

# 基于深度学习的四旋翼无人机单目视觉避障方法

张午阳<sup>1,2</sup>, 章伟<sup>1</sup>, 宋芳<sup>1\*</sup>, 龙林<sup>2</sup>

(1. 上海工程技术大学 机器人智能控制实验室, 上海 201620; 2. 上海工程技术大学 机械与汽车工程学院, 上海 201620)

(\* 通信作者电子邮箱 songfang@sues.edu.cn)

**摘要:** 针对无人机避障问题, 提出一种基于深度学习的四旋翼无人机单目视觉避障方法。首先通过目标检测框选出目标在图像中的位置, 并通过计算目标选框上下边距的长度, 以此来估量出障碍物到无人机之间的距离; 然后通过协同计算机判断是否执行避障动作; 最后使用基于 Pixhawk 搭建的飞行实验平台进行实验。实验结果表明, 该方法可用于无人机低速飞行条件下避障。该方法所用到的传感器只有一块单目摄像头, 而且相对于传统的主动式传感器避障方法, 所占用无人机的体积大幅减小。该方法鲁棒性较好, 能够准确识别不同姿态的人, 实现对人避障。

**关键词:** 深度学习; 目标检测; 单目视觉; 无人机避障

**中图分类号:** TP242 **文献标志码:** A

## Monocular vision obstacle avoidance method for quadcopter based on deep learning

ZHANG Wuyang<sup>1,2</sup>, ZHANG Wei<sup>1</sup>, SONG Fang<sup>1\*</sup>, LONG Lin<sup>2</sup>

(1. Laboratory of Intelligent Control and Robotics, Shanghai University of Engineering Science, Shanghai 201620, China;

2. College of Mechanical and Automobile Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

**Abstract:** A monocular vision obstacle avoidance method for quadrotor based on deep learning was proposed to help quadrotors to avoid obstacles. Firstly, the position of object in the image was obtained by object detection, and by calculating the height of the object box in the image, the distance between quadcopter and obstacle was estimated. Then, whether performing obstacle avoidance was determined by synergetic computer. Finally, experiments were conducted on a flight test platform based on Pixhawk flight control board. The results show that the proposed method can be applied to quadcopter obstacle avoidance with low speed. Compared with traditional active sensor methods, the proposed method greatly reduces the occupied volume with only one monocular camera as sensor. This method is robust and can identify people with different postures as obstacles.

**Key words:** deep learning; object detection; monocular vision; quadcopter obstacle avoidance

## 0 引言

随着机器人技术以及微机电系统(Micro-Electro-Mechanical System, MEMS)的发展, 无人机在各领域中的应用越来越广泛。围绕着无人机相关的研究也日趋增多, 其中, 无人机的自主飞行技术一直是研究热点之一。四旋翼无人机具有体积小、结构简单、机动性强、可垂直起降等固定翼无人机所不具备的优点, 因此具有很强的适应性和广泛的应用场景, 可用在航拍、巡检、物流等低空域任务当中, 很多科研机构 and 高校均对此展开了不同层次的深入研究。

对于固定翼无人机, 因其飞行高度较高, 在其任务高度上基本没有障碍物阻拦, 可以人为地为其划定航线, 让其沿着预定航线在预定高度飞行。然而, 四旋翼无人机执行任务时, 受限于通信距离以及续航时间, 其飞行高度一般较低, 飞行路径上可能会分布着各式各样的障碍物, 比如楼房、树木、人类等, 若四旋翼无人机不具有避障功能, 则会沿着预定航线与障碍物相撞, 造成财产损失甚至造成人员伤亡事故。这是一个客观存在而且棘手的问题, 本文致力于提出此类问题的解决方案。

目前, 有很多科研机构 and 高校都对小型四旋翼无人机的避障问题展开了深入研究, 并提出了很多行之有效的解决方案<sup>[1-6]</sup>。若要实现避障, 首先要使无人机感知到障碍物的方位, 文献[6]提出了一种基于 RealSense 相机的避障方式, 通过 RealSense 相机获得当前帧的景深图像, 通过对景深图像的分层和分割, 获得障碍物的方向和距离信息, 以此达到感知障碍物的目的。在感知到障碍物的空间相对位置后, 无人机需要对任务路径进行重新规划, 针对这一问题, 文献[7]提出了一种基于分布感知与决策的避障算法, 该算法将避障问题分为了局部决策和全局决策两部分, 这样做不仅避免了无人机与障碍物碰撞, 同时也能够在任务区域内规划出一条最优路径。为了降低无人机避障系统的成本, Gageik 等<sup>[8]</sup>提出了一种成本较低的避障方案, 该方案在尽量减少冗余模块的基础上, 作了一定的整合优化, 让无人机实现障碍物检测和避障。除此之外, 也可以通过同步定位与地图构建(Simultaneous Localization and Mapping, SLAM)<sup>[9-11]</sup>技术、并行跟踪与地图构建(Parallel Tracking and Mapping, PTAM)<sup>[12]</sup>技术等实现无人机避障功能。

传统的避障方式是通过高精度传感器感知障碍物的空间

收稿日期: 2018-09-20; 修回日期: 2018-11-15; 录用日期: 2018-11-26。 基金项目: 国家自然科学基金资助项目(51505273)。

作者简介: 张午阳(1994—), 男, 安徽淮北人, 硕士研究生, 主要研究方向: 无人机避障、计算机视觉; 章伟(1977—), 男, 安徽桐城人, 副教授, 博士, 主要研究方向: 非线性控制与观测器、鲁棒控制、无人机控制; 宋芳(1980—), 女, 黑龙江大庆人, 副教授, 博士, 主要研究方向: 机器人、运动控制; 龙林(1976—), 男, 安徽桐城人, 硕士, 主要研究方向: 智能化控制、物联网。

相对位置,然后据此执行避障动作<sup>[7-9,12]</sup>。此种避障方式精度高,但是由于所依赖的传感器体积较大,导致无人机载荷明显增大,其续航能力将会明显下降;并且其无法对识别出的障碍进行分类,即其并不知道前方障碍物到底为何物,这样,无人机就无法根据不同障碍物类型作出最优的规避动作。近年来,随着图像处理及计算机视觉技术的发展,基于机器视觉的无人机避障逐渐成为当前无人机领域的一个研究热点<sup>[4-6,13]</sup>。图像采集和处理设备体积小,非常适合无人机在任务中携带。传统的视觉识别是通过特征提取方法来处理图像信息,然而,它会受到光照、物体摆放位置等外部因素影响,鲁棒性较差,故这种方法适合于特定条件下对特定物体进行识别。

近年来,深度学习发展迅速<sup>[14-16]</sup>,其优势在于不需要人为地指定需要提取的特征值,而是使用大数据去训练得到模型,使其自己学习总结目标所具备的特征<sup>[17]</sup>。

文献[18]使用基于深度学习的目标检测算法对机械臂将要抓取之物进行检测分类,并通过实验验证了其有效性,其泛化性和鲁棒性较好。受其启发,本文提出一种基于深度神经网络的四旋翼无人机避障方法,通过训练集对模型进行训练,使其能够辨认出一定数量的障碍物,并能识别前方障碍的类型,即具有障碍物分类的功能。通过机载协同计算机将目标障碍物识别并框选出来,得到目标障碍物在图像中的位置,进而计算出目标障碍物的图上尺寸,再利用相似三角形原理估量出目标障碍物到无人机之间的间隔,再由机载协同计算机判断是否执行避障动作。该方法能应对各类复杂场景中的目标检测任务,例如目标被遮挡、亮度变化、目标姿态变化等。系统结构及算法流程如图1所示。

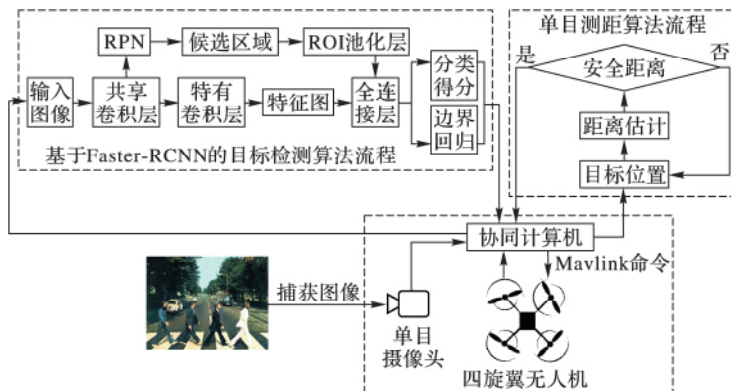


图1 系统结构及单目测距算法流程

Fig. 1 System structure and flow chart of monocular ranging algorithm

## 1 基于深度学习的目标检测

在目标检测领域,卷积神经网络(Convolutional Neural Network, CNN)具有优异的特性,其权重共享网络结构能够降低网络模型的复杂性和权重数量<sup>[19-20]</sup>。CNN的优势是识别准确率高,然而大量的卷积运算使其运行效率偏低,实时性较差。为了克服这一缺陷,区域卷积神经网络(Regional-CNN, R-CNN)应运而生<sup>[21]</sup>。R-CNN算法分为四个步骤:首先将一张图像划分为2000~3000个候选区域(Region Proposals),然后利用CNN提取每个候选区域的特征,再训练SVM分类器对这些特征进行分类,最后通过边界框回归算法重新定位目标边界框。虽然R-CNN克服了传统方法的缺点,但是其运算效率依然不高,因为其对于每个候选区域都要重新计算整个网络。快速R-CNN(fast R-CNN)则可以较好地克服以上缺

点<sup>[22]</sup>。相对于R-CNN,fast R-CNN主要优势在于提出了ROI(Region Of Interests)策略,将候选区域映射到CNN模型的特征层上,直接在特征层上提取对应区域的深层特征,避免了不断输入不同区域图像的部分,克服了R-CNN提取卷积特征时冗余操作的缺点<sup>[22]</sup>;但其实时性并未得到很大的提升。因此Ren等<sup>[23]</sup>在fast R-CNN的基础上提出了Faster R-CNN,其模型框架如图2所示。考虑到无人机在飞行过程中需要对障碍物进行实时检测并作出相应的规避动作,本文采用Faster R-CNN作为目标检测模型。

作为fast R-CNN的改进,Faster R-CNN引入了区域生成网络(Region Proposal Network, RPN)的概念,RPN模型如图3所示。在RPN中,损失函数为分类损失(Classify Loss)和回归损失(Regression Loss)的联合损失<sup>[23]</sup>:

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

其中: $i$ 表示第 $i$ 个锚点; $p_i$ 表示锚点 $i$ 为目标的概率; $p_i^* = 1$ 时表示第 $i$ 个锚点为正样本; $t_i$ 表示预测的边界框的4个参数化坐标向量; $t_i^*$ 表示候选区域边界和真实目标边界之间的偏差; $L_{cls}$ 表示分类损失; $L_{reg}$ 表示回归损失; $N_{cls}$ 和 $N_{reg}$ 分别表示分类损失和回归损失的归一化参数; $\lambda$ 表示平衡权重。

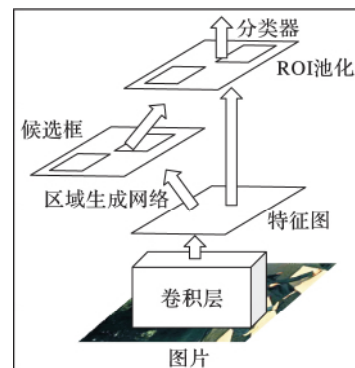


图2 Faster R-CNN 模型框架

Fig. 2 Framework of Faster R-CNN model

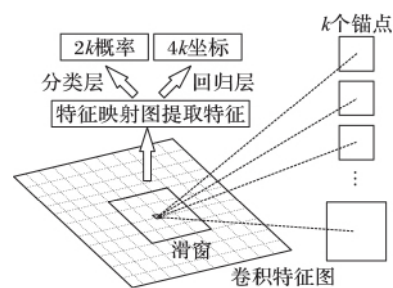


图3 RPN 模型

Fig. 3 RPN model

Faster R-CNN的数据分为训练集和测试集两部分,分别记为:

$$\left\{ \begin{aligned} \{x_n^{Tr}, y_n^{Tr}\}_{n=1}^{N_{Tr}} \\ \{x_n^{Te}, y_n^{Te}\}_{n=1}^{N_{Te}} \end{aligned} \right. \quad (2)$$

其中: $Tr$ 表示训练集; $Te$ 表示测试集; $x_n \in \mathbf{R}^{n \times m}$ 为输入; $y_n$ 为输出; $y_n$ 包括两部分:一是场景中所有目标的位置,二是目标区域所对应的物体类别。可写为:

$$y_n = [P_{x_n-tar}^{(s)} \in \mathbf{R}^4, L_{x_n-tar}^{(s)} \in \{1, 2, \dots, C\}] ; \quad s = 0, 1, 2, \dots, S \quad (3)$$

其中:  $S$  为场景  $x_n$  中目标的个数;  $C$  为场景中的目标种类;  $P_{x_n-tar}^{(s)}$  和  $L_{x_n-tar}^{(s)}$  分别表示第  $s$  个目标对应的坐标和类标。

对于区域生成网络, 输入与输出之间存在以下关系:

$$\begin{cases} X_1 = \text{ConvNet}^{(1)}(x, \theta_1) \in \mathbf{R}^{u \times v \times r} \\ RP_x = [RP_x^1 \in \mathbf{R}^{2 \times S}, RP_x^2 \in \mathbf{R}^{4 \times S}] = \\ \quad \text{RPN}(X_1(x), \theta_{RP}) \end{cases} \quad (4)$$

其中:  $x \in \mathbf{R}^{n \times m}$  为输入场景;  $X_1 \in \mathbf{R}^{u \times v \times r}$  为经过 CNN 提取特征值之后的输出, 即 RPN 的输入, 有  $r$  个特征图, 每个尺寸为  $u \times v$ ;  $\theta_1$  和  $\theta_{RP}$  为待学习的参数;  $x$  为输入图像;  $RP_x$  为输出;  $S$  为候选目标区域的个数;  $RP_x^1$  为每个候选区域判断为目标区域的得分;  $RP_x^2$  为候选目标区域的位置。

结合 RPN 的输出, Faster R-CNN 网络的输入和输出之间存在如下关系:

$$\begin{cases} X_1 = \text{ConvNet}^{(1)}(x, \theta_1) \in \mathbf{R}^{u \times v \times r} \\ X_2 = \text{ConvNet}^{(2)}(X_1, \theta_2) \\ X_3 = \text{ROI}^{(3)}(X_2, RP_x) \\ y = [RP_x^{(s)}, L_x^{(s)}]; \hat{s} = 0, 1, \dots, \hat{S} \end{cases} \quad (5)$$

其中:  $X_2$  为经过特有卷积层处理后的特征图, 由  $X_1$  得到;  $\theta_2$  为待学习参数;  $X_3$  为 ROI 池化层输出;  $y$  为预测输出, 包括目标区域的位置和类别两部分。

Faster R-CNN 网络优化目标函数有两个:

1) RPN 优化通路。优化目标函数为:

$$\min_{\theta_1, \theta_{RP}} J(\theta_1, \theta_{RP}) = \frac{1}{N} \sum_{n=1}^N L(x_n, \sigma_n; \theta_1, \theta_{RP}) + \lambda_1 R(\theta_1) + \lambda_2 R(\theta_2) \quad (6)$$

其中:  $L(x_n, \sigma_n; \theta_1, \theta_{RP})$  为风险评估损失项; 余下部分为正则项。风险评估损失项表达式形式为:

$$\begin{cases} L(x_n, \sigma_n; \theta_1, \theta_{RP}) = \frac{1}{N_{\text{cls}}} \sum_{i=1}^K L_{\text{cls}}(\hat{p}_i^{(n)}, p_i^{(n)}) + \\ \quad \frac{1}{N_{\text{reg}}} \sum_{j=1}^K L_{\text{reg}}(\hat{t}_j^{(n)}, t_j^{(n)}) \\ L_{\text{cls}}(\hat{p}_i^{(n)}, p_i^{(n)}) = -\log[\hat{p}_i^{(n)}(1) \cdot p_i^{(n)}(1) + \\ \quad (1 - \hat{p}_i^{(n)}(1)) \cdot (1 - p_i^{(n)}(1))] \\ L_{\text{reg}}(\hat{t}_j^{(n)}, t_j^{(n)}) = R(\hat{t}_j^{(n)} - t_j^{(n)}) = \\ \quad \sum_{p=1}^4 R(\hat{t}_j^{(n)}(p) - t_j^{(n)}(p)) \end{cases} \quad (7)$$

其中:  $L_{\text{cls}}(\hat{p}_i^{(n)}, p_i^{(n)})$  为区域得分损失;  $L_{\text{reg}}(\hat{t}_j^{(n)}, t_j^{(n)})$  为候选区域位置;  $p_i^{(n)}(1)$  和  $t_j^{(n)}(p)$  分别为回归器  $L_{\text{reg}}(\hat{t}_j^{(n)}, t_j^{(n)})$  和分类器  $L_{\text{cls}}(\hat{p}_i^{(n)}, p_i^{(n)})$  中的权重因子。式(7)中的回归器的非线性函数为:

$$R(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & |x| > 1 \end{cases} \quad (8)$$

2) fast R-CNN 优化通路:

$$\begin{cases} \min_{\theta} J(\theta_1, \theta_2, \theta_3, \{ \text{Re}(RP_x, \rho) \}_{c=1}^C) = \\ \quad \sum_{n=1}^N [L_{\text{cls}}(x_n, y_n; \theta) + L_{\text{loc}}(x_n, y_n; \text{Re}(RP_x))] \\ L_{\text{cls}}(x_n, y_n; \theta) = \frac{1}{S} \sum_{s=1}^S \text{loss}(\hat{L}_{x_n}^{(s)}, L_{x_n-tar}^{(s)}) \\ L_{\text{loc}}(x_n, y_n; \text{Re}(RP_x)) = \frac{1}{S} \sum_{s=1}^S \frac{1}{C} \sum_{c=1}^C R(RP_{x_n}^{(s)}, \rho_{x_n-tar}^{(s)}) \end{cases} \quad (9)$$

其中:  $\theta_1, \theta_2, \theta_3$  均为待学习参数;  $L_{\text{cls}}$  为区域分类损失;  $L_{\text{loc}}$  为目标区域的位置精修损失;  $\text{Re}(RP_{x_n})$  为每一类所对应的目标区域进行精修的参数。

采用随机梯度下降的方法进行端到端的训练, 求解优化目标函数式。

## 2 四旋翼无人机避障

通常的无人机避障方案是通过距离传感器测得障碍物到无人机之间的距离, 比如激光雷达(Lidar)<sup>[24]</sup>、超声波<sup>[25-26]</sup>等, 然后根据测得的距离判断无人机与障碍物之间的位置关系, 从而实现避开障碍物。这类主动传感器需要时刻对无人机周围的环境进行扫描, 并且其体积相对较大, 造成了大量的能量损耗, 缩短了无人机的续航时间。本文利用单目视觉方案进行目标与摄像头之间距离的估计, 属于一种软测量方案, 体积相对于传统传感器测量大幅减小, 成本也大幅缩减。

利用上文提到的 Faster R-CNN, 被测目标能够有效并完整地框选出来, 并计算选择框上下边距之间的像素尺寸。利用相似三角形的原理, 通过计算得到的目标选择框上下边距之间的像素尺寸, 以此来估计出障碍物距无人机当前位置的距离, 原理如图4所示。由相似三角形原理, 物体到摄像头之间的距离满足以下关系:

$$d = (f \cdot h) / l \quad (10)$$

其中:  $d$  表示物体到摄像头之间的实际距离;  $f$  表示摄像头的焦距, 本文采用的摄像头为定焦摄像头;  $h$  代表被检测目标的实际尺寸;  $l$  表示物体经过摄像头成像后的尺寸。

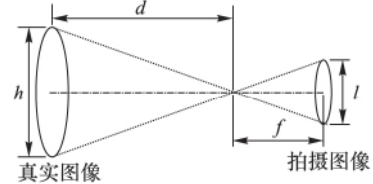


图4 相似三角形估计距离原理示意图

Fig. 4 Distance estimation using similar triangle principle

## 3 实验与结果分析

实验平台为一架四旋翼无人机, 基于 Pixhawk 飞控板搭建, 并通过 Pixhawk 的 Telem 2 接口将 Raspberry 3B+ 与其连接, 作为协同电脑。RaspberryPi 3B+ 操作系统为 Raspbian stretch 系统, 搭载一分辨率为  $640 \times 480$  的单目摄像头, 负责图像采集以及目标检测, 并通过 DroneKit-Python 控制无人机的动作(升降、俯仰、横滚以及偏航)。在对目标检测网络进行训练时, 采用的是 Windows 10 操作系统, 并使用 Nvidia GeForce GTX 1060 显卡对训练过程进行加速。Faster R-CNN 模型所依赖的深度学习框架为 Tensor flow。图5为本次实验所采用的飞行实验平台。

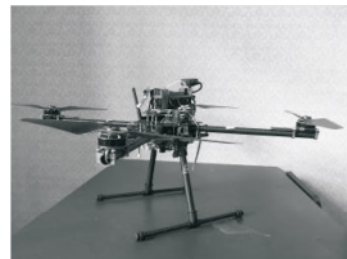


图5 基于 Pixhawk 的飞行实验平台

Fig. 5 Flight test platform based on Pixhawk



### 3.1 目标检测

因为无人机避障要优先考虑避开人体,因此本文先训练模型识别人体。将“people”作为标签,训练模型识别人体的能力。本文从互联网上随机选取了6张带有人物的照片,以此来检验模型对于人体识别的准确率。为了实现障碍物分类的功能,分别将“screen”“trash can”“laptop”“chair”等作为标签,训练模型对于障碍物分类的能力。经实验验证,本文所训练的模型识别人体的准确率均在99%以上,并且无论人体是以何种角度出现在图片中(正面、侧面或背面),模型均能将其识别出来,鲁棒性较好。对于其他类型的障碍物,比如椅子、垃圾桶等地面小型目标,也有一定的识别能力。目标检测结果如图6所示。



(a) 室外仅识别人体实验结果



(b) 实验室内人体和潜在障碍物的多目标检测实验结果

图6 基于深度学习的目标检测实验结果

Fig. 6 Detection results based on deep learning

### 3.2 距离估计

实验选取人体的参考身高为1.75 m,以此为参考,通过计算图像中目标选择框上下边框之间的长度便可估计出人体距离无人机之间的距离。考虑到无人机的反应时间以及图像传输延迟,本文将安全距离设置为1.5 m。当检测到无人机与人体之间的距离小于或等于1.5 m时,RaspberryPi 3B+将产生报警信号,并向Pixhawk发送避障指令。经对比,估计出的距离和测量得到的实际距离误差在 $\pm 0.5$  m。图像处理结果如图7所示,图中所示距离信息为摄像头到人体之间的距离估计值。

### 3.3 四旋翼无人机避障

是否执行避障动作由协同计算机 RaspberryPi 3B+ 作出判断,再通过 DroneKit-Python 向 Pixhawk 发送 Mavlink 指令。

通过实际飞行,无人机确实可以在距离人1.5 m之外避让行人,该实验验证了本文方法的有效性。避障实验结果如图8所示,其中,图8(a)为从机载协同计算机处理得到的距离估计结果,图8(b)为在两种不同的距离下进行实验的实拍图。



图7 基于深度学习的单目视觉距离检测结果

Fig. 7 Detection result of monocular visual distance estimation based on deep learning



(a) 机载图像处理模块对人体检测的结果



(b) 避障实验过程

图8 四旋翼无人机避障实验结果

Fig. 8 Experiment results of quadcopter obstacle avoidance

## 4 结语

本文利用单目摄像头,基于深度学习实现了对人体的目标检测,并将人体在图像中的位置通过框选的方式标注出来,再利用相似三角形原理估算出人体到无人机之间的近似距离。与传统避障方法相比,此方法具有占用体积小、鲁棒性好、成本低等优点。通过仿真实验验证了该方法能对人体进行高准确率识别,并且能估算出人体到无人机之间的距离,误差为 $\pm 0.5$  m。在飞行实验平台上进行了实际飞行避障实验,实验结果验证了该方法的有效性。

由于受到 RaspberryPi 3B + 机能的限制, 在进行飞行实验时, 它对人体的检测和距离估计往往会出现延迟现象。为保证实验人员的安全, 目前无人机只能进行低速飞行以抵消延迟带来的影响; 并且本文只训练了模型用于识别人体, 并不能做到行为预测, 并且在实际应用中还有许多种障碍要闪避, 这些都是未来需要研究的内容。

#### 参考文献(References)

- [1] 杨秀霞, 刘小伟, 张毅. 基于时间约束的无人机避障研究[J]. 飞行力学, 2015, 33(2): 125 - 129. (YANG X X, LIU X W, ZHANG Y. Research on obstacle avoidance of UAV based on time constraint[J]. Flight Mechanics, 2015, 33(2): 125 - 129.)
- [2] 高迪. 无人机避障雷达目标探测与跟踪算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2017: 10 - 17. (GAO D. Research on target detection and tracking algorithm for UAV obstacle avoidance radar [D]. Harbin: Harbin Institute of Technology, 2017: 10 - 17.)
- [3] 朱平, 甄子洋, 覃海群, 等. 基于立体视觉和光流的无人机避障算法研究[J]. 电光与控制, 2017, 24(12): 31 - 35. (ZHU P, ZHEN Z Y, QIN H Q, et al. Research on UAV obstacle avoidance algorithm based on stereo vision and optical flow[J]. Electronics Optics & Control, 2017, 24(12): 31 - 35.)
- [4] YANG Y, WANG T T, CEHN L, et al. Stereo vision based obstacle avoidance strategy for quadcopter UAV [C]// Proceedings of the 2018 Chinese Control and Decision Conference. Piscataway, NJ: IEEE, 2018: 490 - 494.
- [5] HU Y Y, WANG Y X. Stereo vision-based fast obstacles avoidance without obstacles discrimination for indoor UAVs [C]// Proceedings of the 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce. Piscataway, NJ: IEEE, 2011: 4332 - 4337.
- [6] HU J, NIU Y, WANG Z. Obstacle avoidance methods for rotor UAVs using RealSense camera [C]// Proceedings of the 2017 Chinese Automation Congress. Piscataway, NJ: IEEE, 2017: 7151 - 7155.
- [7] XU Z, WER R, ZHANG Q, et al. Obstacle avoidance algorithm for UAVs in unknown environment based on distributional perception and decision making [C]// Proceedings of the 2016 IEEE Chinese Guidance, Navigation and Control Conference. Piscataway, NJ: IEEE, 2016: 1072 - 1075.
- [8] GAGEIK N, BENZ P, MONTENEGRO S. Obstacle detection and collision avoidance for a UAV with complementary low-cost sensors [J]. IEEE Access, 2015, 3: 599 - 609.
- [9] RIISGAARD S, BLAS M R. SLAM for dummies: a tutorial approach to simultaneous localization and mapping [EB/OL]. [2018-05-10]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.208.6289&rep=rep1&type=pdf>.
- [10] 陈伟楠, 刘冠峰, 李俊良, 等. 室内环境的元胞自动机 SLAM 算法[J]. 机器人, 2016, 38(2): 169 - 177. (CHEN W N, LIU G F, LI J L, et al. Cellular automaton SLAM algorithm for indoor environment[J]. Robot, 2016, 38(2): 169 - 177.)
- [11] 徐伟杰, 李平, 韩波. 基于 2 点 RANSAC 的无人直升机单目视觉 SLAM[J]. 机器人, 2012, 34(1): 65 - 71. (XU W J, LI P, HAN B. Unmanned helicopter monocular vision SLAM based on 2 points RANSAC[J]. Robot, 2012, 34(1): 65 - 71.)
- [12] BAI G, XIANG X, ZHU H, et al. Research on obstacles avoidance technology for UAV based on improved PTAM algorithm [C]// Proceedings of the 2015 IEEE International Conference on Progress in Informatics and Computing. Piscataway, NJ: IEEE, 2015: 543 - 550.
- [13] CHEN Z, LUO X, DAI B. Design of obstacle avoidance system for micro-UAV based on binocular vision [C]// Proceedings of the 2017 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration. Piscataway, NJ: IEEE, 2017: 67 - 70.
- [14] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504 - 507.
- [15] ABBAS M A. Improving deep learning performance using random forest HTM cortical learning algorithm [C]// Proceedings of the 2018 First International Workshop on Deep and Representation Learning. Piscataway, NJ: IEEE, 2018: 13 - 18.
- [16] WANG Y. Cognitive foundations of knowledge science and deep knowledge learning by cognitive robots [C]// Proceedings of the 2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing. Piscataway, NJ: IEEE, 2017: 5.
- [17] GOODFELLOW I, BENGIO Y, COURVILLE A, et al. Deep learning [M]. Cambridge, MA: MIT Press, 2016: 11 - 12.
- [18] 杜学丹, 蔡莹皓, 鲁涛, 等. 一种基于深度学习的机械臂抓取方法[J]. 机器人, 2017, 39(6): 820 - 828, 837. (DU X D, CAI Y H, LU T, et al. A mechanical arm grabbing method based on deep learning[J]. Robot, 2017, 39(6): 820 - 828, 837.)
- [19] ZOU J, SONG R. Microarray camera image segmentation with faster-RCNN [C]// Proceedings of the 2018 IEEE International Conference on Applied System Invention. Piscataway, NJ: IEEE, 2018: 86 - 89.
- [20] LECUN Y, KAVUKCUOGLU K, FARABET C. Convolutional networks and applications in vision [C]// Proceedings of the 2010 IEEE International Symposium on Circuits and Systems. Piscataway, NJ: IEEE, 2010: 253 - 256.
- [21] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2014: 580 - 587.
- [22] GIRSHICK R. Fast R-CNN [C]// Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2015: 1440 - 1448.
- [23] REN S, HE K, 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.
- [24] ZHENG W, XIAO J, XIN T. Integrated navigation system with monocular vision and LIDAR for indoor UAVs [C]// Proceedings of the 2017 12th IEEE Conference on Industrial Electronics and Applications. Piscataway, NJ: IEEE, 2017: 924 - 929.
- [25] GUANGLEI M, HAIBING P. The application of ultrasonic sensor in the obstacle avoidance of quad-rotor UAV [C]// Proceedings of the 2016 IEEE Chinese Guidance, Navigation and Control Conference. Piscataway, NJ: IEEE, 2016: 976 - 981.
- [26] DIXIT K R, KRISHNA P P, ANTONY R. Design and development of H frame quadcopter for control system with obstacle detection using ultrasound sensors [C]// Proceedings of the 2017 International Conference on Circuits, Controls, and Communications. Piscataway, NJ: IEEE, 2017: 100 - 104.

This work is partially supported by the National Natural Science Foundation of China (51505273).

**ZHANG Wuyang**, born in 1994, M. S. candidate. His research interests include quadcopter obstacle avoidance, computer vision.

**ZHANG Wei**, born in 1977, Ph. D. associate professor. His research interests include nonlinear control and observer, robust control, unmanned aerial vehicle control.

**SONG Fang**, born in 1980, Ph. D. associate professor. Her research interests include robot, motion control.

**LONG Lin**, born in 1976, M. S. His research interests include intelligent control, Internet of things.