的情形;图4(b)、(c)描述了在塔身区域中斜线段会聚到支撑的主框架上的情形;图4(d)、(e)描述了在铁塔内部线段的交叉特征。可以发现杆塔具有如下的特点:1)杆塔由不同长度、位置、方向的钢件组成,错综复杂的直线型钢结构具有非常丰富的梯度信息;2)杆塔具有镂空的3D线结构,相邻两个主塔面交界处的节点形态往往呈现

出“K”型或者交叉“X”型,钢件的多样性(包括长度、位置、方向)也使得其梯度在幅值和方向上具有较多的变化;3)杆塔由4个面组成,当在不同的拍摄角度下使用无人机对电力线杆塔进行拍摄时,杆塔塔面具有多样化的结构特征。

综合考虑了无人机对杆塔的拍摄方式及杆塔3D镂空结构的特点,本文提出一种由远及近的杆塔检测框架。如图5所示,该框架包括3个阶段:

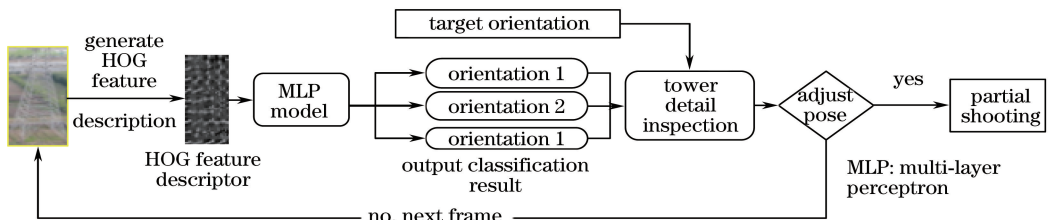


图5 由远及近杆的塔检测框架

Fig. 5 Framework of tower detection from far to near

第1阶段,图像训练。提取对3种不同方位的杆塔图像的HOG特征,并将其输入到MLP训练神经网络模型。

第2阶段,杆塔检测。将无人机在不同方位下拍摄的杆塔图像输入到预训练好的网络模型中,检测杆塔方位。

第3阶段,局部细节巡检。调整无人机的姿态,并对杆塔局部目标(如:均压环、绝缘子等部件)进行拍摄与检测。

### 3 HOG特征提取与方位、区域检测

HOG特征在图像分类识别方面也有一些应用,如:文献[10-11]。本文基于杆塔不同方位下的结构特征,使用杆塔区域内的梯度方向直方图特征,构建相应的描述子。该方法分为训练、检测两个阶段。在训练阶段,提取不同方位下杆塔的梯度方向特征,将统计得到梯度方向特征的描述向量输入到MLP中,训练分类器模型。检测阶段分为两步。第1步,输入无人机拍摄到的杆塔图像,提取杆塔区域的HOG特征描述,并将其输入到分类器中进行分类,检测当前杆塔方位。使用基于图像的视觉伺服<sup>[12]</sup>方法,根据无人机坐标系与相机等部件坐标系之间的转换关系构建无人机动态模型,通过求解非线性优化问题来计算无人机的最佳飞行轨迹,并调整无人机的飞行姿态。第2步,使用文献[13]中的

方法找到电力线与杆塔连接位置,对杆塔部件细节进行拍摄。

#### 3.1 杆塔HOG特征提取

参考Dalal等<sup>[14]</sup>提出的方法,物体的局部外观和形状通常可以通过梯度的强度或者边缘的分布来描述。HOG特征是一种在计算机视觉和图像处理中进行物体检测的描述子,通过计算和统计图像局部区域梯度的方向直方图来构成特征。

以无人机悬停时采集到的RGB图像作为输入,经过图像预处理、分类检测等步骤后输出图像中的杆塔信息,即是否包含杆塔、杆塔所在位置及杆塔类型。提取杆塔区域后对图像进行预处理,将RGB图像转换为灰度图像。经过直方图均衡化后对图像进行目标增强,采用Sobel边缘提取算法提取图像边缘。这可用于计算像素点 $(x,y)$ 处的梯度值和方向 $(0\sim 180^\circ)$ ,梯度计算方法可表示为

$$G_x(x,y) = H(x+1,y) - H(x-1,y), \quad (1)$$

$$G_y(x,y) = H(x,y+1) - H(x,y-1), \quad (2)$$

$$G(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)}, \quad (3)$$

$$\alpha(x,y) = \arctan\left[\frac{G_x(x,y)}{G_y(x,y)}\right], \quad (4)$$

式中: $G_x(x,y)$ 、 $G_y(x,y)$ 、 $H(x,y)$ 分别表示输入图像中像素点 $(x,y)$ 处的水平方向梯度、垂直方向梯度和像素值。提取过程如图6所示。

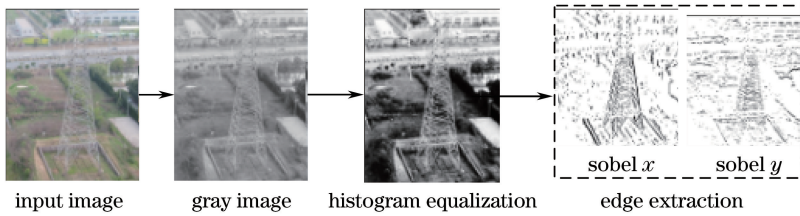


图6 各像素点梯度的计算

Fig. 6 Calculation of gradient for each pixel point

图7给出梯度的角度分布,在得到水平和垂直方向大小为 $M\times N$ 的边缘图像之后,将其分成若干个网格(cell),每个网格中包含 $8\text{ pixel}\times 8\text{ pixel}$ 。考

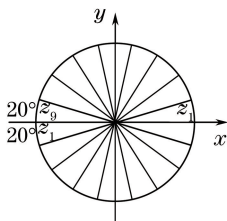


图7 梯度的角度分布

Fig. 7 Angular distribution of gradient

考虑梯度方向,将cell内部的梯度方向分为9个方向块,每 $20^\circ$ 分为一个区间,即采用9个bins( $Z_1\sim Z_9$ 区块)的直方图来统计上述 $8\text{ pixel}\times 8\text{ pixel}$ 内的梯度信息。对cell内每个像素的梯度方向的幅值在直方图上进行加权,将加权结果映射到对应的角度范围块内,得到该cell的梯度方向直方图,即为该cell对应的9维特征向量。区域HOG特征直方图的描述如图8所示。每个不同网格内的直方图梯度特征有所不同,把各个cell单元组合成空间上连通的区域(block),一个block内所有cell的特征向量串联,得到该block的HOG特征。这些区间是互有重叠

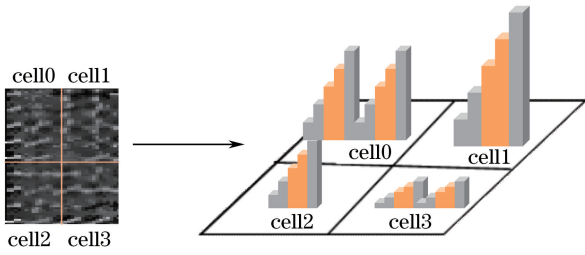


图 8 每个 cell 中的梯度统计直方图

Fig. 8 Gradient statistical histogram in each cell

的, 每一个单元格的特征会以不同的结果多次出现在最后的特征向量中, 通过这种归一化的处理可得到该 block 描述向量, 即杆塔区域特征的 HOG 描述符。通常使用的 HOG 结构大致有 3 种: 矩形 HOG、圆形 HOG 和中心环绕 HOG。Dalal 等的实验证明矩形 HOG 和圆形 HOG 的检测效果基本一致, 而环绕形 HOG 效果相对较差<sup>[15]</sup>, 因此本文采用矩形 HOG 方式提取杆塔局部位置的 HOG 特征的向量描述。

### 3.2 多层感知机方位分类

MLP 的结构如图 9 所示, 包含输入层、隐藏和输出层。设  $n$  为特征维度,  $w$  为权重,  $b$  为偏置量, 在训练时, 先前向传播。设  $(x_1, x_2, \dots, x_n)$  为提取的杆塔 HOG 特征, 中间层为隐藏层, 每层神经元与下一层是全连接的, 前向传播计算神经元之间的权重和偏置, 隐藏层的输出为  $f(w \times x + b)$ , 输出层采用 softmax 损失函数输出分类结果  $f_{\text{softmax}}[f(w \times x + b)]$ 。利用输出值和样本值计算总误差, 使用后向传播(BP)算法来更新各个神经元的权重参数, 获得损失最小时的权重和偏置。检测阶段输出的杆塔 ROI(region of interest)作为分类阶段的输入值, 提取 HOG 特征, 得到一个表示该 ROI 内杆塔方位梯度的 HOG 特征向量。将该特征向量输入训练好的 MLP 分类器, 判断区域中杆塔所对应的杆塔方位。

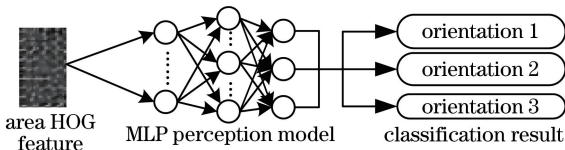


图 9 HOG 特征提取和 MLP 感知模型分类

Fig. 9 HOG feature extraction and MLP perception model classification

## 4 杆塔细节巡检

一般地, 无人机将遥控信号提供的方位值  $x_d$ 、

$y_d, z_d, \psi_d$  作为机体控制的期望输入, 进而实现对无人机的位置及姿态的控制。相机将遥控信号提供的  $\varphi_c, \theta_c$  作为机载相机控制的期望输入, 进而实现拍摄角度的控制<sup>[16]</sup>。其中,  $x_d, y_d, z_d$  表示机体坐标系下无人机期望位置;  $\psi_d$  为无人机的期望方位角, 即机体轴在水平面上的投影与正北方向的夹角;  $\varphi_c$  为横滚角, 即机体轴在水平面上投影与竖直地轴之间的夹角;  $\theta_c$  为俯仰角, 即机体轴与水平面之间的夹角。本文使用王亚先等<sup>[17]</sup>提出的基于图像的视觉伺服方法实现无人机姿态的调整及对电力线杆塔连接区域的拍摄。

设  $r(t)$  表示系统的期望值,  $y(t)$  表示系统的实际输出值,  $e(t)$  表示期望值与输出值之间的误差,  $u(t)$  表示系统的控制量,  $t$  为时间。比例-积分-微分(PID)控制的调节机制通过 PID 电路对误差  $e(t)$  进行调节来获得被控对象所需的控制量, 即

$$u(t) = K_P e(t) + K_I \int_0^t e(t) dt + K_D \frac{\Delta e(t)}{\Delta t}, \quad (5)$$

式中:  $K_P, K_I, K_D$  分别为比例、积分和微分调节系数。以当前杆塔图像和目标杆塔方位信息作为输入, 根据王亚先等<sup>[17]</sup>提出的视觉定位算法确定无人机的飞行控制状态参量替代机体状态反馈值, 构成基于无人机的视觉控制。确定期望拍摄角度  $\varphi_c, \theta_c$ , 替代遥控信号提供的角度值, 构成基于视觉的相机拍摄控制。在由无人机位置控制器、姿态控制器及电机控制器组成的位置控制回路中, 由 PID 控制算法<sup>[18]</sup>可知位移加速度与位移误差量和速度误差量之间存在一定的线性关系。基于视觉的相机拍摄角度控制方法可表示为

$$\ddot{\theta} = K_{P\theta}(\theta_c - \theta_d) + K_{I\theta} \int_0^t (\theta_c - \theta_d) dt + K_{D\theta} \frac{\Delta(\theta_c - \theta_d)}{\Delta t}, \quad (6)$$

$$\ddot{\varphi} = K_{P\varphi}(\varphi_c - \theta_d) + K_{I\varphi} \int_0^t (\varphi_c - \varphi_d) dt + K_{D\varphi} \frac{\Delta(\varphi_c - \theta_d)}{\Delta t}, \quad (7)$$

式中:  $\ddot{\varphi}, \ddot{\theta}$  为无人机姿态角加速度;  $K_{P\theta}, K_{I\theta}, K_{D\theta}$  为位置控制回路中比例、积分和微分调节系数;  $K_{P\varphi}, K_{I\varphi}, K_{D\varphi}$  为姿态控制回路中比例、积分和微分调节系数。拍摄步骤为: 1) 在实现杆塔方位全局识别的基础上, 根据连接点的结构特征识别出局部连接区域; 2) 由特写拍摄方式下局部巡检点在图像坐标上的物理变化来计算局部定位的拍摄角度; 3) 根据获得的拍摄角度来完成局部定位, 实现对杆塔和电力线连接部位图像的拍摄。

## 5 实验验证

所提算法通过 QT 编写的 C++ 代码实现,并在 CPU 为 2 GHz 四核 ARM A57,内存为 8 G 的 NVIDIA Jetson TX2 上运行。搭建如图 10(a)所示的无人机视觉感知平台实验,实验选用的操作系统是 ubuntu16.04。

### 5.1 数据集准备

数据集来源于无人机高清摄像头采集的图像,季节上包含了春夏秋冬 4 个季节。图像原始分辨率

为 4096 pixel $\times$ 2160 pixel,所以先将图像分辨率降低至 1024 pixel $\times$ 540 pixel,再从原始图像中裁剪出目标或背景区域。样本图像大小为 64 pixel $\times$ 128 pixel,制作成自己的数据集,最终得到 1500 张训练样本(包含 150 张背景、1350 张 3 个部位各 3 类方位的杆塔图像,其中每个部位有 150 张图像),900 张杆塔图像(每类 100 张图像)和 100 张背景图像用于训练,剩下每类 50 张图像用于测试,得到如图 10(b)~(d)所示的数据集。

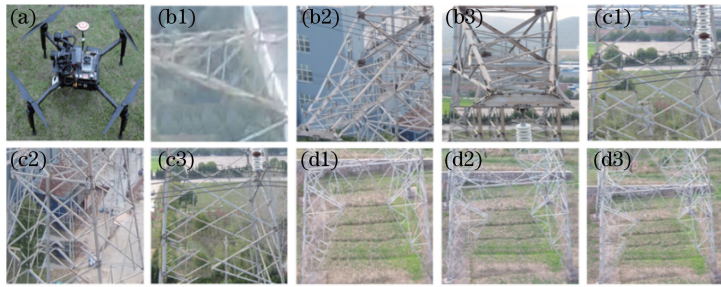


图 10 无人机视觉感知平台及在不同方位下杆塔局部区域。(a)无人机视觉感知平台;(b1)正面塔顶;(b2)正侧面塔顶;(b3)侧面塔顶;(c1)正面塔身;(c2)正侧面塔身;(c3)侧面塔身;(d1)正面塔底;(d2)正侧面塔底;(d3)侧面塔底

Fig. 10 Visual perception platform for UAV and local areas of tower under different orientations. (a) Visual perception platform for UAV; (b1) front of tower top; (b2) front side of tower top; (b3) side of tower top; (c1) front of tower body; (c2) front side of tower body; (c3) side of tower body; (d1) front of tower bottom; (d2) front side of tower bottom; (d3) side of tower bottom

### 5.2 评价指标

实验室先前已经采用了 Faster RCNN 目标检测网络对杆塔识别和检测,本文则采用基于 HOG 特征的杆塔方位检测方法,并从准确率、识别速度两个维度对两种方法进行比较。准确率为目标方位类别识别正确的样本数目与该类所有样本的总数之比;识别速度为样本从输入网络到输出识别的方位结果所需要的时间。实验过程中,分别记录每一类方位识别的准确率和识别时间。

### 5.3 实验结果对比分析

实验使用 caffe 框架实现 Faster RCNN 卷积神经网络,使用 5.1 节中构建的数据集对 Ren 等<sup>[7]</sup>在 ImageNet<sup>[19]</sup>上预训练好的模型进行微调。实验所设置的网络参数与 Ren 等<sup>[7]</sup>使用的参数典型值基本一致,即 dropout 比例为 0.6,非极大值抑制(NMS)前后 proposal 提议区域数分别为 2000 和 300,anchor 数为 9,学习率为 0.001,权重衰减率为 0.0005,批处理尺寸为 128,最终分类输出类别为 3 类。MLP 网络为 3 层,输入层为最终提取到的 HOG 特征,隐藏层有 6 个神经元,输出层有 3 个神经元,使用反向传播方法训练网络模型。使用

5.1 节中准备的数据集进行训练、测试。

如表 1 所示,使用 Faster RCNN 对杆塔局部方位进行分类,分别用 ZF-Net(Zeiler Fergus Net)、VGG-16(Visual Geometry Group Net)和 ResNet-101(Residual Net)作为基础特征提取网络,本文所提的方法与 ResNet-101+Faster RCNN 的识别准确率最高,二者比较接近。后者得益于 ResNet-101 的深层卷积神经网络结构能够提取到较为高层的语义特征,经全连接层输出的高维度特征能够对目标进行较好的表示。而本文所提算法可针对性地提取杆塔局部区域的 HOG 特征,该梯度统计特征可更好地对杆塔自身的线性钢结构特征进行描述,能够有效地区分出同一区域不同方位下的杆塔方位特征。结合 MLP 浅层网络来对杆塔方位进行分类识别,本文所提方法同样获得较高的识别准确率。本文所提算法匹配准确率均高于其他算法,可达到 87.9%,分别比 ZF-Net、VGG-16 和 ResNet-101 的平均准确率高了 27.9%、17%和 11.1%。

如表 2 所示,使用 ZF-Net 浅层卷积神经网络虽然识别时间较短,但识别准确度较差,无法实际应用。理论上,不同层学习到的特征具有层次结

表1 杆塔局部区域方位识别的准确率

Table 1 Accuracy rate of identification of local area orientation of tower

Algorithm	Fig. 10 (b)			Fig. 10 (c)			Fig. 10 (d)		
	Front	Front side	Side	Front	Front side	Side	Front	Front side	Side
ZF-Net+Faster RCNN	0.682	0.691	0.702	0.618	0.699	0.664	0.677	0.686	0.629
VGG16+Faster RCNN	0.740	0.724	0.743	0.641	0.756	0.751	0.765	0.731	0.763
ResNet-101+Faster RCNN	0.772	0.780	0.788	0.693	0.801	0.793	0.787	0.759	0.789
HOG+MLP(Proposed)	<b>0.826</b>	<b>0.818</b>	<b>0.859</b>	<b>0.879</b>	<b>0.891</b>	<b>0.902</b>	<b>0.878</b>	<b>0.869</b>	<b>0.821</b>

构,层数越深,其特征不变性越强、分类能力越强。例如,VGG-16等的深层网络可提取出具有方位表示性的不变性特征,但这种通过提高基础网络的复杂度来更好地学习特征的方式会使得网络的复杂度增大、分类识别时间变长,难以满足实际应用的要求。此外,这种方式的深度神经网络容易出

现梯度消失或梯度爆炸问题,造成梯度不稳定,难以训练。本文提出的HOG特征结合3层感知机分类方式,可较好地对不同杆塔的区域进行分析和描述,且复杂度较低,分别比ZF-Net、VGG-16和ResNet-101方法的平均处理时间缩短了70.6%、93.8%和94.1%。

表2 杆塔局部区域方位识别所需时间

Table 2 Time required for identification of local area orientation of tower

ms

Algorithm	Fig. 10 (b)			Fig. 10 (c)			Fig. 10 (d)		
	Front	Front side	Side	Front	Front side	Side	Front	Front side	Side
ZF-Net +Faster RCNN	31	33	29	26	30	37	29	33	31
VGG16+Faster RCNN	140	159	154	145	148	143	141	153	157
ResNet-101+Faster RCNN	172	175	168	169	172	174	176	167	168
HOG+MLP(Proposed)	<b>8</b>	<b>10</b>	<b>9</b>	<b>11</b>	<b>9</b>	<b>10</b>	<b>8</b>	<b>9</b>	<b>8</b>

## 5.4 数据增广

对5.1节中的数据集作镜像、加高斯噪声处理,扩大样本集,使用增广后的数据集再次对Faster RCNN模型进行微调,对MLP模型进行训练。训练参数与5.4节一致。使用测试集对模型进行重新测试,结果如表3和4所示,本文所提方法的识别准确率均用加粗表示。

由表3和4可知,基于ZF-Net、VGG16和ResNet-101的Faster RCNN方法的识别准确率有所波动,并且由于整体网络结构未发生调整,FasterRCNN方法的识别速度基本不变。本文所提方法与上述方法在平均准确率上分别提高32.6%、16.5%和10.7%,平均处理耗时分别缩短了71.9%、94.2%和94.9%。

表3 数据增广后杆塔局部区域方位识别的准确率

Table 3 Accuracy rate of local area orientation of tower after data augmentation

Algorithm	Fig. 10 (b)			Fig. 10 (c)			Fig. 10 (d)		
	Front	Front side	Side	Front	Front side	Side	Front	Front side	Side
ZF-Net+Faster RCNN	0.653	0.671	0.708	0.621	0.652	0.634	0.650	0.619	0.639
VGG16+Faster RCNN	0.732	0.738	0.749	0.661	0.754	0.752	0.762	0.725	0.778
ResNet-101+Faster RCNN	0.781	0.783	0.798	0.701	0.801	0.808	0.787	0.762	0.789
HOG+MLP(Proposed)	<b>0.833</b>	<b>0.828</b>	<b>0.855</b>	<b>0.876</b>	<b>0.895</b>	<b>0.893</b>	<b>0.874</b>	<b>0.866</b>	<b>0.834</b>

表4 数据增广后杆塔局部区域方位识别所需时间

Table 4 Time required for local area orientation of tower after data augmentation

ms

Algorithm	Fig. 10 (b)			Fig. 10 (c)			Fig. 10 (d)		
	Front	Front side	Side	Front	Front side	Side	Front	Front side	Side
ZF-Net+Faster RCNN	31	30	29	27	32	33	30	34	27
VGG16+Faster RCNN	139	148	140	139	148	141	143	142	149
ResNet-101+Faster RCNN	168	171	163	165	177	168	179	165	161
HOG+MLP(Proposed)	<b>10</b>	<b>7</b>	<b>6</b>	<b>8</b>	<b>9</b>	<b>11</b>	<b>9</b>	<b>9</b>	<b>8</b>

## 6 结 论

无人机在行业领域的拍摄内容主要是野外建造的人造设施,拍摄方式需以精准路径及角度来特写人造设施。研究了一种由远及近的杆塔检测方法:1)基于杆塔地理位置信息和无人机地理位置信息建立无人机的控制规则;2)分析杆塔自身及其位置方位特点,并结合杆塔部件区域在不同方位下的不同梯度特征,构建出杆塔区域 HOG 特征描述符;3)使用 MLP 模型对 HOG 特征进行分类,并识别杆塔区域的方位。实验表明,在无人机执行巡检任务时,通过提取出不同方位下杆塔区域的不同 HOG 特征描述,可实现对杆塔的快速巡检。本文所提出的方法在检测准确率上较 Faster RCNN 方法最大提高了 27.9%,运算时间缩短 70.6%,并且具有一定的普适性。但是,对于无人机在识别杆塔局部区域方位后如何更稳健地用于视觉伺服,本文并没有很好地解决,所以,下一步研究的重点是杆塔区域方位与无人机的视觉反馈控制。

## 参 考 文 献

- [1] Cheng W, Song Z. Power pole detection based on graph cut[C] // 2008 Congress on Image and Signal Processing, May 27-30, 2008, Sanya, Hainan, China. New York: IEEE, 2008, 3: 720-724.
- [2] Han B, Shang F. A frame model of power pylon detection for UAV-based power transmission line inspection[J]. Zhejiang Electric Power, 2016, 35(4): 6-11.  
韩冰, 尚方. 面向无人机输电线路巡检的电力杆塔检测框架模型[J]. 浙江电力, 2016, 35(4): 6-11.
- [3] Cetin B, Bikdash M, McInerney M. Automated electric utility pole detection from aerial images[C] // IEEE Southeastcon 2009, March 5-8, 2009, Atlanta, GA, USA. New York: IEEE, 2009: 44-49.
- [4] Golightly I, Jones D. Corner detection and matching for visual tracking during power line inspection[J]. Image and Vision Computing, 2003, 21(9): 827-840.
- [5] Cerón A, Mondragón I, Prieto F. Real-time transmission tower detection from video based on a feature descriptor[J]. IET Computer Vision, 2016, 11(1): 33-42.
- [6] Hui X L, Bian J, Yu Y J, *et al.* A novel autonomous navigation approach for UAV power line inspection [C] // 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO), December 5-8, 2017, Macau, China. New York: IEEE, 2017: 634-639.
- [7] Ren S Q, He K M, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [8] Henriques J F, Caseiro R, Martins P, *et al.* High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583-596.
- [9] Hui X L, Bian J, Zhao X G, *et al.* Deep-learning-based autonomous navigation approach for UAV transmission line inspection [C] // 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), March 29-31, 2018, Xiamen, China. New York: IEEE, 2018: 455-460.
- [10] Lei J, Fu J P, Zhang P L. Research on flaw classification of rifle cannon bore image[J]. Laser & Optoelectronics Progress, 2011, 48(12): 121002.  
雷洁, 傅建平, 张培林. 线膛火炮内膛疵病图像分类研究[J]. 激光与光电子学进展, 2011, 48(12): 121002.
- [11] Xu Y, Wei Z N. An improved traffic sign image recognition algorithm [J]. Laser & Optoelectronics Progress, 2017, 54(2): 021001.  
徐岩, 韦镇余. 一种改进的交通标志图像识别算法[J]. 激光与光电子学进展, 2017, 54(2): 021001.
- [12] Potena C, Nardi D, Pretto A. Effective target aware visual navigation for UAVs [C] // 2017 European Conference on Mobile Robots (ECMR), September 6-8, 2017, Paris, France. New York: IEEE, 2017: 1-7.
- [13] Han J, Zhang J J, Wang B H. Method on recognizing the structure of transmission line based on perceptual organization [J]. Infrared and Laser Engineering, 2013, 42(12): 3458-3463.  
韩军, 张晶晶, 王滨海. 基于感知组织的输电线路结构识别方法[J]. 红外与激光工程, 2013, 42(12): 3458-3463.
- [14] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), June 20-25, 2005, San Diego, CA, USA. New York: IEEE, 2005: 886-893.
- [15] Dalal N. Finding people in images and videos [D].



- Grenoble: Institut National Polytechnique de Grenoble-INPG, 2006: 1-8.
- [16] Wu X L, Shi Z Y, Zhong Y S. An overview of vision-based UAV navigation[J]. Journal of System Simulation, 2010, 22(s1): 62-65.  
吴显亮, 石宗英, 钟宜生. 无人机视觉导航研究综述[J]. 系统仿真学报, 2010, 22(s1): 62-65.
- [17] Wang Y X. Research on visual ranging and visual positioning based on electric power inspection[D]. Shanghai: Shanghai University, 2017: 17-27.  
王亚先. 基于电力巡检的视觉测距与视觉定位的研究[D]. 上海: 上海大学, 2017: 17-27.
- [18] Xue W L. Research of quadrotor UAV flight attitude control [D]. Shenyang: Shenyang University of Technology, 2016: 6-29.  
薛伟霖. 四旋翼无人机姿态控制的研究[D]. 沈阳: 沈阳工业大学, 2016: 6-29.
- [19] Deng J, Dong W, Socher R, *et al.* ImageNet: A large-scale hierarchical image database [C] // 2009 IEEE Conference on Computer Vision and Pattern Recognition, June 20-25, 2009, Miami, FL, USA. New York: IEEE, 2009: 248-255.