

系统综述

无人系统之“眼”——计算机视觉技术与应用浅析.....	01
张 丹, 单海军, 王 哲, 吴陈炜	
基于深度学习的单目视觉深度估计研究综述.....	12
郭继峰, 白成超, 郭 爽	

技术研究

垂面栖息微型无人机飞行原理与技术进展.....	22
昌 敏, 孙 杨, 白俊强	
无人机编队避障与控制技术研究现状及发展趋势.....	32
方洋旺, 欧阳楚月, 符文星, 程昊宇	
基于 Voronoi 图质心的多无人机协同区域搜索算法	39
朱 利, 符小卫	
无人潜航器仿生推进技术专利分析.....	52
张鸣宇, 李 峰, 周 鹏	

前沿动态分析

边缘计算的发展趋势综述.....	60
葛悦涛, 尹晓桐	
5G 在无人驾驶汽车中的应用前景分析	65
张朝阳, 李星宏	

《无人系统技术》办刊宗旨是刊载无人系统领域新进展、新成果、新技术，促进学术交流，推动成果转化，提高我国在该领域的科研装备水平。

现诚挚地向无人系统技术研究领域的专家学者、学术带头人以及致力于无人系统理论与技术创新研究的科技工作者约稿。稿件内容以综述、研究论文、技术报告、专题研讨和科技评述等形式为主，广泛报道无人空中系统、无人地面系统、无人水面水下系统、智能机器人等研究领域的理论研究成果、技术创新、学术见解和经验总结。主要包括无人系统平台技术、自动控制技术、信息技术以及人工智能、通信、传感器等技术在无人系统领域的新应用。同时，也包括无人系统自主技术、协同技术、跨域同步技术、智能无人系统伦理及其与有人系统的关系等。



Systematic Review

The Eye of Unmanned Systems:A Brief Analysis on Techniques and Applications of Computer Vision	01
ZHANG Dan, SHAN Haijun, WANG Zhe, WU Chenwei	
A Review of Monocular Depth Estimation Based on Deep Learning	12
GUO Jifeng, BAI Chengchao, GUO Shuang	

Technical Study

Flight Principles and Research Progress on Vertical-Perching Micro Aerial Vehicle (MAV)	22
CHANG Min, SUN Yang, BAI Junqiang	
Research Status and Development Trend of Obstacle Avoidance and Control Technology for UAV Formation	32
FANG Yangwang, OUYANG Chuyue, FU Wenxing, CHENG Haoyu	
Multiple UAVs Cooperative Area Search Algorithm Based on Centroid of Voronoi Diagram	39
ZHU Li, FU Xiaowei	
Patent Analysis of Bionic Propulsion Technology of Unmanned Underwater Vehicle	52
ZHANG Mingyu, LI Feng, ZHOU Peng	

Frontier Dynamic Analysis

A Survey on Edge Computing	60
GE Yuetao, YIN Xiaotong	
Analysis of the Application Prospect of 5G in Driverless Cars	65
ZHANG Zhaoyang, LI Xinghong	



无人系统之“眼” ——计算机视觉技术与应用浅析

张丹^{1,2}, 单海军¹, 王哲³, 吴陈炜¹

(1. 之江实验室, 杭州 311121; 2. 约克大学, 多伦多 M3J1P3; 3. 商汤集团有限公司, 香港 999077)

摘要: 计算机视觉 (CV) 是指赋予机器某种视觉洞察力, 让机器具备视觉感知以及认知的能力。计算机视觉是人工智能领域最重要的研究方向之一, 主要包含图像理解和视频理解两大研究方向。主要从图像识别与理解、视频识别与理解两大角度进行技术综述和应用浅析, 在图像领域重点回顾了图像分类、图像目标检测、图像分割与图像描述的研究现状, 在视频领域重点回顾了视频目标检测与识别、视频目标跟踪、视频行为识别与视频描述等领域的研究现状, 并对各自领域的研究方向进行了展望。最后简述了计算机视觉在智慧医疗、公共安全和自动驾驶等领域的应用现状。

关键词: 计算机视觉; 图像理解; 目标检测; 视频理解; 行为识别; 视觉描述; 人工智能
中图分类号: TP391.4 **文献标识码:** A **文章编号:** 2096-5915 (2019) 02-0379-11

The Eye of Unmanned Systems: A Brief Analysis on Techniques and Applications of Computer Vision

ZHANG Dan^{1,2}, SHAN Haijun¹, WANG Zhe³, WU Chenwei¹

(1. Zhejiang Lab, Hangzhou 311121, China; 2. York University, Toronto M3J1P3, Canada;
3. SenseTime Group Limited, HongKong 999077, China)

Abstract: Computer Vision (CV) refers to the ability to give the machine some visual insight, especially the visual perception and recognition. CV is one of the most important research direction in the field of artificial intelligence (AI), and it mainly includes two research directions: image understanding and video understanding. This paper mainly reviewed the research progress of computer vision from the above two directions. In the field of image understanding, the research status of image classification, image object detection, image segmentation and image caption are reviewed. In the field of video understanding, the research status of video object detection and recognition, video object tracking, video action recognition and video caption are reviewed, and the research directions in their respective fields are prospected. Finally, the applications of computer vision in intelligent medicine, public security and automatic driving are briefly introduced.

Keywords: Computer Vision; Image Understanding; Object Detection; Video Understanding; Action Recognition; Visual Caption; Artificial Intelligence

1 引言

近年来，人工智能和深度学习获得突破，成为了大众关注的焦点。如LeCun Y, Bengio Y, Hinton G等^[1]提出的深度卷积网络在图像识别领域大放异彩就是典型。计算机视觉是人工智能领域最重要的研究方向，因为视觉也是人体获得信息最多的感官来源，Treicher^[2]等通过实验证实：视觉信息占人类获得全部信息的83%。因此，计算机视觉也一直是各大研究机构和企业争相研究的热点，也是人工智能领域研究历史最长、技术积累最多的方向。本文尝试从计算机视觉的两大研究内容——图像和视频，来阐述图像识别与理解、视频识别与理解的研究历史和最新技术进展，并对计算机视觉技术在公共安防、智慧医疗、自动驾驶等领域的应用现状进行介绍。通过对技术及应用的归纳和总结，希望能为关注该领域的专家和学者提供最新且有价值的内容。

2 图像识别与理解进展综述

计算机视觉的研究主体之一是图像，目的是对图像进行各种各样的处理与分析。按照对图像理解层次的不同，一般可以把计算机视觉分为低级视觉和高级视觉，低级视觉包括图像去噪、拼接、超分辨率等传统任务。通常而言，这些任务不需要计算机理解图像上有什么，而高级视觉则明确要求计算机理解图像上有什么，以及解决基于此而引出的新问题。本文将重点对高级视觉进行介绍。

人们通常喜欢把计算机和人脑进行对比，以更好地理解 and 探讨计算机视觉技术在做什么。我们也由一个例子开始：对于一个牙牙学语的婴儿，父母会拿着一张图，问这是什么，这其实就是在教婴儿做最简单的图像分类。等婴儿大了一些，他可以指着图上的树、房子、小狗，依次说出这些是什么，聪明的孩子还可以用手指把房子的轮廓勾勒出来，这就是计算机视觉中的图像检测和

图像分割。随着大脑发育，孩子对图像理解地更深了，学会看图说话了，这也就对应了我们要介绍的图像描述任务。

本节将以图像分类、图像检测、图像分割和图像描述的顺序，由易到难地介绍计算机视觉的几个重要方面。

2.1 图像分类进展综述

图像分类是计算机视觉最基础的任务之一，也叫图像识别，主要通过算法识别图像中的物体类别，如一匹马、花朵、小狗等。该领域长期以来吸引了大量研究者的关注。2012年，由Hinton团队^[3]研发的基于神经网络的算法在ImageNet图像分类挑战赛上夺得第一，并以15.3%的错误率大幅领先第二名的26.2%，一举震惊学术界。在此之前，传统算法大多使用手工设计的特征配合当时最好的分类算法支持向量机来进行分类^[4]，这类方法最大的问题就是当分类种类变多后，人一般很难设计出相应的图像特征来表示新的类别，需要相关领域的专家花费以年为单位的时间反复试验。而Hinton团队的巨大成功让人们看到：神经网络可以通过学习大量标注后的数据以及梯度反传算法，端到端地优化整个图像识别任务，从而使得人们不再需要关注为每个新的种类设计独特的特征，而只需要设计网络结构使网络更容易被优化。

神经网络一般由一层的神经元组成，神经元之间通过权重进行关联，不同层的神经元之间还会设置激活函数，使得神经网络具有非线性的表达能力。在其后的几年至今，基于神经网络的深度学习算法几乎统治了所有计算机视觉相关领域。一个最重要的发现就是随着神经网络的层数加深，参数增多，它的表达能力越强，能拟合的数据量越大，VGG和Inception网络都是其中的典型代表。但是随着网络变深，梯度消失问题逐渐明显，较浅的神经层几乎接收不到深层网络传来的梯度，导致网络无法收敛到较好的局部最优。此后最重要的工作就是2015年何恺明等^[5]提出

的残差网络，该方法通过恒等映射解决了梯度消失的问题，并破记录的将原来几十层的网络加深到了几百层，基于该方法在多个挑战中都取得了冠军。2016年围棋人机大战中，以4:1战胜世界围棋冠军李世石的算法AlphaGo采用的也是残差网络的结构。虽然现在图像分类问题上的研究已有很多，该领域的研究也相对比较成熟。但是还面临较多的细分问题挑战，如同种类物体内的识别，图像遮挡下的物体识别与分类，视点变化和尺度变化等，这些问题都值得继续探索。

2.2 图像目标检测进展综述

图像目标检测的任务是要求算法用一个矩形框框出图上每一个物体，并对框里的物体进行分类。从任务难度上看，图像检测比图像分类增加了一个定位的功能，即需要找到图上所有目标的位置，然后再对框进行图像分类的处理。

经典的检测算法包括基于梯度直方图(HOG)特征的算法以及Deformable Part Model(DPM)。HOG特征算法主要提取图上的方向梯度直方图，并通过支持向量机来得到目标的模板。然而，对于一些非刚体、可以形变的目标来说，要用模板来穷尽其所有的姿态几乎是不可能的。DPM的解决办法就是对目标的每个部分单独建立子模型，并且建立了子模型与主模型之间的形变约束。比如人的手臂可以摆到头部附近，但是不可能离开身体两米远。这些先验的位置关系通过损失函数的形式定义出来，就可以有效建模目标的变形情况。

深度学习兴起后，区域卷积神经网络(RCNN)首次将其应用到目标检测问题上，并在检测准确率上大幅超越以往的HOG和DPM算法，成为了新的研究热点^[6]。RCNN首先通过Selective Search算法为图像生成上百个候选框，再将每个候选框送入神经网络提取特征，最后用支持向量机进行分类^[7]。RCNN在训练和测试时都有多个步骤，只有其中提取特征的部分是由神经网络完成的，之后的研究重点集中在如何将其余所有步骤

都通过神经网络来完成，从而可以端到端的优化整个系统的性能。其中，快速区域卷积神经网络(Faster-RCNN)算法提出了Region Proposal network(RPN)来进行候选框的生成，由神经网络来预测候选框的位置和大小^[8]，同时，RPN和RCNN两部分可以共享底层网络并联合优化。这个方法大大提高了网络效率，而且由网络预测的候选框质量也优于Selective Search和Edgebox这类传统算法。Faster-RCNN推理时分为两个阶段，第一阶段是RPN生成上万个候选框，网络预测框里是否有物体，按框得分筛选出可信度较高的候选框，进入第二阶段，进一步细分类。为了让网络运行更快，研究者又提出了以SSD和YOLO为代表的一阶段检测器，这些方法的特点是对所有的候选框直接进行分类，跳过了是否有目标的步骤。

2.3 图像分割进展综述

图像分割是将整个图像分成一个个像素组，并对每个像素进行标记和分类。它试图在语义上理解图像中每个像素的类别，是行人、道路、建筑物还是其他类别。因此，与图像分类不同，研究人员需要用模型对密集的像素进行预测。较早的算法一般通过对每个像素的局部领域提取纹理、光照等特征，再对该像素进行粗分类，然后基于条件随机场对分割结果进行建模。粗分类的结果由于是基于每个像素独立预测的，一般有比较大的不稳定性。条件随机场假设相邻的像素应该具有相同的类别，除非两个像素刚好落在图像梯度较大的位置。通过优化单像素的预测结果和领域约束，条件随机场可以优化出一个较平滑的分割结果。

由于神经网络可以通过数据驱动，并能自动优化参数，特征提取和分类两部分可以在网络中统一完成。只需要将以每个像素为中心选取一个邻域作为网络的输入，网络就可以较好地预测出该区块中心像素的类别，比如2013年Lecun等^[9]提出的方法是提取多尺度的图像区块进行学习，让网络从多个层级的信息中学到更全面的信息；2014

年Pinheiro等^[10]提出的基于循环神经网络来对图像区块进行预测，每次的输入是图像块以及上一轮的预测概率图，网络可以基于上一轮的预测对自己进行修正。

但是，这类方法在推理时，需要将相邻像素的重合部分反复输入网络，非常低效。2015国际计算机视觉顶会上，Darrell等^[11]提出的全卷积神经网络(FCN)因为解决此问题而获得了最佳论文的荣誉。FCN的原理就是通过卷积操作等效替换了原本分割网络中的全连接层，并且用反卷积层解决了网络特征层降采样后分辨率下降的问题。这样，FCN的输入不再是某个像素为中心的图像区块，而直接是原图像，输出则是相同大小的分割结果图。此后，几乎所有的图像分割方法都采用了FCN或其相似的原理。较有代表性的是Deeplab，将概率统计模型条件随机场通过神经网络建模，自动学习其中的参数。图像分割(UNet)在网络中添加了跳层连接，增加了多尺度的信息使预测的细节更完善。

2.4 图像描述进展综述

图像描述是指给定一张图像，算法自动生成一段描述性文字，即看图作文。图像描述是上述几个任务的能力递进与语义理解递进，它需要算法不仅要理解图上有什么物体，而且要理解物体之间的相互关系，最后要用文字将其描述出来。因此，主流的图像描述框架一般分为两步，先将图像进行编码，提取语义特征，然后再将特征通过长短期记忆网络进行解码，逐字输出描述语句。

由于图像上的物体可能有很多，但真正重要的目标只有少量，最早的方法在图像编码中引入了注意力机制，即让网络在无监督的情况下自己学习哪些空间区域是比较重要的，这样会比直接不加区分地对整张图编码要更合理，比如SCA-CNN和Show, Attend and Tell就是其中的代表。但是，由于注意力机制是靠网络自动学习的，本身的精度无法保证，并且一般来说图像描述的主语

都是图上的目标物体，因此研究者又引入了目标检测和属性分析的模型。通常可以把在专门做目标检测和属性分析的数据集上训练得到的神经网络引入到系统中，作为一种强有力的图像编码方式。这种编码方式不仅得到了显式的目标类别、属性等信息，同时也通过这些网络得到了目标的外观特征，可以进一步在解码中帮助生成语句。

在训练方式上，一般深度学习的算法使用交叉熵损失来优化网络，每个算法预测的句子需要和人类描述的句子逐个单词地一一对应起来。但是，由于自然语言具有多义性的特点，即同一个意思可以由多种相同的方式表达，逐词比较的损失函数显然过于死板。同时，由于生成单词的过程是不可微的，无法通过反向传播算法来直接优化，因此很多工作引入了强化学习来解决这些问题^[12-14]。

3 视频识别与理解进展综述

视频是图像在时序上的序列组成。相对图像而言，视频有其特点：不仅仅只有视觉画面，视频还包含声音，具有多模态信息；另外视频中包含运动，有物体的运动信息，它对于行为类的内容有强有力的表现形式，需要理解的层次和维度更深。计算机视觉技术的迅速发展和深度学习的突破不仅拓宽了图像领域的应用，同时也给视频内容的识别与理解带来了新的契机。相对图像识别与理解，视频信息更丰富，研究也更富挑战性。目前，视觉领域研究还比较少，随着AI能力的进一步发展，以及与5G时代的深度融合，视频将成为未来社会主要的沟通交流的信息媒介。所以说视频是下一个风口，值得大力投入进行研究。

下文将由浅入深地介绍视频识别与理解的几大领域：视频目标检测与识别、视频目标跟踪、视频行为识别和视频描述，阐述各分支的研究现状。

3.1 视频目标检测与识别进展综述

目标检测是计算机视觉的一个热门研究方向，

应用广泛，如视频智能监控、机器人导航、工业视觉检测等诸多领域。视频目标检测是图像目标检测的延伸，找出图像或视频帧中所有感兴趣的物体，并定位出目标物体位置和类别。不过在图像中，目标是静态的，目标和位置均固定不变；在视频中，目标和位置均随视频动态变化，要进行目标检测难度更大。

计算机视觉对于目标运动的分析可以大致分为三个层次：运动分割与目标检测，目标跟踪，动作识别与行为描述。其中，目标检测既是计算机视觉领域要解决的基础任务之一，同时也是视频监控技术的基本任务。视频目标检测本质上解决两个问题：图像识别和目标定位。

简单来讲，相比单张图片检测，视频检测多了时间上下文的信息。充分利用好时序上下文关系，可以解决视频中连续帧之间的信息冗余问题，减少计算量，从而提高检测速度。同时关注如何提高检测质量，解决视频中的运动模糊、视频失焦等问题。

按照技术实现的原理来看，现在的视频目标检测方法主要有以下三类：一是与光流相结合，如微软亚洲研究院的Dai的Deep Feature Flow (DFE)算法^[15-16]，该模型首先将检测任务分为特征提取任务和检测任务两部分，通过算法识别出关键帧、非关键帧，在关键帧上提取特征得到特征图，在非关键帧上使用FlowNet估计光流，通过关键帧提取的特征做双线性的形式得到非关键帧的特征图，在两种方式得到特征图后使用检测网络完成目标检测。二是与RNN模型结合，如Xiao等^[17]提出了用于视频对象检测的空间-时间记忆网络，其核心是一种新颖的时空记忆模块 (STMM) 作为反复计算单元，用于模拟长期时间外观和运动动力学。STMM的设计实现了预训练骨干卷积神经网络 (CNN) 重量的完全集成，这对于精确检测至关重要。此外，针对视频中的对象运动，提出了一种新颖的MatchTrans模块，用于在帧与帧之间对齐空间-时间记忆。三是与检测相结合，如Feichtenhofer等^[18]提出了一种ConvNet架构，它能

同时执行检测和跟踪。总的来说，视频目标检测相关的研究还不多，该领域处于起步阶段。但是不管是从实用性而言，还是从学术研究的角度来说，都是很不错的研究课题。

3.2 视频目标跟踪进展综述

目标跟踪是计算机视觉中一类非常重要的问题，也是视频中特有的研究问题。简单来讲，目标跟踪是在动态连续的视频序列中，建立所要跟踪物体的位置关系，得到目标物体完整的运动轨迹，并判断其运动趋势。按照跟踪目标数量的多少，可分为单目标跟踪与多目标跟踪。前者跟踪视频画面中的单个目标，后者则同时跟踪视频画面中的多个目标，得到这些目标的运动轨迹。

单目标跟踪一般默认为目标一直在视频中，视频中跟踪的目标也唯一。这类问题研究历史已久，相对比较成熟。对于多目标跟踪，一般情况下，对象位置变化很大，跟踪目标个数不固定。相对单目标跟踪来说，多目标跟踪问题更加复杂，通常需要解决以下问题：跟踪目标的自动初始化和自动终止，即如何判断新目标的出现，旧目标的消失；跟踪目标的运动预测和相似度判别，即准确的区分每一个目标；跟踪目标之间的交互和遮挡处理；跟丢目标再次出现时，如何进行再识别等。

目标跟踪是计算机视觉研究领域的热点问题之一，过去几十年以来，涌现出了一批性能出色的技术和方法，按照技术的原理，大致可以分为三大类：一是经典跟踪方法，如Meanshift、粒子滤波 (Particle Filter) 和卡尔曼滤波等；二是基于检测 (Track By Detection) 或相关滤波 (Correlation Filter) 的方法；三是近几年来出现的深度学习相关方法。

2010年以前，目标跟踪领域大部分采用一些经典的跟踪方法，比如Meanshift、粒子滤波和卡尔曼滤波，以及基于特征点的光流算法等。Meanshift的本质是一种基于概率密度分布的跟踪方法，其原理是使目标的搜索一直沿着概率梯度

上升的方向，直到迭代收敛到局部峰值。粒子滤波方法是一种基于粒子分布统计的方法，首先对跟踪目标进行建模，并定义一种相似度度量方法，以确定粒子与目标的匹配程度。卡尔曼滤波常被用于描述目标的运动模型，它不对目标的特征建模，而是对目标的运动模型进行建模，以估计目标在下一帧的位置。另外，基于特征点的光流跟踪也属于经典跟踪方法，它的原理是在目标上提取一些特征点，然后在下一帧计算这些特征点的光流匹配点，统计得到目标的位置。

经典的跟踪方法有很多缺陷：如无法处理和适应复杂的跟踪变化，另外，算法也不够鲁棒，准确性也偏低。随着技术的发展，逐渐提出了基于检测的跟踪方法，基本原理是把跟踪的目标作为前景，利用在线学习或离线训练的检测器来区分前景目标和背景，从而得到前景目标的位置。相比经典方法，基于检测的跟踪方法更能适应跟踪过程中的复杂变化。随着技术的发展，还出现了相关滤波的跟踪算法，如Martins等^[19]在2012年提出的CSK方法，作者提出了一种基于循环矩阵的核跟踪方法，并解决了密集采样的问题，最后利用傅立叶变换快速实现了检测过程。其后又衍生出了基于HOG特征的KCF方法、考虑多尺度或颜色特征的方法以及用深度学习提取的特征结合KCF的方法等。

第三大类是深度学习相关的视频目标跟踪方法。在目标跟踪上，深度学习的应用方式是吧神经网络学习到的特征直接应用到相关滤波或Struck的跟踪框架里面，从而实现目标跟踪。本质上，深度学习方法是提取到了更好的特征表达，跟踪模型有效地利用深度学习的特征提升目标跟踪性能。根据特征的层次不同，基于深度学习的多目标跟踪可以分为表观特征的深度学习、基于相似性度量的深度学习，以及基于高阶匹配特征的深度学习。另外，基于深度学习的跟踪方法有比较好的特性，是端到端输出的跟踪框架，如GOTURN。目前，基于深度学习的跟踪框架还在不断发展中，从SiameseFC到CFNet等，这个领域

正在蓬勃发展。

总的来说，基于深度学习的单目标跟踪算法取得了长足的进步。相对而言，深度学习在多目标跟踪领域的应用还比较多的局限于匹配度量的学习。在多目标跟踪领域中，研究如何采用生成式网络模型和深度强化学习来学习跟踪场景的适应性，以提升跟踪算法的性能是未来的研究趋势。

3.3 视频行为识别

视频行为识别，也称为动作识别，在智能监控、人机交互、视频序列理解、医疗健康等众多领域扮演着很重要的角色。简单来讲，是给定一个视频，让计算机判断感兴趣的人或物体在进行什么行为，比如跑步、打太极、骑自行车等。视频行为识别是计算机视觉的一个基本问题，也是计算机视觉中视频理解的关键基础技术，只有识别视频中的单个行为，才能整合起来对视频整体有所理解。

行为识别是非常有难度的，其处理对象是视频，相对于图像分类来说多了一个需要处理的时序维度。行为识别还有一个难点是视频段长度不一，而且开放环境下视频中存在多尺度、多目标、摄像机移动等众多问题，实际问题非常复杂。

按照技术的发展历程来看，行为识别技术主要分为两类：一是传统方法，如iDT(improved Dense Trajectories)方法；二是基于深度学习的相关方法。

在传统方法中，iDT算法最为经典和有效，其框架主要包含密集采样特征点、特征轨迹跟踪和基于轨迹的特征提取三个部分。通过网格划分的方式在多尺度图像中分别密集采样特征点，特征轨迹跟踪通过计算特征点领域内的光流中值得到特征点的运动方向，运动描述完成基于轨迹的特征提取。

基于深度学习的行为识别方法是目前行为识别领域的主流技术。按照实现原理的不同，也可以分为几大类：时空结合网络、双流网络、深层

生成网络。

时空网络 (Spatialtemporal Networks) 建模本质上是三维卷积网络, 受到卷积神经网络在图像分类领域取得成功的启发, 研究人员把卷积神经网络用于视频分类, 即在 2D 卷积网络的基础之上把时间信息加入卷积运算。利用三维核从空间和时间维度提取特征, 从而捕捉编码在相邻帧中的时空信息和运动, 最经典的方法就是 Facebook 提出的 C3D 模型。

双流网络模型 (Two-Stream Networks) [20] 主要借鉴人类视觉皮层织物的原理, 视觉皮层包含两个通路: 腹流参与物体识别, 背流参与空间位置信息处理以及识别动作。Simonyan 和 Zisserman 设计了一类深层神经网络, 引入两个并行网络用于行为识别, 即空间流 (Spatial stream) 接受原始静止视频帧, 时间流网络 (Temporal stream) 接受光流场为输入, 这两个流在最后融合进行识别和分类。

深层生成网络 (Deep Generative Networks) 本质是一种无监督技术, 因为数据注释的成本很高, 如果训练不需要标签, 模型还能够很好地捕捉到序列的内容和运动, 则可以实现准确的预测。目前主要的模型有 Dynencoder [21], 它是一种深度自动编码器, 用于捕获视频动态; 还有 LSTM 自动编码器, 它能接受序列为输入并学习相应的紧凑表示, 用于发现长期线索 [22]。

行为识别现在是一个正在大力进行研究的领域, 还有很多问题远远没有到解决: 从短时特征的提取到长时时间序列的建模, 包括多尺度的检测、跟踪、姿态估计以及相关问题。另外, 行为分析识别还和视频描述有很大的相关性, 是后者的基础, 都属于视频理解。

3.4 视频描述

在对视频的物体识别分类, 目标的行为识别之后, 最重要的工作是对视频内容的理解。图像描述 (Image Captioning) 是为一张图像生成一句描

述, 视频描述 (Video Captioning) 与其类似, 是为一个视频片段生成一句描述。因为短视频包含多帧视频图像, 所以相比于图像描述, 视频描述更复杂更具难度, 需要考虑帧与帧之间的关系 (时序因素) 对生成句子的影响, 以及视频上下文之间的联系。

在视频描述中, 理解视频时不仅要理解每一帧中的物体, 也要理解物体在多帧之间的运动。所以, 视频理解往往比图像理解更为复杂。主要表现为如下几点: 一、视频既可以被视为帧流, 亦可以视为 Trimmed Video Clip; 当其被视为帧流的时候, 首先将视频帧作为图片, 最后对其进行池化操作; 对于 Trimmed Video Clip 的情况, 可以使用时域频域模型来学习视频中的一些行为。二、视频中的多模态信息, 如音频和字符, 都可以作为视频描述的输入特征。三、在视频描述中可以使用不同的池化方式, 如基于时间轴的均值池化和最大值池化, 还可以使用注意力机制, 比如 Hard-Attention 和 Soft-Attention, 或者使用编解码的方法对其进行处理。

S2VT 模型 (Sequence to Sequence-Video to Text) 是视频描述领域最早的研究之一, 它使用了编码器解码器结构 [23]。编码器的主要作用是将有用的视频信息使用深度神经网络 (CNN, RNN) 转化成特征给解码器使用。

SA 模型和 S2VT 模型一样是比较经典的视频描述任务模型, 两者的创新性各有偏重: S2VT 将 seq2seq 模型引入了视频描述任务, SA 模型则是将注意力机制引入了视频描述任务。注意力机制本质上是在视频中有权重地选择特征, 即视频提取多帧的图像特征后, 在生成描述的时候, 会给每个图像特征一个权重系数, 决定这个权重系数的是模型解码阶段每个 RNN (LSTM) 隐藏单元, 所有权重系数的和为 1。

总结来讲, 在视频描述领域, 目前主要有两大类方法: 一类是基于 CNN-LSTM 框架, 另一类是基于编解码框架。

4 计算机视觉的应用概况

4.1 在智慧医疗领域的应用

随着近几年来计算机视觉技术的进步，智慧医疗领域受到了学术界和产业界的持续关注，其应用也越来越广泛和深入。Topol等^[24]指出，面向智慧医疗，人工智能技术从三个层面将产生深刻的影响：第一层面对于临床医生，计算机视觉技术能帮助其更快速、更准确地进行图像分析工作；第二层面对于卫生系统，其能通过人工智能的方式改善工作流程、减少医疗差错；第三层面对于患者，通过增强的云存储能力，他们可以处理自己的数据，以促进自我健康。

深度学习在医学图像等领域的应用有大量的研究成果被报道。如图像配准技术，是在医学图像分析领域进行量化多参数分析与视觉评估领域的关键技术^[25]。Monti^[26]等比较并且评估了PET与MR两种图像模式下的配准方式，通过实验论证了PET与MR混合的配准方式能在头颈部这类复杂的解剖区域提供更高的配准精度。Esener等^[27]则针对乳腺癌的诊断，采用多阶段分类的方式研究了一种特征集合技术，获得了相对于单阶段分类更高的分类准确率。Gargiulo等^[28]将CT和MRI图像与DTI跟踪成像结合起来，使用图像分割协议对颅骨基底、肿瘤和关键纤维束进行三维建模，为神经外科手术提供了先进的治疗方式。

Gurovich等^[29]通过训练能代表200多种综合症的17000多张罕见病患者的面部图像，得到了DeepGestIt算法模型，该模型能够提高识别罕见遗传综合症的准确率，在实验的502张不同的图像中，其正确识别综合症的准确率达到91%。Friedman等人训练了基于卷积神经网络的人工智能来识别心室功能障碍患者，当在52870名患者上测试该神经网络，灵敏度、特异性和准确度的值分别为86.3%、85.7%和85.7%。而且在没有心室功能障碍的患者中，人工智能筛查阳性的患者未来

发生心室功能障碍的风险是阴性的4倍。

4.2 在公共安全领域的应用

公共安防领域是计算机视觉技术的重要应用场景，尤其是人脸识别技术，作为构建立体化、现代化社会治安防控体系的重要抓手和技术突破点，在当前的图侦实战中具有重要应用价值^[30]。近十年来，街道摄像头等视觉传感器的普及为智能安防的应用提供了硬件基础与数据基础，为深度学习算法模型提供了大量的训练数据，从而大幅提升了人脸识别的技术水平^[31-33]。

国内多家人脸识别产品已经被公安部门用于安防领域。完整的人脸识别系统包括人脸检测、人脸配准、人脸匹配、人脸属性分析等模块^[34]，其主要应用包括静态人脸识别、动态人脸识别、视频结构化等。如1:1比对的身份认证，相当于静态环境下的人脸验证任务，用于比对输入图像与指定图像之间的匹配，已经成熟应用于人脸解锁、身份验证等场景，在2008年北京奥运会期间作为国家级项目投入使用，在奥运会历史上第一次使用该项技术保障了开闭幕式安检的安全通畅^[35]。

动态人脸识别技术则通过摄像头等视觉传感设备在视频流中获得动态的多个人脸图像，从数据库中的大量图像中找到相似度最高的人脸图像，用于人群密集场所当中的布控，协助安全部门进行可疑人口排查、逃犯抓捕等情报研判任务。视频结构化则是面向人、车、物等对象，从视频流中抽象出对象的属性，如人员的体貌特征、车辆的外形特征等。这些技术能够预警打架斗殴、高危车辆等社会治安问题的发生，为公安工作提供了一大利器，成为打击违法犯罪活动、建设平安城市的重要技术^[36]。

4.3 在无人机与自动驾驶领域的应用

无人机与自动驾驶行业的兴起，让计算机视觉在这些领域的应用也成为了近年来的研究热点。以无人机为例，简单至航拍，复杂至救援救灾和空中加油等应用，都需要高精度的视觉信号以保

障决策与行动的可靠性^[37]。无人机的核心导航系统中，很重要的一个子系统就是视觉系统，通过单摄像头、双摄像头、三摄像头甚至全方向的摄像头布置，视觉系统能克服传统方法的限制与缺点，结合SLAM、VO等技术，应用近几年热门的深度学习算法，能够提升位姿估计、高度探测、地标跟踪、边缘检测、视觉测距、障碍检测与规避、定位与导航等任务的进度。从外界获取的信号与无人机飞控系统的视觉伺服系统形成闭环能提升飞行器的稳定性^[38]。目前，商用的无人机已被广泛地应用于活动拍摄、编队表演、交通检测乃至载人飞行等领域。

计算机视觉软硬件技术的齐头并进加速了自动驾驶汽车技术的发展，特别是在摄像头普及、激光雷达、毫米波雷达、360°大视场光学成像、多光谱成像等视觉传感器配套跟进条件下，在卷积神经网络深度学习算法等的配合下，基于计算机视觉系统的目标识别系统利用计算机视觉观测交通环境，从实时视频信号中自动识别出目标，为自动驾驶，如起步、加速、制动、车道线跟踪、换道、避撞、停车等操作提供判别依据^[39]。自动驾驶的车辆可以完成道路及道路边沿识别、车道线检测、车辆识别、车辆类型识别、非机动车识别、行人识别、交通标志识别、障碍物识别与避让等任务。目前国内外一些企业在自动驾驶领域都取得了不错的成绩。

4.4 在其他社会领域的应用

计算机视觉应用非常广泛，除了上文提到的三个重要的领域之外，在三大产业（工业、农业、服务业）都有着大量的应用实践，为人类生活提供了越来越多的便利。如在工业领域，计算机视觉是工业机器人领域的关键技术，配合机械装置能够实现产品外观检测、缺陷检测、质量检测、产品分类、部件装配等功能。ABB公司研发的IRB360工业机器人借助FlexPiker视觉系统实现了跟踪传送带物品并且完成分拣，大大提升了工

作效率。在工业互联网大力推进的大背景下，计算机视觉的应用将越来越普及，在智能化、无人化的工业上发挥出更大的作用。

在农业领域，计算机视觉的应用成果涉及农产品品质检测、作物识别与分级、农副产品出厂质量监测、植物生长监测、病虫害的探测与防治、自动化收获等领域，为精细农业和农业生产自动化奠定了基础。腾讯报道了其在2018年利用摄像头等传感器收集温室气温等环境数据，再通过深度学习模型计算、判断与决策，远程控制黄瓜的生产，减少了人力资源的投入。

在第三产业，“智慧城市”概念带动了诸如智慧交通、智慧教育、智慧社区、智慧零售、智慧政务等基于计算机视觉技术的应用场景。在泛娱乐行业，应用深度卷积神经网络的AlphaGo在2016年击败了围棋世界冠军李世石九段，成为了人工智能发展历史上的一个里程碑事件。而不久后，其研发团队推出了更为先进的AlphaGoZero系统，采用新的强化学习算法，获得了超人的性能，同时打败围棋、将棋、西洋棋的世界顶尖人类选手。这也预示以计算机视觉为代表的人工智能技术未来将深刻改变人类的生活方式乃至社会形态。

5 结束语

近年来，互联网、移动互联网带来了海量的数据积累，以图形处理单元为代表的计算力显著提升，深度学习这一关键技术蓬勃发展，计算机视觉技术在多个领域突破了实用的技术红线，并且开始在越来越多公众熟知的领域起到至关重要的作用。本文主要从图像和视频的识别与理解两个角度总结了计算机视觉领域的最新研究成果，并解释了这些技术背后的原理。最后，介绍了计算机视觉在智慧医疗、公共安全和自动驾驶等领域的应用。可以预见未来以计算机视觉为代表的人工智能技术将深刻改变人类的生活方式乃至社会形态。

参 考 文 献

- [1] LeCun Y, Bengio Y, Hinton G. Deep learning [J] . Nature, 2015, 521 (7553): 436.
- [2] Treichler D G. Are you missing the boat in training aids [J] . Film and AV Communication, 1967, 1: 14–16.
- [3] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks [J] . Advances in Neural Information Processing Systems, 2012, 25 (2) .
- [4] Corinna C, Vapnik V. Support–vector networks [J] . Machine learning, 1995, 20 (3): 273–297.
- [5] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C] . Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
- [6] Girshick R, Donahue J, Darrelland T, et al. Rich feature hierarchies for object detection and semantic segmentation [C] . 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014.
- [7] Uijlings J R R, Sande K E A, Gevers T, et al. Selective search for object recognition [J] . International Journal of Computer Vision, 2013, 104 (2): 154–171.
- [8] Ren S, He K, Girshick R, et al. Faster r–cnn: Towards real–time object detection with region proposal networks [C] . Advances in Neural Information Processing Systems, 2015: 91–99.
- [9] Farabet C, Couprie C, Najman L, et al. Scene parsing with multiscale feature learning, purity trees, and optimal covers [J] . ArXiv Preprint ArXiv: 1202.2160, 2012.
- [10] Pinheiro P, Collobert R. Recurrent convolutional neural networks for scene labeling [C] . International conference on Machine Learning, 2014.
- [11] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C] . Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [12] Rennie S J, Marcheret E, Mroueh Y, et al. Self–critical sequence training for image captioning [C] . Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [13] Liu S, Zhu Z, Ye N, et al. Improved image captioning via policy gradient optimization of spider [C] . Proceedings of the IEEE international Conference on Computer Vision, 2017.
- [14] Ren Z, Wang X, Zhang N, et al. Deep reinforcement learning–based image captioning with embedding reward [C] . Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [15] Zhu X, Xiong Y, Dai J, et al. Deep feature flow for video recognition [C] . Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [16] Zhu X, Wang Y, Dai J, et al. Flow–guided feature aggregation for video object detection [C] . Proceedings of the IEEE International Conference on Computer Vision, 2017.
- [17] Xiao F, Jae L Y. Video object detection with an aligned spatial–temporal memory [C] . Proceedings of the European Conference on Computer Vision (ECCV), 2018.
- [18] Feichtenhofer C, Pinz A, Zisserman A. Detect to track and track to detect [C] . Proceedings of the IEEE International Conference on Computer Vision, 2017.
- [19] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis [J] . IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002 (5): 603–619.
- [20] Zhu G, Wang J, Wu Y, et al. MC–HOG correlation tracking with saliency proposal [C] . Thirtieth AAAI Conference on Artificial Intelligence, 2016.
- [21] Yan X, Chang H, Shan S, et al. Modeling video dynamics with deep dynencoder [C] . European Conference on Computer Vision. Springer, Cham, 2014.
- [22] You Q, Jin H, Wang Z, et al. Image captioning with semantic attention [C] . Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
- [23] Shin H C, Orton M R, Collins D J, et al. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data [J] . IEEE transactions on pattern analysis and machine intelligence, 2013, 35 (8): 1930–1943.
- [24] Topol E J. High–performance medicine: the convergence of human and artificial intelligence [J] . Nature Medicine,

- 2019, 25 (1): 44.
- [25] Junfeng G, Yong Y, Pan L, et al. Computer Vision in Healthcare Applications [J]. Journal of Healthcare Engineering, 2018, 2018:1-4.
- [26] Serena M, Carlo C, Mario C, et al. An evaluation of the benefits of simultaneous acquisition on PET/MR coregistration in head/neck imaging [J]. Journal of Healthcare Engineering, 2017, 2017:1-7.
- [27] Idil I E, Semih E, Tolga Y. A new feature ensemble with a multistage classification scheme for breast cancer diagnosis [J]. Journal of Healthcare Engineering, 2017, 2017:1-15.
- [28] Paolo G, árnadóttir í, Gíslason M, et al. New directions in 3D medical modeling: 3D-printing anatomy and functions in neurosurgical planning [J]. Journal of Healthcare Engineering, 2017, 2017:1-8.
- [29] Gurovich Y, Hanani Y, Bar O, et al. Identifying facial phenotypes of genetic disorders using deep learning [J]. Nature Medicine, 2019, 25 (1): 60.
- [30] 胡雪地, 庄唯, 罗忆. 人脸识别技术在警务实战中的应用探索 [J]. 警察技术, 2017(4): 65-68.
- [31] Sun Y, Wang X, Tang X. Deep learning face representation from predicting 10, 000 classes [C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, 2014.
- [32] Taigman Y, Yang M, Ranzato M, et al. DeepFace: closing the gap to human-level performance in face verification [C]. Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, 2014.
- [33] Trigueros D S, Meng L, Hartnett M. Face recognition: from traditional to deep learning methods [J]. ArXiv preprint arXiv:1811.00116, 2018.
- [34] 汪海洋. 人脸识别技术的发展与展望 [J]. 中国安防, 2015 (21):62-65.
- [35] 苏光大. 人脸识别在社会公共安全领域的应用 [J]. 中国安防, 2015 (14):12-14.
- [36] 刘晗. 人脸识别技术在十大领域的创新应用 [J]. 中国安防, 2015 (21):66-69.
- [37] Erdos D, Erdos A, Watkins S E. An experimental UAV system for search and rescue challenge [J]. IEEE Aerospace and Electronic Systems Magazine, 2013, 28 (5): 32-37.
- [38] Milford M J, Schill F, Corke P, et al. Aerial SLAM with a single camera using visual expectation [C]. 2011 IEEE International Conference on Robotics and Automation, 2011.
- [39] SAE International, Surface Vehicle Recommended Practice, J3016. Taxonomy and Definitions for Terms Related to Driving Automation Systems for on-Road Motor Vehicles [S]. June 2018.

作者简介:

张丹(1964-), 男, 博士, 加拿大工程院院士, 主要研究方向为并联机器人构型创新、运动静力学建模及操作性能优化等。

单海军(1986-), 男, 博士, 主要研究方向为计算机视觉、云计算和深度学习。

王哲(1989-), 男, 博士, 主要研究方向为计算机视觉、深度学习以及自动驾驶相关的感知技术。

基于深度学习的单目视觉 深度估计研究综述

郭继峰, 白成超, 郭 爽

(哈尔滨工业大学航天学院, 哈尔滨 150001)

摘 要: 随着智能化程度的不断提升, 对深度估计的要求也越来越高, 激光雷达和立体视觉被广泛应用, 取得了不错的效果。但受限于传感单元重量、体积以及成本等因素, 逐渐涌现出一种新的研究思路, 即仅利用成本低廉的单目视觉实现对深度信息的精确测量。首先分析了现有深度信息提取方式的特点及缺陷, 给出了单目深度估计的研究意义; 其次对近年来基于深度学习进行单目深度估计的方法进行了分类及特点分析, 包括监督学习、无监督学习、半监督学习、基于条件随机场 (CRF) 的方法、联合语义分割、引入其他信息辅助深度估计的方法; 最后对此领域的未来发展趋势做出了简要分析。

关键词: 深度估计; 深度学习; 单目视觉; 监督/无监督学习; 随机条件场; 语义信息

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 2096-5915 (2019) 02-0390-10

A Review of Monocular Depth Estimation Based on Deep Learning

GUO Jifeng, BAI Chengchao, GUO Shuang

(School of Astronautics, Harbin Institute of Technology, Harbin 150001, China)

Abstract: With the continuous improvement of intelligence, the requirement of depth estimation is becoming higher and higher. Lidar and stereo vision are widely used, and good results have been achieved. However, limited by the weight, volume and cost of the sensor unit, a new research idea has emerged gradually, that is, precise measurement of depth information using only low-cost monocular vision. Firstly, the characteristics and shortcomings of the existing methods of extracting depth information are analyzed, and the research significance of monocular depth estimation is given. Secondly, the methods of monocular depth estimation based on deep learning in recent years are classified and analyzed, including supervised learning, unsupervised learning, semi-supervised learning, conditional random field (CRF) based method, joint semantics segmentation, and information-aided depth estimation method. Finally, the future development trend of this field is briefly analyzed.

Keywords: Depth Estimation; Depth Learning; Monocular Vision; Supervised/Unsupervised Learning; Stochastic Conditional Field; Semantic Information

收稿日期: 2019-01-15; 修回日期: 2019-02-21

基金项目: 国家自然科学基金 (11472090)

1 引言

传统获取高精度的目标深度信息常采用激光雷达实现, 但因其价格昂贵, 目前多用于技术研发和测试阶段, 距离大规模市场化应用还有一定的距离; 除此之外, 近些年随着人工智能技术的快速发展, 视觉逐渐成为研究的热点, 同时也暴露出一些弊端, 其中基于双目视觉的深度估计受基线长度限制, 导致设备体积与载具平台不能很好的匹配; 基于RGB-D的深度估计量程较短, 在实际应用中能力有限, 同时在室外环境中的表现也不尽理想, 受环境变化影响较大。而单目相机具有价格低廉, 获取信息内容丰富, 体积小等优点, 可以有效克服上述传感器的诸多不足。因此, 利用单目相机获取深度信息具有重要的研究意义, 且已逐渐成为当前计算机视觉领域的研究热点之一。

一方面, 近年来几何视觉的理论体系逐渐完备, 其在增强现实和视觉SLAM中的应用已日趋成熟, 寻求理论上的重大突破有较大难度。另一方面, 深度学习的发展掀起了计算机视觉领域的又一次发展狂潮, 大量文献表明, 深度学习在解决目标识别、目标跟踪和图像分割等传统计算机视觉问题中发挥了巨大的作用, 在效率和精度上均有大幅度提升。立体视觉与学习视觉形成了各自

的理论体系, 有着各自的适应场景。在这样的背景下, 人们开始将目光转向了二者的交叉与融合, 利用单目相机提取深度信息就是其中一个重要的研究方向。

通过整理分析, 本文将基于深度学习进行单目深度估计的方法分为以下六种: 监督学习、无监督学习、半监督学习、基于随机条件场(CRF)的方法、联合语义信息、引入其他信息辅助深度估计。在实际的应用当中, 上述六种方法相互交叉, 并没有严格的界限, 为了更加清晰的梳理研究框架, 本文将基于上述分类进行逐一阐述。

2 现状分析

2.1 有监督方式训练神经网络

2.1.1 真实深度图作监督+直接拟合

文献[1-2]用于神经网络训练的数据集均是原始图像和与其对应的深度图。文献[1]中David E等首次提出用CNN卷积神经网络进行单目深度估计, 基本思想是用一个多尺度神经网络对深度图进行估计。网络架构主要为两部分, 分别是全局粗略尺度网络(Global Coarse-Scale Network)和局部精细尺度网络(Local Fine-Scale Network)。前者

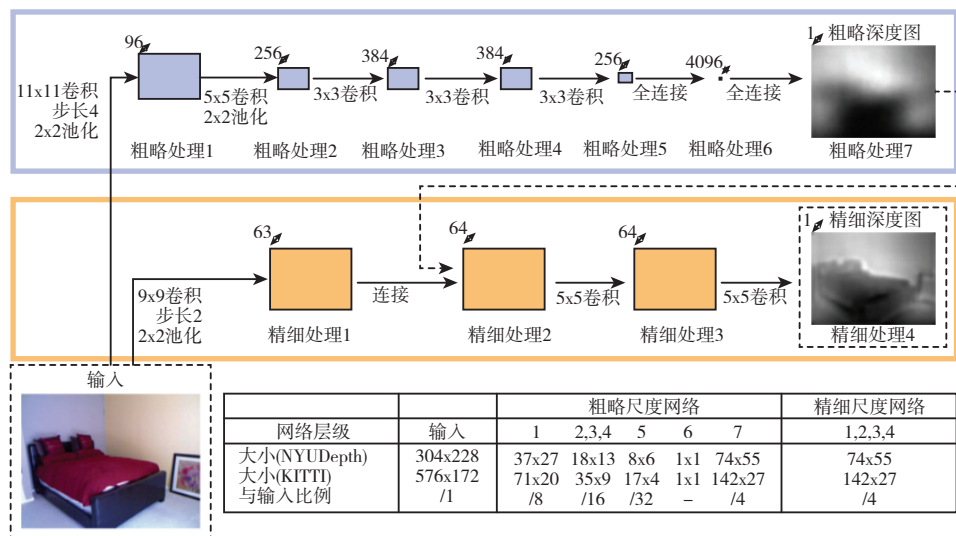


图1 多尺度网络的结构

Fig.1 Architecture of multi-scale network

得到一个低分辨率的粗略深度图,再用后者去改善前者的输出得到最后的精细深度图。值得注意的是,此文采用的是具有真实深度的数据集进行有监督训练,然而在实际场景中深度图的获取是比较困难的,因此此类数据集制作难度较大,数量也很少,算法的适应场景和泛化能力受到数据集的限制。

文献[2]在文献[1]的基础上做出了一些改进:(1)增加了第三个尺度的网络,可以输出更高分辨率的图片(分辨率从文献[1]的 55×77 提高到 109×147);(2)网络训练思路有所不同,文献[1]采用的方式是,先训练尺度1的网络,训练完成后,以网络的输出作为尺度2网络的输入。此文所采用的方式是:尺度1网络和尺度2网络联合训练,固定这两个尺度的网络之后,再训练尺度3网络,其目的仍然是对前面网络估计出的粗略深度图进行改善。

Jun L^[3]等在此基础上进行了一些改进,主要的贡献在于:(1)在文献[2]的基础上加入了多尺度网络之间的跳跃连接,可以加速网络的收敛。(2)提出了一种新的图像损失函数构造方式,使得经过数据增强后所得相关图像估计深度的差异最小化,即使由相同的数据生成的训练数据恢复得到的深度也要尽可能接近,损失函数更好地利用了增强的数据,并增强了网络的泛化能力。(3)联合估计深度和深度的梯度,并将二者融合之后输

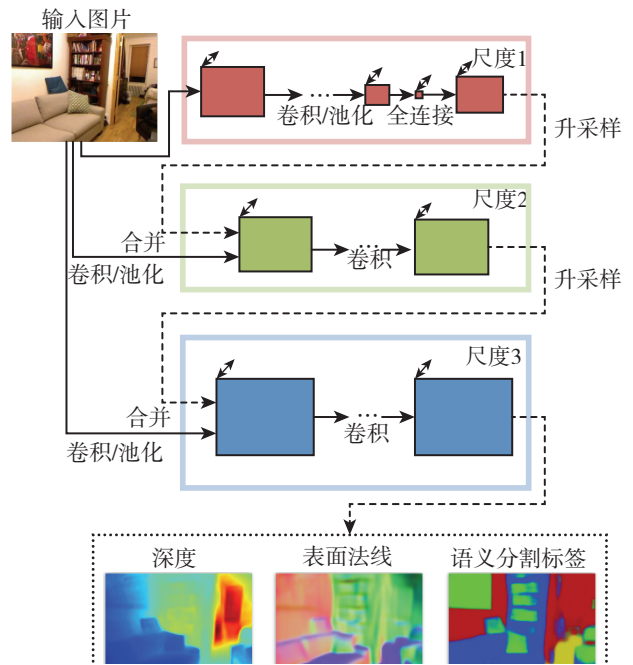


图2 改进后的多尺度网络结构:相比图1中的网络又增加了一个尺度

Fig.2 Improved multi-scale network architecture: an additional scale is added compared with the network shown in Fig.1

出最终的深度图,有效提高了估计精度。

Iro L等^[4]提出了一种包含残差学习网络的全卷积网络结构对单目图像和深度图之间的模糊映射进行建模的方法。为了提高输出分辨率,作者还提出了一种在网络中通过向上采样对卷积后得到的特征图(Feature Map)进行高效学习的新方法,并引入了反向Huber损失函数进行优化,结果表明

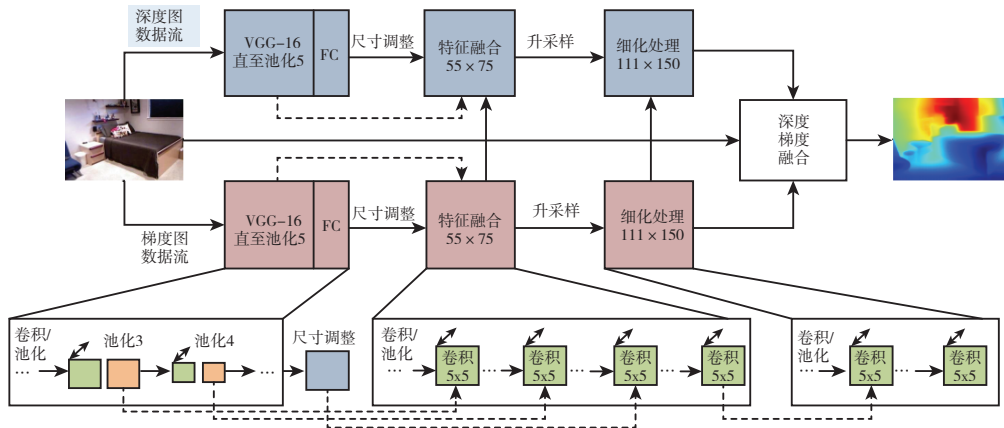


图3 双数据流深度估计网络结构

Fig.3 Two-streamed depth estimation network architecture

这个形式的损失函数特别适合于深度映射的问题。

主要贡献在于：(1)在深度预测网络中引入了一种全卷积的结构，包含升采样模块，可以输出高精度密集深度图；同时该网络结构具有比当时其他网络少一个数量级的训练参数，需要的训练数据更少。(2)新颖地使用反向Huber函数作为损失函数对网络进行优化训练，并从理论和实验上论证了为什么更适合于深度映射问题。

值得一提的是，该方法经作者改进后融合到单目SLAM技术当中^[5]，提出了一种将卷积神经网络预测的密集深度图与直接法单目SLAM的深度测量相结合的方法。这种融合方法可以在传统单目SLAM方法容易失效的位置（如缺少纹理特征的区域）进行深度预测，并利用深度预测估计地图重建的绝对尺度，克服了单目SLAM缺少尺度信息的局限性。

2.1.2 基于相对深度的标记数据集

此类方法从人类是如何对深度进行估计的角度出发，思考并得出了“比起估计准确的量化数据而言，人们更善于估计场景中各点之间、各种特征之间的相对关系”的结论，亮度和深度等特征都具有这个特点。例如，人们在观察远处的两座山时，很难判断出这两座山的实际距离是多少，却可以很轻松地说出哪一座山距离自己更近；亮

度也具有同样的特点，人们很难量化地估计出场景中的某个景物有多亮，却可以很容易判断场景中各个点的亮度关系。由此便产生了一种深度估计的思路，用经过相对深度标记的数据集训练神经网络从而使网络具备估计深度的能力。

Daniel Z^[6]等提出了一个框架，通过学习顺序关系推断图像某些特征。该系统没有直接估计特征的度量值，而是对输入图像中的各点两两之间的相对关系进行估计。之后再将这些相对关系的稀疏概率测量序列在全局层面上进行汇总，从而输出连续度量的密集深度映射图。与特征度量值估计相比，估计每一对点之间的相对顺序关系有几个优点：估计点对的相对关系比直接估计度量值相对简单且模型更加清晰明确，这种做法相当于将复杂问题简单化；人类在相对关系判断的问题上做得的更好，因此训练所需数据集的标记不依赖激光雷达等深度传感器，人工即可高效完成，所以数据收集更容易；各点对之间的相对顺序关系对数据的单调变换是不变的，从而增加了系统的鲁棒性。

Chen W F^[7]等构建了一个相对深度的数据库，每张图片仅仅标注两个随机点之间的相对远近关系，通过一个神经网络的训练就能得到输入图像

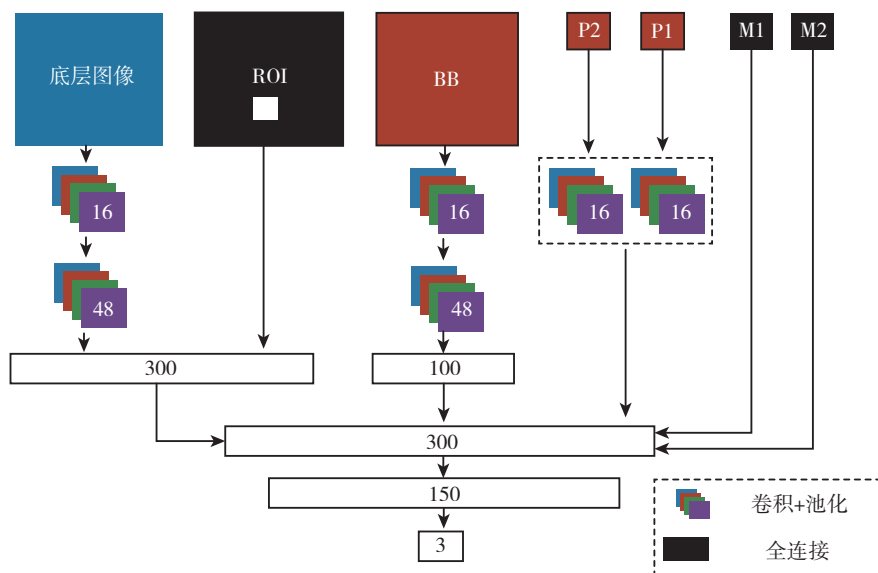


图4 神经网络结构：卷积层中的数字表示通道数

Fig.4 Deep neural network architecture: the numbers in the convolutional layers denote the number of channels

之中的相对深度信息。其将研究重点放在了室外的单目图像深度感知，即场景中存在深度值较大的景物（如天空、远山等），这是现有深度传感器量程所不能达到的。作者提出了一种新的数据集——Depth in the Wild，通过在室外场景的图像中标注随机点对之间的相对深度关系构造数据集。文章提出了一种新的算法，首先训练分类器预测图像中两点之间的顺序关系，给定一幅输入图像，反复使用该分类器预测稀疏点对之间的顺序关系（主要是相邻超像素之间的关系）；再通过求解二次约束优化从预测的顺序关系中重构深度图，该优化强制附加平滑约束，并可协调各点对之间可能存在的相互矛盾的顺序关系，以此来使得到的深度图更加平滑。

这类基于相对深度标注的训练网络的方法相当于引入了一种“QA问答”机制，训练神经网络使其具备回答“给定图像上任意两个点，谁距离相机更近？”的能力，思路使人眼前一亮。并且

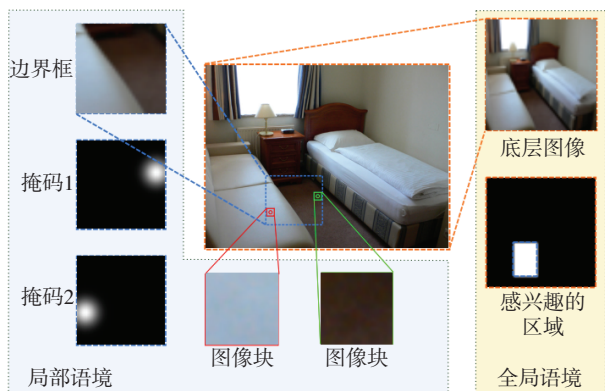


图5 网络输入：提取每个兴趣点周围的一个小块（图像上的红色和绿色方块），这些小块及其边界框（蓝色）以及表示边界框中点的相对位置的掩码一起构成网络的“本地语境”，“全局语境”由缩小后的图像以及表示图像中边界框位置的感兴趣区域（ROI）的掩码提供

Fig.5 Inputs to the network: extract a patch around each point of interest (red and green squares on the image). These patches, together with the bounding box (blue) of the patches and masks denoting the relative position of the points within the bounding box, form the “local context” for the network. The “global context” is provided by a downscaled version of the image, along with a Region of Interest (ROI) mask denoting the location of the bounding box within the image.

模型清晰明确，与人类估计深度的方式更加接近。但是由于其训练过程中始终未包含有关真实深度的信息，所以这种方法得到的深度图只是相对深度图，与真实深度图仍然相差一个尺度信息。



图6 数据集标注两个随机点之间的相对远近关系
Fig.6 The data set marks the relative far and near relationship between two random points

2.2 无监督方式训练神经网络

2.2.1 采用单目相机图片作为训练集

Zhou T H^[8]等提出了利用一段由单目相机拍摄的图片序列作为训练集，采用无监督的方式训练神经网络进行单目深度估计的方法，这也是此类方法的代表作。文中网络架构由两部分组成，即深度估计网络（Depth CNN）和相机位姿估计网络（Pose CNN），利用未做标记的单目图片序列来实现对单目深度网络和相机姿态估计网络的训练。

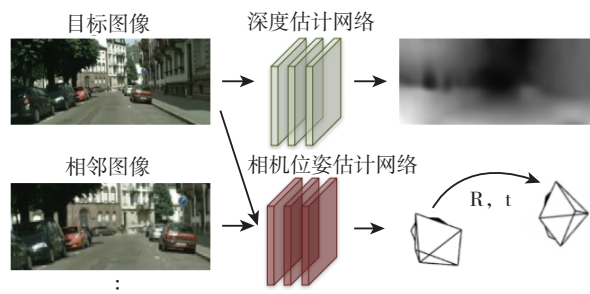


图7 通过单目视频同时估计深度及相邻帧图片间的相机位姿变化

Fig.7 Simultaneously estimate both depth and camera pose through monocular video

值得一提的是文章中提出的损失函数的构造方式：用神经网络估计出的深度和相机姿态从图

片序列的某幅图像恢复出与其相邻的另一幅图像，将恢复出的图像与真实图像之间的像素差作为无监督训练的损失函数，通过最小化损失函数使网络收敛。这种视图合成的方法也被很多无监督训练单目深度估计网络所采用。需要指出的是，该文章中采用单目图片序列作为数据集训练得到的深度估计网络输出的深度图仅为相对深度图，与真实深度图之间相差一个倍数。这是因为单目数据集中不包含深度信息，因此无法用此网络得到绝对深度图，只能计算其与真实深度图的相对误差。在自动驾驶等实际场景中，绝对深度是必不可少的导航信息，所以这种方法在机器人导航领域难以得到实际应用。

Yin Z C^[9] 在以上基础上进行了改进，增加了对光流进行估计的部分，在各个模块的预测中提取几何关系，再将它们合并从而进行图像重建，融合了深度、相机运动与光流三部分的信息进行联合估计。

Reza M^[10] 提出了一种使用三维几何约束的非监督学习单目图像深度和运动估计方法。其主要贡献是：明确地考虑场景的三维特征，维持连续帧之间三维点云的一致性，提出一个新的反向传播算法进行三维结构重建。文章使用深度和运动估计，将这种新的三维损失和基于每一帧的影像质量二维损失结合起来，同时使用了掩码去除一些无用信息存在的区域。为了证明方法的普适性，文献[10]使用了未校准的自制手持式摄像头所采集的视频进行训练，并在KITTI数据集上进行了测试。

2.2.2 采用双目相机图片作为训练集

Clement G^[11] 等利用图像重建损失对网络进行训练，之后由神经网络输出视差图，再利用预测出的视差图计算深度图。与文献[8]不同，文献[11]采用双目相机的图片序列作为数据集，因为双目相机的数据包含场景的深度信息，所以由此训练出的网络输出的视差图进而计算出的深度图为绝对深度，包含场景的尺度信息。

如图9所示，神经网络仅接收双目相机的左图作为输入，便可输出左右两幅视差图，再加上左右

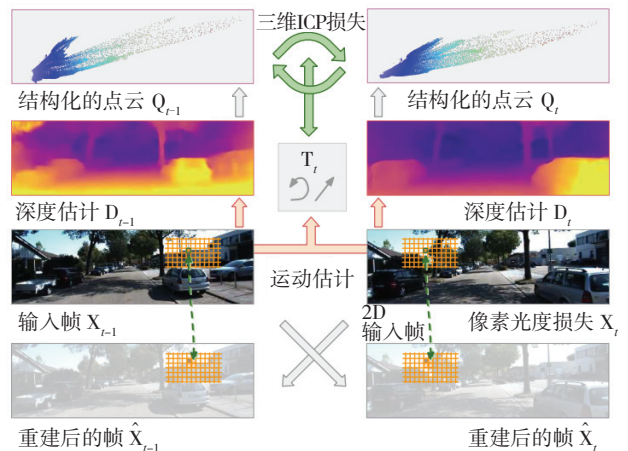


图8 除了2D的光度损失之外，3D几何损失作为监督来提升无监督深度估计的精度

Fig.8 In addition to 2D photometric losses, novel 3D geometric losses are used as supervision to adjust unsupervised depth and ego-motion estimates by the neural network.

两幅原图即可合成出左右原图的估计图，再与数据集中的原图进行比较构造损失函数使网络趋于收敛。需要注意训练集中的右图仅在训练中使用，训练之后的网络仅需左图作为输入，就完成从单张图片估计深度的任务。除此之外，文章在损失函数中引入了左右一致性检验，强制左右视差图中对应点的视差保持一致，进一步提高了对视差图的估计质量，从而更准确地计算出深度图。

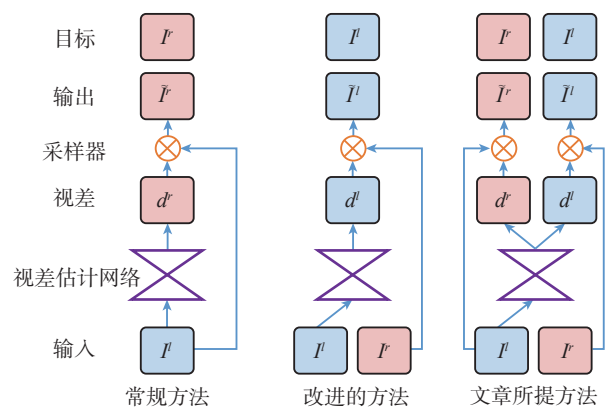


图9 利用双目图像进行无监督训练的图像重建过程示意图
Fig.9 An unsupervised training sketch of left-right image reconstruction process using stereo image pairs

Zhang Y D等^[12] 的研究重点在于解决无监督学习的尺度不确定问题，采用双目数据联合训练深度估计和视觉里程计网络。研究提出了联合学

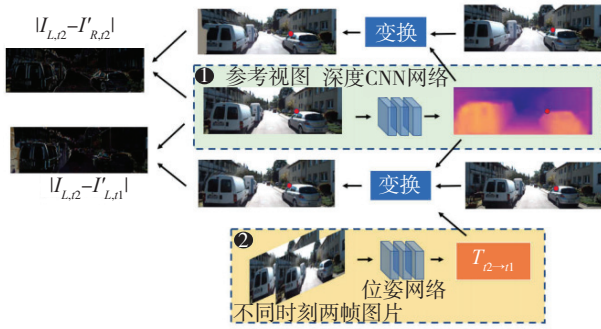


图10 训练过程的演示：立体摄像机TL→R之间已知的摄像机运动限制了深度CNN和里程计CNN以实际比例预测深度和相对摄像机姿态

Fig.10 Training instance example. The known camera motion between stereo cameras TL→R constrains the Depth CNN and Odometry CNN to predict depth and relative camera pose with actual scale

习深度估计和视觉里程计的无监督框架，不存在尺度不确定性问题；利用了左右图像对在时空上的约束，在现有技术基础上改进了单目深度估计的性能；设计了一种基于帧到帧的视觉里程计，性能显著优于同类无监督学习方法；除图像重构的颜色及光度误差损失之外，还考虑了基于深层特征的重建损失，显著提高了深度估计和视觉里程计的精度。

2.3 半监督方式训练神经网络

在单目深度估计的问题上，有监督学习的方法往往受限于真实深度难以获取，训练样本不足，而无监督学习的方法却受限于缺少尺度信息，估计精度难以保证。结合二者的优势，提出了半监

督学习的方式训练神经网络。

Yevhen K^[13]提出了使用稀疏的真实深度对神经网络进行监督，从而形成半监督学习的单目图像深度估计的方法。在损失函数的构造上，共分为三部分：有监督损失 (Supervised Loss)、无监督损失 (Unsupervised Loss) 和正则化损失 (Regularization Loss)。其中有监督损失即为上文提到的使用稀疏的真实深度与预测的深度值之间的差距形成；无监督损失的构造方式与文献 [11] 中利用左右图像的视图重建的方法一致；最后一部分是将梯度的正则化损失作为约束。

此外，文献 [14-16] 同样是用半监督学习的思想训练神经网络。

2.4 基于条件随机场 (CRF) 的方法

条件随机场 (CRF) 在图像语义分割领域一直都有着非常优秀的表现，考虑到深度值的连续性特点，可用连续CRF公式化学习，因此人们开始将CRF应用在解决深度估计问题上，近年来取得了一些成果。

Liu F Y^[17]等将卷积神经网络与连续CRF相结合，提出了一种深度卷积神经场模型。深度值具有连续性的特点，因此可以解析计算概率密度函数中的配分函数，从而求对数似然的最优解，在反向传播训练中精确地计算出梯度。除此之外，由于存在封闭形式的解，可以有效解决预测深度的映射问题。

Xu D^[18]等提出了一种融合了来自多个卷积神

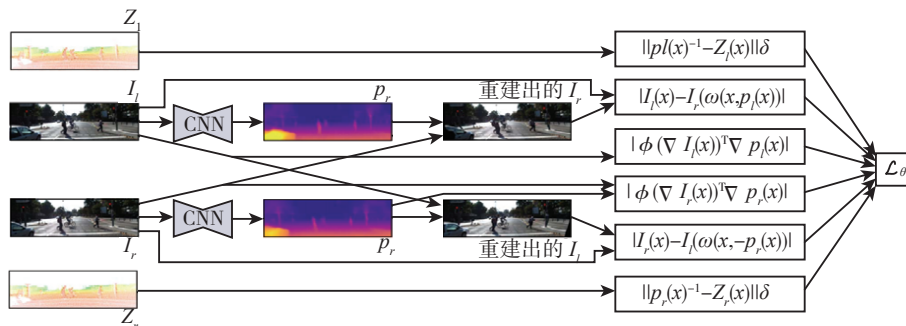


图11 此半监督方法的损失函数构造：在图像有真实深度的地方进行监督学习，无真实深度的地方进行无监督学习
Fig.11 Components and inputs of semi-supervised loss function: supervised learning is employed where groundtruth of depth is available, while unsupervised learning is used where there is no groundtruth

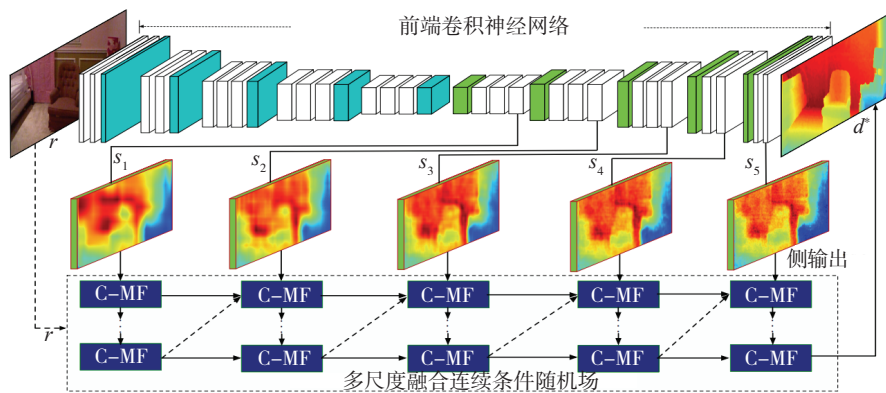


图12 该模型由前端CNN和融合模块两部分组成

Fig.12 Model is composed of two main components: a front-end CNN and a fusion module

神经网络输出的互补信息的深度模型。将连续条件场平均后积分，提出了两种变体，一种是基于多个CRF联级，另一种基于统一的图模型。提供了两种方案通过连续CRF积分多层联级输出实现融合多个尺度特征，一是多个尺度CRF模型，二是级联特定尺度的CRF。

2.5 联合语义分割进行深度估计

文献 [19–22] 将语义分割与深度估计联合起来，利用深度和语义信息之间的相似关系，使二者相辅相成，达到性能提升的目的。Arsalan M^[19] 等提出了一种新的模型，对单个RGB图像同时进行深度估计和语义分割。作者为每个任务分别训练一部分模型，然后利用单个损失函数将两个模型组合成一个完整的模型并进行微调。此外，作者还将深层卷积神经网络与全连接的CRF连接在一起，可以捕获语义和深度信息之间的相互顺序关系，提高了二者的准确性。与此类似，Zhang Z Y等^[20] 研究了单目图像的物体级分割和深度排序问题，在精确分割的基础上，通过MRF（马尔科夫随机场）对局部深度排序进行整合，形成完整的深度图。深度估计与语义分割都是图像场景理解与分析的重要内容，二者具有很强的相关性。将它们联合起来，相互辅助校正的思想，还具有较大的研究价值。

2.6 引入其他信息辅助深度估计

还有一些其他可从单张图片中提取出的，与

深度包含相似信息的特征可以帮助我们更加准确便捷地估计深度。Pratul P S等^[23] 提出用相机孔径效应作为监督信息预测单张图的深度。在网络里用一个可微分的孔径绘制函数，可以训练一个神经网络从一张全聚焦图像回归一个深度图。Qi X J等^[24] 从RGB图像中联合估计深度和表面法线，表面法线可由三维点云估计，深度由表面法线解线性方程组推算。其中，作者利用了平面区域表面法线变化不大的几何规律，来辅助和优化单目深度估计。文献 [25–26] 引入物体表面法线辅助深度估计，Zhang Y D等^[12] 通过从颜色上预测法线，然后求深度，进而在RGB-D图像的深度通道中填充较大的缺失区域，使用经过真实深度填补的深度相机拍摄的深度图和RGB图片作为训练集。引入了颜色信息，补充深度相机图片缺失深度的部

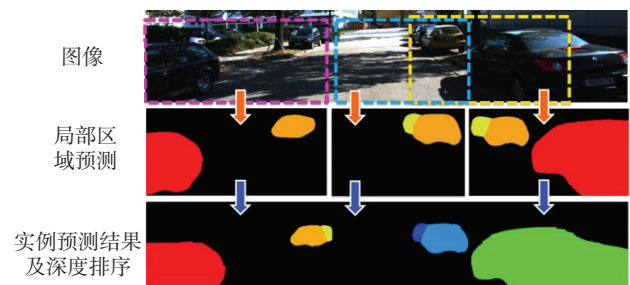


图13 通过神经网络预测每个局部区域的物体分割及深度排序，之后进行整张图片的整合

Fig.13 The neural network is used to predict the object segmentation and depth ordering of each local region, and then the whole picture is integrated.

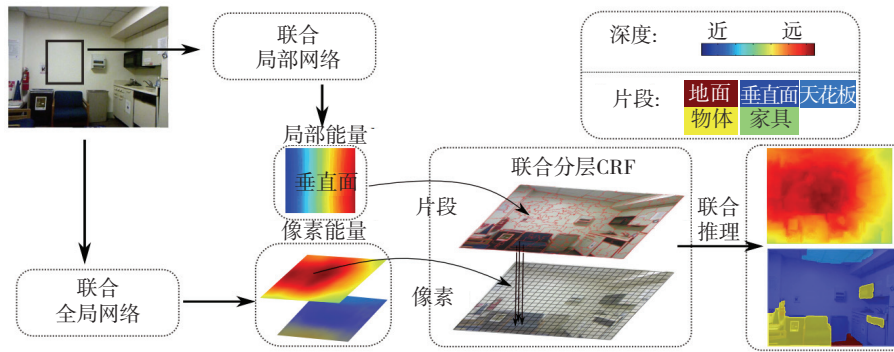


图14 语义分割和深度估计联合框架

Fig.14 Joint framework of semantic segmentation and depth estimation

分, 扩展了单目深度估计的实际应用。文献 [27-28] 把深度估计问题当成一个由远到近层层分离的问题, 按照分类问题的思想来研究。

3 结束语

本文将基于深度学习的单目深度估计方法分为了六种方法: 监督学习、无监督学习、半监督学习、基于条件随机场 (CRF) 的方法、联合语义信息、引入其他信息辅助深度估计, 并对其原理及各类方法的特点进行了分析。

从 2014 年 David E 首次提出使用深度学习的方法解决单目深度估计问题开始, 这一问题就成为计算机视觉领域的前沿课题, 近几年的研究已经取得了初步成果, 但依然存在精度不足的问题, 神经网络的稳定性和泛化能力也有待进一步提高。如何将单目深度估计、视觉 SLAM 技术、三维重建等技术相结合, 使之在实际工程问题中得到较好的应用也是今后需要重点研究的方向。

参 考 文 献

- [1] David E, Christian P, Rob F. Depth map prediction from a single image using a multi-scale deep network [C]. The 28th Conference on Neural Information Processing Systems, Montr é al, Canada, December 8-13, 2014.
- [2] David E, Rob F. Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture [C]. The 15th International Conference on Computer Vision, Santiago, Chile, December 13-16, 2015.
- [3] Jun L, Reinhard K, Angela Y. A two-streamed network for estimating fine-scaled depth maps from single RGB images [C]. The 16th International Conference on Computer Vision, Santiago, Chile, October 22-29, 2017.
- [4] Iro L, Christian R, Vasileios B. Deeper depth prediction with fully convolutional residual networks [C]. The 4th International Conference on 3D Vision, CA, USA, October 25-28, 2016.
- [5] Keisuke T, Federico T, Iro L, et al. CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction [C]. The 30th IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, July 22-25, 2017.
- [6] Daniel Z, Phillip I, Dilip K, et al. Learning ordinal relationships for mid-level vision [C]. The 15th International Conference on Computer Vision, Santiago, Chile, December 13-16, 2015.
- [7] Chen W F, Fu Z, Yang D W, et al. Single-image depth perception in the wild [C]. The 30th Conference on Neural Information Processing Systems, Barcelona, Spain, December 5-10, 2016.
- [8] Zhou T H, Matthew B, Noah S, et al. Unsupervised learning of depth and ego-motion from video [C]. The 30th IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, July 22-25, 2017.
- [9] Yin Z C, Shi J P. GeoNet: unsupervised learning of dense depth, optical flow and camera pose [C]. The 31th IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 19-21, 2018.
- [10] Reza M, Martin W, Anelia A. Unsupervised learning

- of depth and ego-motion from monocular video using 3D geometric constraints [C]. The 31th IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 19–21, 2018.
- [11] Clement G, Oisin M A, Gabriel J B. Unsupervised monocular depth estimation with left-right consistency [C]. The 29th IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 19–21, 2016.
- [12] Zhang Y D, Ravi G, Chamara S W, et al. Unsupervised learning of monocular depth estimation and visual odometry with deep feature reconstruction [C]. The 31th IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 19–21, 2018.
- [13] Yevhen K, Jorg S, Bastian L. Semi-supervised deep learning for monocular depth map prediction [C]. The 30th IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, July 22–25, 2017.
- [14] Yang N, Wang R, Jorg S, et al. Deep virtual stereo odometry: leveraging deep depth prediction for monocular direct sparse odometry [C]. The 15th European Conference on Computer Vision Munich, Germany, September 8–14, 2018.
- [15] Tulyakov S, Ivanov A, Fleuret F. Semi-supervised learning of deep metrics for stereo reconstruction [J]. arXiv Preprint, 2016.
- [16] Stepan T, Anton I. Weakly supervised learning of deep metrics for stereo reconstruction [C]. The 16th International Conference on Computer Vision, Santiago, Chile, October 22–29, 2017.
- [17] Liu F Y, Shen C H, Lin G S. Deep convolutional neural fields for depth estimation from a single image [C]. The 28th IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, June 8–10, 2015.
- [18] Xu D, Elisa R, Ouyang W L, et al. Multi-scale continuous CRFs as sequential deep networks for monocular depth estimation [C]. The 30th IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, July 22–25, 2017.
- [19] Arsalan M, Hamed P, Jana K. Joint semantic segmentation and depth estimation with deep convolutional networks [C]. The 4th International Conference on 3D Vision, CA, USA, October 25–28, 2016.
- [20] Zhang Z Y, Alexander G S, Sanja F, et al. Monocular object instance segmentation and depth ordering with CNNs [C]. The 15th International Conference on Computer Vision, Santiago, Chile, December 13–16, 2015.
- [21] Liu B Y, Stephen G, Stephen G. Single image depth estimation from predicted semantic labels [C]. The 23th IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, June 13–18, 2010.
- [22] Wang P, Shen X H, Lin Z, et al. Towards unified depth and semantic prediction from a single image [C]. The 28th IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, June 8–10, 2015.
- [23] Pratul P S, Rahul G, Neal W, et al. Aperture supervision for monocular depth estimation [C]. The 31th IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 19–21, 2018.
- [24] Qi X J, Liao R J, Liu Z Z, et al. GeoNet: Geometric neural network for joint depth and surface normal estimation [C]. The 31th IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 19–21, 2018.
- [25] Zhang Y D, Thomas F. Deep depth completion of a single RGB-D image [C]. The 31th IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 19–21, 2018.
- [26] Wang P, Shen X H, Bryan R, et al. SURGE: Surface regularized geometry estimation from a single image [C]. The 30th Conference on Neural Information Processing Systems, Barcelona, Spain, December 5–10, 2016.
- [27] Li B, Dai Y, Chen H, et al. Single image depth estimation by dilated deep residual convolutional neural network and soft-weight-sum inference [J]. arXiv Preprint, 2017.
- [28] Cao Y Z, Wu Z F, Shen C H. Estimating depth from monocular images as classification using deep fully convolutional residual networks [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28 (11): 3174 – 3182.

作者简介:

郭继峰(1977-), 男, 博士, 教授, 主要研究方向为空间机器人智能感知与自主规划技术。

白成超(1990-), 男, 博士研究生, 主要研究方向为空间机器人智能感知技术、SLAM。本文通讯作者。

郭爽(1997-), 男, 硕士研究生, 主要研究方向为视觉感知、单目深度估计。

垂面栖息微型无人机飞行原理 与技术进展

昌 敏¹, 孙 杨², 白俊强^{1, 2}

(1. 西北工业大学无人系统技术研究院, 西安 710072; 2. 西北工业大学航空学院, 西安 710072)

摘 要: 微型无人机具有体积小、重量轻、成本低、飞行灵活等优点, 但是微型无人机气动效率低、动力能源不足, 导致任务时间受限, 使用效能低下。受鸟类、蝙蝠等飞行动物启发, 仿生设计的栖息无人机是一种多模态平台, 即可空中巡航、物面停靠栖息或者爬行。停靠栖息或者爬行耗费的能量远远低于空中巡航所需的能量, 因此栖息能力将大大延长任务时间。为了实现飞行模式与栖息模式的灵活转换, 栖息微型无人机一般选择栖息于垂面, 比如建筑的内外墙壁、树干等。概括了典型垂面栖息微型无人机的方案特点, 结合微棘式、仿生足式和针刺式三类典型垂面栖息方式, 阐述了不同栖息原理的着壁、停靠栖息与复飞过程, 为后续开展垂面栖息无人机的研究提供参考。

关键词: 微型无人机; 多模态; 垂面栖息; 栖息原理; 微棘式栖息; 仿生足式栖息; 针刺式栖息

中图分类号: V279+.2 **文献标识码:** A **文章编号:** 2096-5915 (2019) 02-0400-10

Flight Principles and Research Progress on Vertical-Perching Micro Aerial Vehicle (MAV)

CHANG Min¹, SUN Yang², BAI Junqiang²

(1. Northwestern Polytechnical University, Unmanned System Research Institute, Xi'an 710072, China;

2. Northwestern Polytechnical University, School of Aeronautics, Xi'an 710072, China)

Abstract: Nowadays, there is a growing need for micro air vehicle (MAV) for both civilian and military application due to its small size, low cost and smart adaptability. However, mission endurance decreases due to reduction in aerodynamic efficiency and poor battery energy. Inspired by flying animals like birds and bats, perching MAV is a bionic multi-mode platform which can fly, perch on the surface or crawl. Perching and crawling consumes much less energy than flying so that the mission endurance extend greatly. Perching MAV usually perch on vertical surfaces such as walls of buldings and tree trunk. In this paper, the characteristics and system composition of common perching is summarized and flight principles of three kinds of perching MAVs are introduced in detail. Based on these three kinds of perching mechanism (microspine, collapsing truss grasper and dart perching), the process of maneuver of landing, perching and takeoff are investigated respectively.

Keywords: Micro Air Vehicle; Bionic Multimode; Perch on Vertical Surface; Perching Machnism; Microspine; Collapsing Truss Grasper; Dart Perching

收稿日期: 2019-01-03; 修回日期: 2019-02-08

基金项目: 国防科技创新特区资助 (18-H863-03-ZT-002-018-03)

1 引言

微型无人机 (Micro Aerial Vehicle, MAV) 具有飞行速度快、体型小、成本低、便于携带等优点, 在民用和军事方面应用广泛。在民用方面, 微型无人机可以执行航空摄像、环境数据采集等任务。在军事方面, 微型无人机广泛装备于城市反恐特种作战的班、排及单兵作战层级, 提供即时战场情报, 战术级分队的决策辅助, 密闭空间的侦察与搜救, 为班排级特种战斗开辟“鸟瞰视角”等。表1为几款典型的小型或微型无人机的性能参数, 其中Raven^[1]是美军现役手掷小型无人机型号, Black Widow^[2]为美国航境公司研制的翼展15cm的固定翼微型飞行器, DelFly II^[3]为荷兰代尔夫特大学MAV实验室的仿果蝇扑翼飞行器, Black Hornet^[4]为FLIR公司的微型直升机, Spark^[5]为中国大疆公司的可折叠微型多轴无人机货架商品。

表1 不同微小型无人机系统对比

Table 1 Comparison of different MAVs

型号名称	Raven	Black Widow	DelFly II	Black Hornet	Spark
国家	美国	美国	荷兰	挪威	中国
级别	小型	微型	微型	微型	微型
飞行器类型	固定翼	固定翼	扑翼	直升机	多轴
时间 /min	60~90	22	10	20	31
控制距离 /km	10	2	室内	1.6	18
速度 /m·s ⁻¹	10~22	12	0~7	0~6	0~20
翼展 / (直径) /m	1.4	0.15	0.28	0.168	0.143
质量 /kg	1.9	0.060	0.017	0.033	0.3

由表1可以看出: 小型固定翼飞行器利用高效气动布局、更多的能源携带量以及低翼载荷, 可实现小时级的滞空飞行。但是, 由于微型无人机质量轻, 尺寸小, 飞行时间为数分钟或者十几分钟, 任务时间很短。受鸟类长时间飞行后的栖息现象启发, 微型无人机如果具备栖息停靠能力, 比如栖息于自然界的悬崖、树干或者城市地区中的建筑物墙壁等垂面, 将大大延长其任务时间。微型飞行器从栖息态转至飞行态需要垂面具备一定的高度, 提供势能以提高脱离复飞成功率。具

备垂面栖息能力的微型无人机具有更为广泛的应用前景, 如风力发电机故障检测^[6], 电力巡线与桥梁检测^[7-8], 城市移动安防监控, 近程信号中继, 复杂环境的抢险救灾^[9]等。

栖息无人机的微型无人机与传统爬行机器人的有机结合, 随着传感器、伺服系统、控制系统的快速发展, 像鸟类一样既能飞行又能栖息的多模态循环切换成为可能。美国麻省理工学院的Cory与Moore等^[7-8, 10-14]对输电线缆(powerline)栖息开展系统研究。美国斯坦福大学仿生与灵巧操作实验室(BDML)的Morgan^[15]等开展了垂面栖息与垂面爬行的旋翼无人机SCAMP(Stanford Climbing and Aerial Maneuvering Platform)研究。BDML的Alexis^[16-17]等开展了垂面栖息与垂面爬行的固定翼无人机研究。加拿大谢布克大学的Dino^[18]提出了基于微棘机构的动力辅助垂面栖息微型无人机解决方案。瑞士洛桑联邦理工学院的Mirko等^[19]采用针刺(dart)插入垂面实现无人机垂面栖息。

2 典型工作流程与系统组成

2.1 典型流程

无人机垂面栖息的典型流程分为4个阶段, 如图1所示。

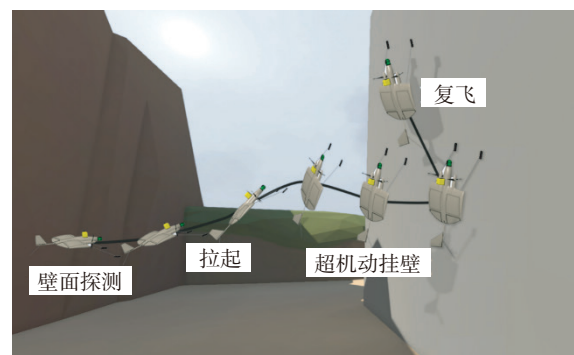


图1 栖息过程示意图

Fig. 1 Artistic representation of perching

(1) 壁面探测

栖息无人机在飞向垂面过程中, 测距传感器测出无人机与垂面之间的距离。

(2) 拉起

当壁面探测距离达到某一设定值，栖息机动指令开始执行。无人机升降舵上偏，拉起机动使无人机俯仰角不断增大，然后借助阻力增大来实现减速，在这个过程中需要严格控制俯仰角的范围，使得俯仰角速度趋近于零。无人机以较低的飞行速度、接近90°的俯仰角以及较低的俯仰角速度接近垂面，实现成功栖息。对于前期飞行速度较快，无人机俯仰角速度过大的情况，在后一阶段，可能需要操作舵面下偏以实现抑制上仰趋势。

(3) 超机动挂壁

从无人机贴上墙壁开始便展开一系列机构联动，实现栖息组件与垂面的紧连接。不同栖息组件的挂壁过程的机构动力学原理不同，差异较大。

(4) 复飞

栖息结束时，可再次进入飞行模态。栖息组件释放收回，无人机即可脱离垂面。利用重力势能，无人机下降加速，当飞机在达到合适的速度之后，舵面能够提供足够的操纵力矩，控制无人机恢复正常飞行。栖息无人机的工作全流程如图2所示。

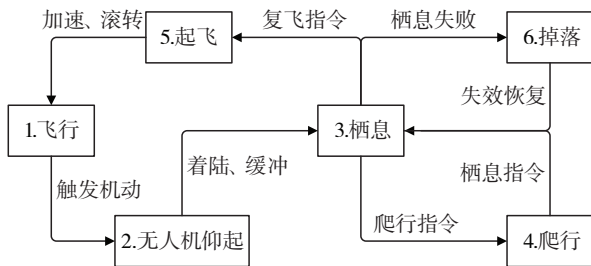


图2 典型无人机栖息流程图
Fig. 2 Flow chart of perching

2.2 系统组成

垂面栖息无人机的系统组成主要包括超机动飞行平台以及栖息组件两大部分，具体构成如图3所示。

垂面栖息微型无人机的方案众多，例如，粘性垫、吸盘抽吸、微棘、针刺、仿生足、磁力、电荷吸引等，接下来将着重介绍微棘式、仿生足式与针刺式三种栖息组件的栖息无人机飞行原理。

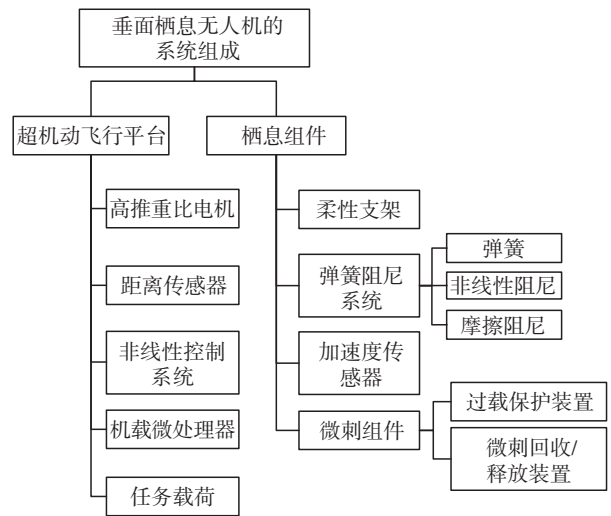


图3 栖息无人机系统组成
Fig. 3 System components of perching UAV

3 微棘式栖息

3.1 栖息装置

微棘式栖息装置布置在柔性支架的末端，并作用在垂面的粗糙颗粒上，粗糙颗粒对微棘的作用力用于平衡冲击过程中动态载荷以及最终稳定后的重力和力矩。微棘具体结构如图4所示。

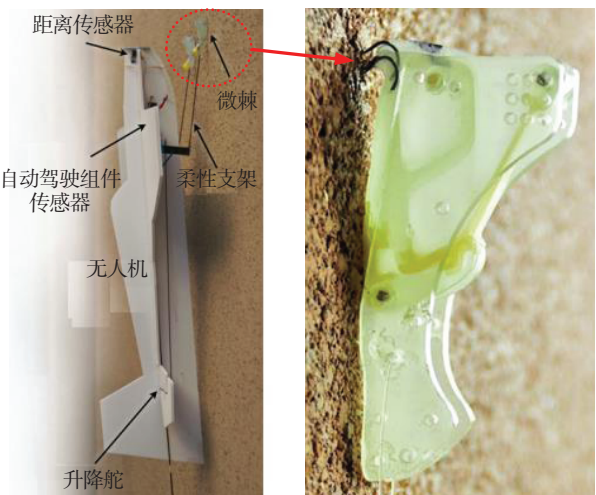


图4 微棘简化模型^[9]
Fig. 4 Illustration of microspine^[9]

微棘的简化模型如图5所示，图中3、4为切向弹簧阻尼系统，用以减缓切向速度所带来的冲击载荷，5为法向弹簧阻尼系统，用以减缓法向速

度所带来的冲击载荷。

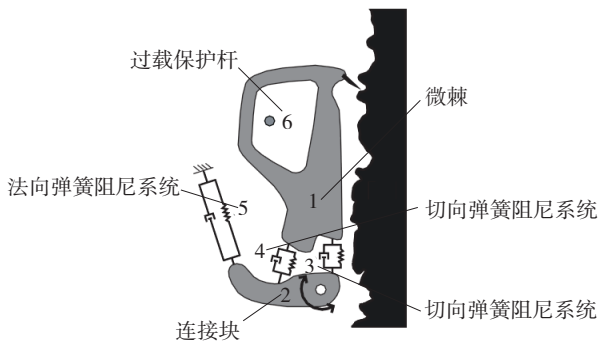


图5 微棘式栖息无人机模型^[9]

Fig.5 Artistic representation of microspine^[9]

微棘在贴壁过程中，受到较大的作用力或当法向力与切向力的比值超过一定值时，都会导致栖息失败。通常对微棘式栖息装置的研究都将其简化为弹簧阻尼系统，实验与计算结果^[9]表明，同一刚度、阻尼特性条件下，微棘受到的作用力与自身位移正相关，因此位移越大，越容易导致失败，所以必须限制微棘的最大许用行程。当最大许用行程减小时，系统阻尼比也需要增大，因为位移缩短，其时间也会缩短，耗散同样的能量，位移变小，就需要提高阻尼来增加做功的行程。但是阻尼增大，会导致初始作用力增大，对结构强度提出更高的要求，使得结构更重。如果采用库伦摩擦阻尼来取代粘性阻尼，摩擦阻尼平衡位置不确定，并且很难设定其摩擦力水平。将两种阻尼机制结合到一起，可以获得一个较好的结果。

3.2 栖息原理

栖息能否成功，直接取决于微棘的受力情况，因此首先对微棘的动力学特性进行分析。微棘工作过程中受到的切向力与法向力构成的区域即为限制平面，如图6所示。法向力为负值，表明受到的作用力使无人机远离垂面。切向力为正值，无人机受到沿重力方向的作用力。

微棘的作用力包括法向力 f_n 和切向力 f_s ， $f_n > 0$ 表示墙面与微棘是相互挤压的， $f_s > 0$ 表明微棘有下滑的趋势。静止时，正压力为负值，因为重心在墙面之外；无人机有远离的趋势，切向力为正值。

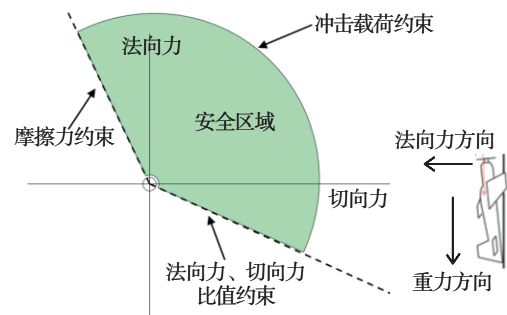


图6 微棘栖息的限制平面^[9]

Fig.6 Limit surface of microspine^[9]

微棘式栖息组件在挂壁栖息过程中的边界约束如表2所示。

表2 微棘作用力约束

Table 2 Constraints on spine forces

类型	约束	适用情况
摩擦力	$-f_s \leq \mu f_n$	$f_s < 0, f_n > 0$
附着力	$-f_n \leq a f_s$	$f_s > 0, f_n < 0$
过载	$\sqrt{(f_n/c)^2 + f_s^2} \leq f_{\max}$	始终适用

对限制平面以及约束表进行分析，可以得出脱离表面主要是三种原因：

(1) 从颗粒表面滑落，因为受力方向不当，导致法向力与切向力比值超过限制曲线的边界。

(2) 冲击力过大，使得原本适合栖息的颗粒被破坏。

(3) 微棘的作用力超过了最大值，弹性变形触发了过载保护，导致微棘收回。

限制平面由两条边界曲线和冲击力约束的边界所围成，所包围区域视为安全区域。但是安全区域内的某些情况下，依然会引起偶然性的栖息失败。之所以会出现偶然的失败，是由微棘与墙面接触的概率特性所决定的。虽然理论上能保证栖息成功，但在实际中会受很多因素影响，例如，粗糙颗粒在表面的分布是随机的，其所能承受的载荷角也是随机的，取决于其形状和当地摩擦条件。因此，可以采用冗余设计理念，既可以承受较大的动态过载，又可以保证成功率。工作过程中，多组微棘会同时在表面滑动，寻找合适接触点，因此，所有微棘全部没有接触到合适

的粗糙颗粒并获得反作用力的情况是低概率事件。即使初始接触失败，只要保证无人机与墙面仍然紧贴，无人机下滑之后会再次接触，实现栖息。

不同垂面表面特征的限制曲面有所不同，但是其规律大致相同，不同之处在于哪种失效机制占主导，比如普通沥青墙面，对于法向力和切向力比值比较敏感。而混凝土表面，受限于每个微棘允许的最大作用力和挠曲变形，主要取决于过载保护特性。

微棘式栖息能否成功，直接取决于微棘在冲击过程中所受到的动态载荷能否满足前述约束平面。而相同情况（飞行速度和角度以及系统质量、惯量等）下，微棘的动态载荷则取决于悬架系统的弹性和阻尼特性。因此，基于前述对微棘的定性分析，还要分析悬架系统的动力学特性。加装微棘的栖息组件如图7所示，其简化模型如图8所示。将两侧的栖息组件合并为一个机构，机身简化为刚体，并忽略其上的气动力作用，因为此时的速度已经很低。相比机身的质量，悬架系统本身质量很小，因此建模时忽略其影响。

因为涉及多个运动机构，所以引入4个右手法则的参照系，墙面参照系为垂直和平行于表面的单位矢量 W_x 、 W_y ，沿机身轴线和垂直机身轴线的机身参照系 a_x 、 a_y ，平行和垂直于根部短杆的矢量 f_x 、 f_y 以及平行和垂直于贴面的矢量 t_x 、 t_y 。因为需要考虑柔性支撑杆的弯曲刚度，所以简化是为方便起见，并没有将其用梁来代替，而是用绕有转动刚度和阻尼的节点的杆来代替，如此简化的柔性支撑杆分为两部分，一部分与根部短杆连接视为刚体，另一部分视为可以转动的杆，转动点赋予转动刚度和阻尼。

悬架系统包括3个衔接部位：根部连接节（hip）、转折点（knee）和末端结点（ankle）。两个支撑杆：根部短杆（femur）和柔性支撑杆（tibia）。根部关节和末端结点处都布置有泡沫，提供刚度和阻尼，通过在根部放置泡沫，限制根部仅可有小范围运动，来防止机身相对垂面有较大的位移，从而保护电机免受可能出现的撞击。连接转折点

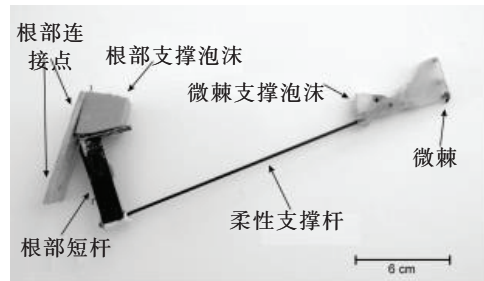


图7 微棘式栖息组件^[9]

Fig.7 Compliant leg suspension of microspine^[9]

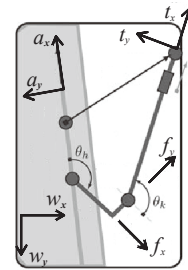


图8 微棘式栖息组件的动力学建模^[9]

Fig.8 Dynamics model of suspension^[9]

与末端结点的柔性支撑杆由于其长度较长，自由端受力时有较大弯曲，因此必须考虑其弯曲刚度，而连接根部结点与转折点的根部短杆则可认为是刚体。

在超机动挂壁过程中，如果将无人机剩余能量直接传递到微棘上，将会很容易导致栖息失败，一是直接作用力更大，触发过载保护，二是切向速度引起较大的切向力，超出限制平面的边界约束。因此，栖息组件的作用就是吸收冲击时的能量，利用自身的弹性与阻尼特性，简化机身对微棘的冲击，同时减小向机身传递的反作用力，使微棘所受的作用力满足限制曲面，确保栖息成功。通过设计栖息组件的几何结构与参数，使其能够满足较大范围的着陆速度与俯仰角。其适应性越强，对机载传感器和飞行控制的要求就越低，着陆过程就越简单和鲁棒。

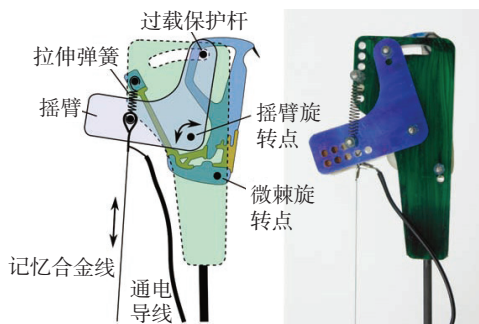
3.3 复飞

复飞过程如图9所示，其过程包括起飞时刻的状态信息获取、无人机与垂面分离以及复飞后的姿态控制。

图9 复飞实验^[17]Fig.9 Multiple exposure photograph of the takeoff sequence^[17]

(1) 微棘收放

微棘机构如图10所示, 通电后, 记忆合金线收缩, 摇臂绕其旋转点旋转, 牵动微棘旋转, 使得微棘脱离墙面。通电大约0.15s内, 记忆合金线可以收缩0.6cm。

图10 微棘收放机构^[17]Fig.10 Spine detachment^[17]

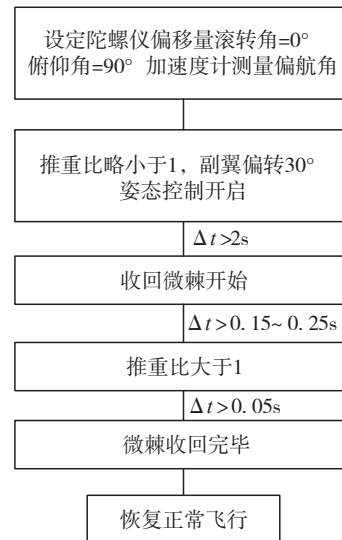
(2) 姿态测量

姿态信息的获取至关重要, 它是接下来的控制的输入。但是, 无人机在起飞时刻前后的姿态会受到多种因素干扰, 如阵风干扰、时间延迟等。起飞前, 借助加速度计获得初始俯仰角, 对陀螺仪角速度信息进行积分, 获得实时的姿态信息。

(3) 复飞控制

测量初始信息, 设定陀螺仪偏移量, 利用副翼偏转应对力矩问题, 推重比保持略小于1, 减小无人机与垂面的作用力, 易于微棘收回。准备完成之后, 0.15s内记忆合金收缩, 电机增大推重比, 辅助分离, 流程如图11所示。

当无人机获得足够大的速度, 舵面获得足够的操纵力矩时, 就可以偏转舵面, 实现180°滚转, 回到正常飞行状态。高推重比的飞机虽然效率上

图11 起飞流程图^[17]Fig.11 Flow chart of takeoff^[17]

低一些, 但是可以在紧凑的空间里实现起飞控制, 也有助于着陆失败恢复, 一旦着陆失败, 可以立即转入起飞模式, 飞离墙面。通常来说, 飞行器着陆时会立即抓住表面, 一旦失败, 其速度和加速度的变化会产生变化, 因此不难探测, 所以可以在较小的下降速度时, 立即收缩微棘, 并增大推重比, 进入飞行状态。

4 仿生足式栖息

微棘悬挂适用于具有一定粗糙度的表面, 但针对一些光滑曲面, 如玻璃等, 仿生足更具有优势。仿生足的栖息装置工作过程中仅仅依靠范德华力, 接触与分离时耗能少, 适用于多种墙面, 能够反复使用。

4.1 栖息装置

仿生足的栖息装置可以用于无人机的栖息, 如图12、13所示。也可用于捕捉太空碎片, 可以利用自身的附着能力来捕捉太阳能帆板、燃料箱等, 如图14、15所示。无人机冲击垂面的A过程中, 装置内部间距下降, 将冲击能量转化为弹簧的势能, 并将系统锁定在B状态, 从而提供持久切向力, 以维持平衡重力的法向作用力分量。



图12 旋翼仿生足式栖息装置

Fig.12 Quadrotor Micro Air Vehicle hanging from a glass surface using the directional adhesive Collapsing Truss Grasper

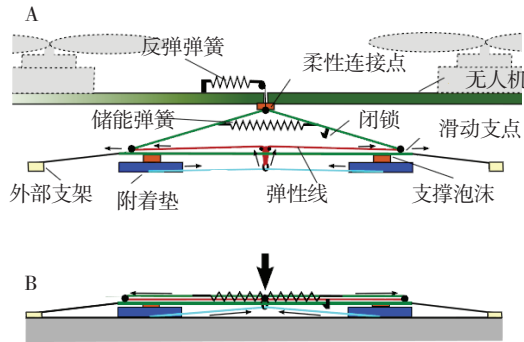


图13 杆束收缩型仿生足简化模型^[22]

A) Schematic showing functional components.
B) Device in locked state (grasping a surface)

Fig.13 Collapsing truss grasper.

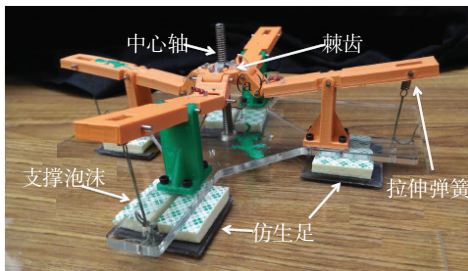


图14 中心连接型仿生足式栖息装置^[22]

Fig.14 MAV pivot linkage grasper^[22]

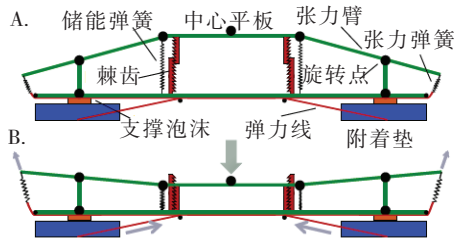


图15 中心连接型仿生足简化模型^[22]

Fig.15 Simplified model of gecko-inspired biomimetic adhesive^[22]

4.2 栖息原理

仿生足式无人机栖息大致过程如图16所示。为保证较小的冲击力，尽量使机体与物面平齐。初始附着可能经过几次反弹来耗散能量。稳定附着后施加切向力的同时产生垂直物面的相互作用力，无人机紧贴物面，将能量储存在弹簧中。如果附着时仍有一定的角速度，还会经历抵抗扭转的过程。

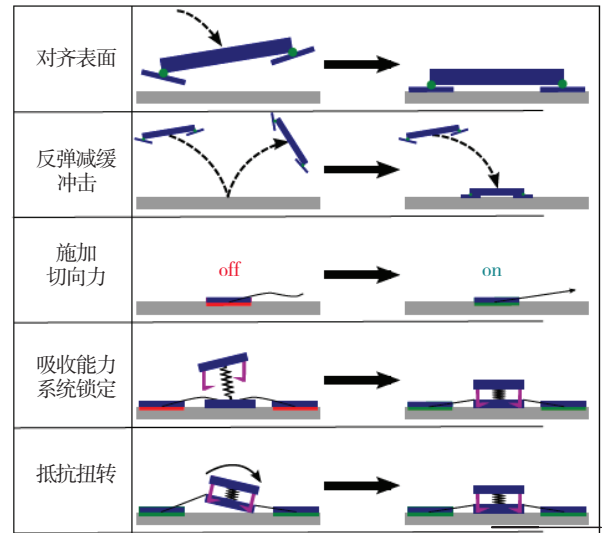


图16 仿壁虎足爪栖息过程^[22]

Fig.16 Illustration of functional principles^[22]

(1) 对齐垂面

无人机在飞向垂面时，可能与垂面存在一定的夹角。因此，无人机首先借助自身两侧外伸支架在与垂面作用时产生的力矩实现自身旋转并对齐垂面。

(2) 反弹缓冲

冲击时的能量将会在初次接触或者后续的再次弹起中被耗散，吸收能量的能力取决于设备的大小以及吸收能量时产生的作用力，后者是其自身的限制。此外，碰撞过程中，不能损坏设备，反弹过程中，不能超过附着装置的行程。

(3) 施加切向力

简单的正向作用力并不能使其附着，除非有切向力作用在正确方向。因此为了能够实现没有外部切向力输入情况下的附着，必须依靠系统本

身产生相应的作用力。

通过一定的机械结构，可以将冲击时产生的作用力转化为切向力，但是大小要适中，否则也会导致失败。多余能量可以被耗散掉或者存储起来，换句话说，力的作用可以来自一种主动机构，在加载之前，所有的触脚都必须与表面接触，因此，主动机构必须有精确传感来保证时间节点准确。

(4) 系统锁定

一旦内部机构产生了切向力，并且冲击过程为系统注入了足够的能量，系统就必须进入锁定状态，以保持切向力，这个过程通过提高棘轮和闭锁来实现。

(5) 抵抗扭转

抓附过程中，由于不能准确的垂直飞向表面，并且自身可能有一定的旋转，或者两个触脚未能同时着陆，都会引起着陆后主体部分的旋转。系统本身应当还能承受一定的扭转，并且能够以最佳方式分散作用力，防止任意触脚在这个过程中达到其自身受力上限引起脱落。如此，才能让各个子系统的组合达到互相提升的目的。

4.3 复飞

当栖息结束时，并不需要施加额外的力使无人机与垂面分离。只需要打开锁定机构，切向力消失，法向的与垂面之间的相互作用力也随之消失。原本储存的能量将会推动无人机与表面分离。



图17 复飞过程

Fig.17 Illustration of resuming flight

5 针刺式栖息

5.1 栖息装置

洛桑联邦理工学院 (EPFL) 提出如图 18 所示三种结构来实现针刺栖息^[19]。

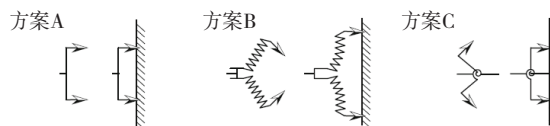


图18 针刺方案^[19]

Fig.18 Illustration of dart perching^[19]

方案A利用对称排列在前段的两组针来栖息，这样的结果没有可动的缓冲部分，简单轻便，缺点就是适应性低，仅在针刺恰好对准表面时，才能很好的刺入，发挥作用。同样，方案A没有缓冲部分，所以整体冲击作用力比较大。假设 MAV 自身质量为 20g，以 4m/s 的速度飞行，针嵌入表面深度为 1mm，那么平均作用力大约为 160N，对于无人机结构强度而言是无法承受的。方案B的装置当其接触表面时，针刺机构可以向表面伸出，针刺沿表面滑动，随后刺入表面，这种方法可以很好的适应表面形状。方案C集成了一个扭转弹簧和两个臂，无人机接触墙面，弹簧就会释放能量，两个携带针刺手臂快速向前，将针插入表面。三种方案的对比如表 3 所示。

表3 三种方案定量对比表

Table 3 Quantitative comparison of three principles

性能指标	方案 A	方案 B	方案 C
重量	5	3	4
尺寸	4	2	0
效果	2	2	5
复杂度	4	2	3
结构应力	1	4	5

分值：1-5，表示指标由劣到优

方案C相比前两个是更为有效的，其针刺的作用力可以通过改变弹力和手臂的重量来调整。减速距离可以通过改变手臂长度来调整，以此来减小冲击过程对结构的反作用力。其缺点在于，需要引入一个扭转弹簧，使得结构上相对复杂一些。方案C针刺装置的细节如图 19 所示。

5.2 栖息原理

针刺栖息过程如图 20 所示：图 (a) 触发机构接触表面，其向后移动；图 (b) 原本紧贴的两块

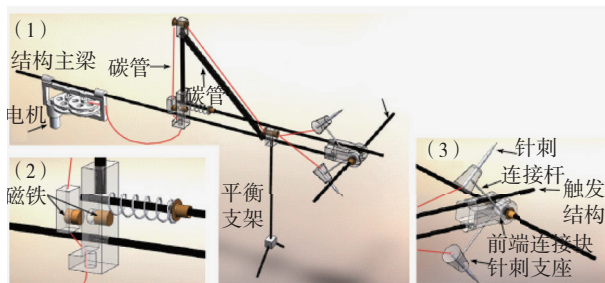


图 19 针刺机械结构示意图^[19]
Fig.19 Device of dart perching^[19]

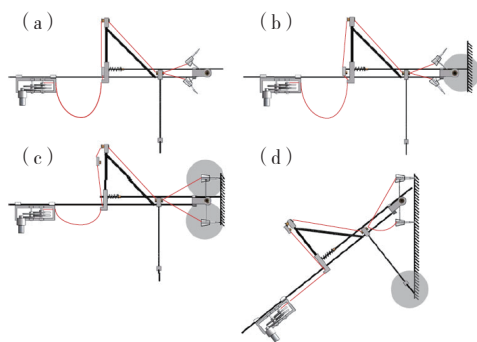


图 20 针刺栖息过程^[19]
Fig.20 Illustration of dart perching^[19]

磁铁分离，牵引线释放，弹簧驱动手臂向前移动；图(c)针刺进入表面，借助于垂面间的相互作用力，将系统能量转化为克服刺入垂面的反作用力的功；图(d)无人机在平衡支架的辅助下，进入平衡状态。

两块用于控制位置的磁铁可以产生推力，但其不是减速的主要原因。为了研究其影响，我们将计算 MAV 所受到的作用力，无人机动能为 10mJ，假设针刺减速距离为 2mm，那么整个过程的作用力约为 5N，其大小是复位弹簧作用力的 28 倍左右，对于质量较大的无人机而言，这个比例将会更大，因此我们可以认为，针刺的阻力是无人机减速的主要因素。

5.3 复飞

栖息结束时，微型电机驱动牵引线收回，将针刺与表面分离。但其自身产生的作用力较小，所以电机启动后还需借助变速箱提高拉力，才可将微棘从垂面中收回。收回线的过程后期，与牵

引线连接的磁铁不断下移，在另一块磁铁的吸附作用下进入凹槽并紧贴，推到触发机构复位，电机产生的拉力将微棘拖出，无人机脱离墙面。

6 结论

本文介绍垂面栖息无人机的优点及其应用，梳理了垂面栖息的研究进展，介绍了垂面栖息的典型流程：探测、拉起、超机动挂壁、复飞，列举了 3 种主要的栖息方式：微棘式、仿生足式、针刺式。结合栖息装置来阐述相应的栖息原理以及如何实现复飞，得出结论如下：

(1) 在某些情况下，地面可能布满杂物，不适合降落，也不适合起飞。超机动飞行以及微棘等栖息装置使得垂面栖息成为可能，避免了水平地面降落与起飞。而且栖息过程以及表面爬行机动相比持续飞行，消耗能量少，从而大大延长任务时间。

(2) 典型的栖息超机动过程：向上偏转升降舵，舵面有效迎角为负值，产生抬头力矩。无人机俯仰角不断增大，角速度呈上升趋势。超过某个临界俯仰角，升降舵有效迎角为正值，产生低头力矩。无人机俯仰角不断增大，但角速度呈减小趋势。无人机利用过程中大攻角下的压差阻力和粘性阻力来实现自身的快速减速。

(3) 微棘式栖息适用于具有一定粗糙度的垂面，该方式鲁棒性强，通过详细的设计可以承受较大范围的着陆情况。仿生足式栖息工作过程仅仅依靠范德华力，相比微棘对玻璃等光滑表面有着更好的附着能力。针刺栖息则避免了栖息超机动过程，减少了对飞行控制的需求，但是无法刺入钢铁等硬度较大的材质。

(4) 研究表明，很多栖息的失败来源于距离探测错误，导致无人机过早或过晚进入栖息模式。使用更加先进的传感器代替目前使用的超声波传感器，将有利于栖息成功。

(5) 目前使用的栖息装置大多数为被动的，未来可以使用主动的栖息装置，使无人机在垂面可以进行爬行，以便调整自己的位置和方向。

参 考 文 献

- [1] Aerovironment. Raven RQ-11A/B Datasheet. [EB/OL]. 2017 [2019]. http://www.avinc.com/images/uploads/product_docs/Raven_Datasheet_2017_Web_v1.pdf.
- [2] Grasmeyer J, Keennon M. Development of the black widow micro air vehicle [J]. Aiaa Aerospace Sciences Meeting & Exhibit, 2001.
- [3] Groen M A. PIV and force measurements on the flapping-wing MAV Delfly II- An aerodynamic and aeroelastic investigation into vortex development [D]. Delft University of Technology, 2010.
- [4] FLIR. Black Hornet PRS Datasheet [EB/OL]. 2018 [2019]. <https://www.flir.com/globalassets/imported-assets/document/black-hornet-prs-spec-sheet.pdf>.
- [5] DJI. Spark User Manual V1.6 [EB/OL]. 2017 [2019]. <https://www.dji.com/spark/info#downloads.pdf>.
- [6] Zhang H, Jackman J. A feasibility study of wind turbine blade surface crack detection using an optical inspection method [C]. Madrid, Spain, October 2013.
- [7] Joseph L. Moore J L. Powerline perching with a fixed-wing UAV [D]. Cambridge: MIT Computer Science and Artificial Intelligence Lab, June, 2011.
- [8] Moore J, Tedrake R. Magnetic localization for perching UAVs on powerlines [C]. 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2011.
- [9] Desbiens A L, Asbeck A T, Cutkosky M R. Scansorial landing and perching [J]. Springer Tracts in Advanced Robotics, 2011, 70:169-184.
- [10] Cory R E, Tedrake R. Experiments in fixed-wing UAV perching [D]. Cambridge: MIT Computer Science and Artificial Intelligence Lab, 2008.
- [11] Cory R E. Perching with fixed wings [D]. Cambridge: Massachusetts Institute of Technology, February 2008.
- [12] Cory R E. Supermaneuverable Perching [D]. Cambridge: Massachusetts Institute of Technology, June 2010.
- [13] Moore J, Cory R, Tedrake R. Robust post-stall perching with a simple fixed-wing glider using LQR-Trees [J]. Bioinspiration & Biomimetics, 2014, 9(2).
- [14] Roberts J W, Cory R E, Tedrake R. On the controllability of fixed-wing perching [C]. Hyatt Regency Riverfront, St. Louis, MO, USA. 2009.
- [15] Pope M T, Kimes C W, Jiang H, et al. A multimodal robot for perching and climbing on vertical outdoor surfaces [J]. IEEE Transactions on Robotics, 2017, 33(1): 38-40.
- [16] Roderick W R T, Cutkosky M R, Lentink D. Touchdown to take-off: at the interface of flight and surface locomotion [J]. Interface Focus, 2017: 1-15.
- [17] Desbiens A L, Asbeck A T, Cutkosky M R. Landing, perching and taking off from vertical surfaces [J]. International Journal of Robotics Research, 2011, 30(3): 355-370.
- [18] Mehanovic D, Bass J, Courteau T, et al. Autonomous Thrust-Assisted Perching of a Fixed-Wing UAV on Vertical Surfaces [J]. Living Machines, 2017: 302-314.
- [19] Kovač M, Germann J, Hürzeler C, et al. A perching mechanism for micro aerial vehicles [J]. Journal of Micro-Nano Mechatronics, 2009, 5(3-4):77-91.
- [20] James L, Tangler J, Kocurek D. Wind turbine post-stall airfoil performance characteristics guidelines for blade-element momentum methods [C]. Reno, Nevada, 2005.
- [21] Desbiens A L, Cutkosky M R. Landing and perching on vertical surfaces with microspines for small unmanned air vehicles [J]. Intell Robot Syst, 2010, 57: 313-327.
- [22] Hawkes E W, Christensen D L, Eason E V, et al. Dynamic surface grasping with directional adhesion [C]. IEEE/RSJ International Conference on Intelligent Robots & Systems. 2013.
- [23] Hao J, Pope M T, Hawkes E W, et al. Modeling the dynamics of perching with opposed-grip mechanisms [C]. IEEE International Conference on Robotics & Automation. 2014.
- [24] William R T, Roderick W R T, Jiang H, et al. Bio-inspired grippers for natural curved surface perching [J]. Living Machines, 2017: 604-610.

作者简介:

昌 敏(1986-), 男, 博士, 高级工程师, 主要研究方向为多域无人系统设计理论与方法。

孙 杨(1996-), 男, 本科生, 主要研究方向为仿生飞行器设计。

白俊强(1971-), 男, 博士, 教授, 主要研究方向为飞行器先进设计理论与技术。

无人机编队避障与控制技术 研究现状及发展趋势

方洋旺¹, 欧阳楚月², 符文星¹, 程昊宇²

(1. 西北工业大学无人系统技术研究院, 西安 710072; 2. 西北工业大学航天学院, 西安 710072)

摘要: 多无人机编队协同控制技术作为人工智能中集群智能方向的热点研究问题, 国内外学者对多无人机编队协同控制的理论和应用都做出了广泛深入的研究, 而作为协同控制的核心问题的多无人机编队避障和控制问题的研究也取得了一些成果。首先从人工势场、速度障碍和Dubins路径三个方面介绍了单无人机避障与控制技术, 并且针对各个算法的不足之处, 提出了一些改进方法。在此基础上, 又对多无人机编队避障与控制技术和无人机机间防碰撞进行了综述, 分析各种方法的优缺点。最后, 对多无人机编队避障及控制发展趋势进行了分析, 并提出了可行的解决思路和方法。

关键词: 集群智能; 多无人机编队; 协同控制; 编队避障和控制; 机间防碰撞

中图分类号: V279 **文献标识码:** A **文章编号:** 2096-5915 (2019) 02-0410-07

Research Status and Development Trend of Obstacle Avoidance and Control Technology for UAV Formation

FANG Yangwang¹, OUYANG Chuyue², FU Wenxing¹, CHENG Haoyu²

(1. Northwestern Polytechnical University, Unmanned System Research Institute, Xi'an 710072, China;

2. Northwestern Polytechnical University, School of Astronautics, Xi'an 710072, China)

Abstract: Multi-UAV formation cooperative control technology is the hot research issues in cluster intelligence direction in artificial intelligence, domestic and foreign scholars have made extensive and in-depth research on the theory and application of multi-UAV formation cooperative control, and as the core issues of collaborative control, the research on multi-UAV formation avoidance and control problems has also achieved some results. In this paper, the single unmanned aircraft obstacle avoidance and control technology is summarized from three aspects: artificial potential field, speed obstacle and Dubins path. Aiming at the deficiencies of each algorithm, some improved methods are proposed. On this basis, the multi-UAV formation avoidance and control technology and the anti-collision between UAVs are reviewed, and the advantages and disadvantages of various methods are analyzed. Finally, the development trend of obstacle avoidance and control of multi-UAV formations is analyzed, and feasible solutions and methods are proposed.

Keywords: Cluster Intelligence; Multi-UAV Formation; Collaborative Control; Formation Obstacle Avoidance and Control; Anti-Collision Between UAVs

收稿日期: 2019-02-22; 修回日期: 2019-03-15

基金项目: 中央高校基本科研业务费专项资金资助 (G2018KY0301)

1 引言

随着当代军事作战方式的变革和生产成本的降低,无人机已经成为全球军事部门、研究机构和高等院校的研究热点。无人机在军事作战上具有巨大的优势,例如无人机在作战时不携带飞行员,对于危险性高的任务,可以做到零人员伤亡;无人机安装自动驾驶系统进行飞行控制,不需要驾驶舱和安全救生等设备,从一定程度上减小飞机体积,降低成本^[1];无人机由于不需要飞行员,因此能够进行更激烈的动作,在空中的停留时间也不会受到飞行员生理限制的影响^[2]。然而当单架无人机在执行大型复杂的任务时,仍会存在诸多问题。从时间来看,单架无人机续航时间较短,对于一些需要大面积、长时间搜索的复杂任务,无人机的最大行程将被限制;从空间来看,单架无人机的活动半径有限,执行大规模军事侦察任务时将严重降低任务完成效率;从任务层面来看,单架无人机的抗干扰能力、探测能力、武器载荷能力均有限,将增加任务失败概率。与此相比,多无人机协同编队能解决时间、空间、任务层面上的冲突^[3]。当任务复杂、飞行区域较大时,可以将总任务分成一些简单的小任务,每架无人机携带不同的设备分别完成各自的任务,使总任务能够一次完成,从而显著提高任务效率。

无人机在编队飞行执行任务过程中,需要解决队形保持、防撞避障、航迹规划等问题。避障问题是无人机编队执行任务过程中必须要解决的关键问题之一,这一问题包含了无人机编队的机间内部碰撞问题和外界障碍物回避问题^[4]。无人机近距编队飞行时,在编队内部,多机之间要保持一定的距离,最小安全距离是不可逾越的红线;在编队外部,空域中的障碍物是重大威胁,同样不能小于最小安全距离。如何避免发生碰撞是确保编队安全性的核心内容,也是目前无人机协同编队制导与控制技术研究的重点和难点。

2 单无人机避障与控制技术

2.1 人工势场法

人工势场法由Khatib于1985年提出^[5],实际上是一种拟物方法,使障碍物的分布情况等信息反映在环境中每一点的势场值当中,根据势场值的大小,决定无人机的行进方向和速度。人工势场法因计算量小、实时性好、意义明确、便于实现等特点得到了广泛的研究和应用,但传统人工势场法存在易陷入局部最优、在狭窄通道中存在航迹抖动、对于动态障碍物规避效果不佳等问题。针对上述问题,国内外相关学者进行了大量的研究来改进人工势场法。

目前解决无人机动态障碍物避障问题比较成熟的是“感知-规避”方法,其中,碰撞锥方法就是典型的“感知-规避”方法。碰撞锥概念由文献[6]提出,该方法在无人机动态障碍物避障、实时碰撞检测和威胁评估方面有着巨大优势。碰撞锥基于无人机和障碍物相对位置矢量和相对速度矢量的简单几何关系,即可实现对两个未知轨迹的移动目标进行碰撞检测,并能实时预测无人机与障碍物的碰撞位置及碰撞时间。碰撞锥因其快速准确的碰撞预测和评估能力,被广泛应用于动态环境下无人机的实时碰撞检测与威胁评估。文献[7]利用碰撞锥的碰撞检测条件与威胁评估能力来改进人工势场函数,提出一种基于碰撞锥改进人工势场的无人机避障路径规划算法。文献[8]针对不同属性的障碍物所构成的威胁分布模型,首先对障碍物分布建立栅格化模型;然后根据静态障碍物,基于稀疏 A^* 搜索算法进行全局航迹规划;最后结合预规划路径和动态障碍物分布,利用改进人工势场法完成无人机的动态避障,提出了一种基于稀疏 A^* 搜索算法预规划和改进人工势场法相结合的无人机动态避障算法。文献[9]基于生物群集理论,针对多智能体系统的避障问题,在传统势场函数的基础上,提出了静态障碍环境下的多智能体编队避障策略,其中包括基于

群集理论的分布式群组运动模型、基于极限环理论的智能体避障控制算法。该方法提出了清晰的多智能体系统的协调控制框架，通过算法方面的改进和优化，有效地克服了人工势场法在智能体避障控制中易陷入局部极值，以及障碍物周围易发生振荡等缺陷，为多智能体的编队控制问题提出了新思路。文献 [10] 基于虚拟结构和“长机-僚机”控制策略，提出一种三维空间复合矢量人工势场方法，复合人工势场由平行 xy 平面和平行 yz 平面的人工势场复合而成，可以避开障碍物周围每个局部最小位置，而且以最优路径快速绕过障碍物，集结编队飞行到达目标点，斥力能够避免无人机之间发生碰撞，同时也避免无人机与障碍物发生碰撞，达到避障的目的。文献 [11] 利用分段思想对势场函数进行修正，解决了目标不可达问题，引入回环力和飞行器间作用力，解决了局部极小值问题和飞行器间的碰撞问题。文献 [12] 增加一个指数项到引力场函数中，然后将一敏感度参数引入斥力场函数，该方法消除了奇异值点，避免了抖动现象，灵活控制运动过程中无人机与障碍物距离的大小，新的人工势场函数不仅有传统人工势场法简洁、快速的优点，还可以将其运用到实时避障控制中，但不适用于高速无人机。

2.2 速度障碍法

最早 Khatib 和 Erdmann 等人开始研究如何使无人机安全导引到目标点并顺利避开动态障碍物，后来，学者们把在静态环境下应用成功的避障方法推广到动态环境中去。文献 [13] 中通过考虑无人机及障碍物的速度，结合到势能函数的构造中，得到适用于动态环境的势场法。文献 [14] 中通过动态窗口方法进行路径规划及避障，这些方法在静态环境下得到较好的结果，但在无人机高速运动情况下，避障效果并不理想。主要原因是在动态环境下，碰撞判断及避障情况下对速度的控制与静态环境相比存在较大区别。为解决动态环境下的碰撞判断问题，文献 [15-16] 提出

了基于运动学方程的碰撞判断方法，文献 [15] 定义了碰撞锥的概念，根据视线的距离及角速度来确定碰撞条件，文献 [16] 中基于视线角速度方程，对直接碰撞和间接碰撞进行了比较，文献 [17] 通过对某一时刻，在考虑无人机及障碍物相对位置情况下，对可能的碰撞速度进行时间积分，得到非线性速度障碍集进行避障。文献 [18] 基于无人机与障碍物之间的几何关系，在极坐标系下建立了无人机与障碍物之间的运动学方程，通过引入基于角度线性化的比例导引律，使无人机能够顺利避开障碍物。文献 [19] 在极坐标系下建立无人机与障碍物之间的几何关系并且建立无人机与障碍物之间的运动学方程，通过设计滑模变结构有限时间收敛制导律，使无人机与避障点的视线角速率收敛到零，相对速度方向收敛到期望的避障方向，使无人机能够顺利避开运动障碍物。无人机与障碍物几何关系如图 1 所示。

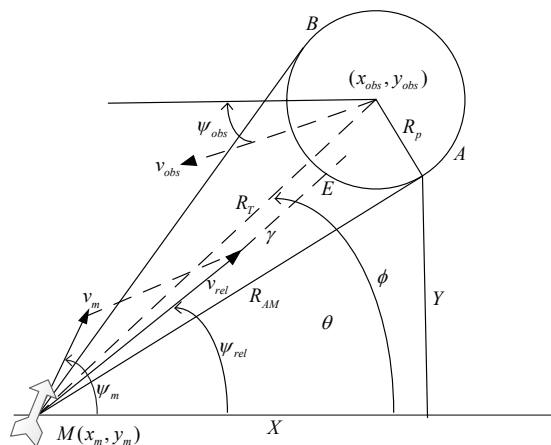


图1 无人机与障碍物几何结构示意图

Fig.1 Schematic diagram of the geometry of the drone and obstacle

2.3 Dubins 路径法

Dubins 路径定义为：在最大曲率限制下，平面内两个有方向的点间的最短可行路径是 CLC 路径或 CCC 路径，或是其子集，其中， C 表示圆弧段， L 表示与 C 相切的直线段。文献 [20] 提出一种基于路径的无人机的避障规划算法，采用遗传算法，结合无人机的飞行性能和最小转弯半径得到一种在已知障碍空间位置前提下的无人机路径规划方

法, 通过算法改进, 将其推广成为在未知障碍物位置等先验知识的前提下的无人机实时避障算法。该方法可以解决多障碍环境下无人机避障控制实时生成的问题, 其路径较传统避障方法获取的路径而言, 考虑了无人机自身的运动学特性, 在工程应用中有很好的实用性。

3 多无人机编队避障与控制技术

以上文献的研究主要集中在单无人机避障与控制技术上, 对于多无人机编队避障与控制技术, Shin Hyo-Sang用几何分析法预测碰撞冲突, 研究了只控制偏航角的冲突规避法, 然而多无人机编队避障与控制技术还有很广阔的研究空间。

3.1 多无人机编队防碰撞

多无人机编队避障与控制是指在多无人机协同执行任务的过程中, 必须保证编队中每一架无人机规避战场环境中的障碍物, 并且编队中的无人机之间也不能发生碰撞。如果多无人机编队规模较小, 任意一架无人机的损失将使多无人机编队的战斗力下降; 而如果多无人机编队规模较大, 无人机之间的碰撞或是无人机与障碍物的碰撞可能会引发连锁反应, 造成其它无人机的碰撞。多无人机编队控制中的实时避障问题可归结为单个无人机进行实时航迹规划, 但是当编队规模较大时, 需要进行协同避障, 这样才能具有更好的控制效果和良好的扩展性。

文献 [21] 利用协同航迹规划方法解决多机防碰撞问题, 但只适用于合作型无人机, 仅在很少情况下适用。文献 [22] 使用几何最优法, 以入侵机为原点修正本机速度和位置, 速度矢量和保护圆相交则冲突, 选择改变偏航角使速度和圆相切, 但只在二维平面内适用。文献 [23] 以两架固定翼无人机在同高度、有障碍物环境下的路径规划为应用背景, 针对传统基于网格的 A^* 搜索算法没有考虑飞行性能约束的问题, 首先利用 Dubins 曲线找到有效节点, 结合 A^* 启发式搜索的

思想, 分别为每架无人机离线构建由 Dubins 曲线组成的最短避障路径; 在两机同高度飞行过程中, 通过相对运动关系判断是否碰撞, 使用“向量共享”解得航向改变量以及在线路径重规划, 得到两机的防撞避障路径。该方法的不足之处是仅考虑了两架无人机的航迹安全和避障等基本问题, 没有考虑如何拓展至更多无人机执行复杂任务的情况。文献 [24] 针对现有防撞控制方法中的无人机编队不能充分利用机间通信、无法保持固有队形的缺陷, 提出人工势场与信息一致性相结合的方法, 融合这两种方法的优点, 在防撞的同时不仅充分利用了机间通信信息, 又能够实现队形保持。下图所示为人工势场示意图:

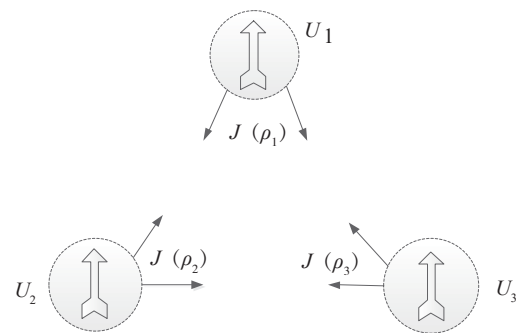


图2 人工势场示意图

Fig.2 Schematic diagram of artificial potential field

基于人工势场与信息一致性的联合防撞碰撞控制算法如下所示:

$$V_i^{NSB} = [V_i^{NSB} \cos \vartheta_i^{NSB} \cos \psi_i^{NSB}, V_i^{NSB} \cos \vartheta_i^{NSB} \sin \psi_i^{NSB}, V_i^{NSB} \sin \vartheta_i^{NSB}]^T$$

文献 [25] 通过增加方向协调力解决人工势场法在动态复杂的环境中存在局部极小点及受迫性碰撞的问题, 然后将改进的人工势场法应用于每个无人机中, 通过多机之间的协同和避障的条件约束, 能够实现多无人机同时到达目标, 且实现无人机之间的避障和在工作区间内躲避障碍物。文献 [26] 基于虚拟结构和人工势场的方法设计了无人机编队避障控制, 虚拟结构法定义虚拟结构来表示编队队形, 各无人机通过将队形误差反馈到自身控制器中, 能够实现高精度的队形保持,

但在编队飞行过程中的障碍物规避问题没有得到解决。人工势场法中，通过设计合适的势场函数可以实现障碍物的规避，但对队形的保持能力较差。将两个算法相结合，编队可以保持队形，在有障碍物时，各无人机可以顺利避障。文献 [27] 详细研究了密集编队的避障问题，建立了透光性假说的避障模型，并基于 LMPC 控制理论设计了密集编队无人机防碰撞的飞行控制器。但其只研究了密集编队无人机内部防碰撞的问题，并没有研究如何避障的问题。

3.2 多无人机机间防碰撞

文献 [28] 提出一种改进人工势场方法解决多无人机机间防碰撞问题。首先通过建立无人机的引力与斥力受力模型，在此基础上，将每架无人机视为移动障碍物，自身产生的斥力势场来避免与其他无人机之间的相互碰撞。当人工势场法陷入局部极小值时，通过构造虚拟障碍物的方法，使得无人机逃离局部最优；提出分段 Bezier 曲线对路径进行在线实时平滑优化，在两个分段曲线连接处保证了曲线的平滑性，防止无人机在飞行过程中做大角度转弯，平滑后的路径不仅不再出现路径振荡现象，而且整条航线变得更加平滑顺畅。文献 [29] 针对传统人工势场法进行航迹规划易陷入局部极小点、发生胁迫性碰撞、目标点无法到达的问题，提出一种基于改进人工势场法的无人机航迹规划算法。该方法通过在合力场中引入方向协调力及无人机与目标点的距离来解决传统人工势场法存在的问题，引入目标点和无人机之间距离的影响，找到一个新的斥力函数，当无人机接近目标点时，斥力逐渐变为零，使整个合势场能够在目标点达到最小，使障碍物与无人机在目标点附近，从而能够到达目标点。通过新建立的势场函数可以保证在目标点达到全局最小，避免了目标不可达问题，但局部最小问题仍然存在。当障碍物和目标点在同一条直线上，无人机受力的控制，只能在直线上反复运动。为解决不能到达目标点的问题，改进方案中加入方向协调

力，在计算 x 轴方向的合力时，引入 y 轴方向的方向协调力，在计算 y 轴方向的合力时，引入 x 轴方向的方向协调力。文献 [30] 针对无人机编队内部避障问题，首先采用基于预测控制的方法，解决了在即将发生碰撞情况下无人机的航向调整问题；然后建立航路回归模型，解决了在无人机避障后迅速回归预定航路的问题。这种方法的一个优点是反应速度快，能够实时在线处理各种信息，符合文中对无人机编队避障控制的需求。无人机预测控制模型流程图如图 3 所示。

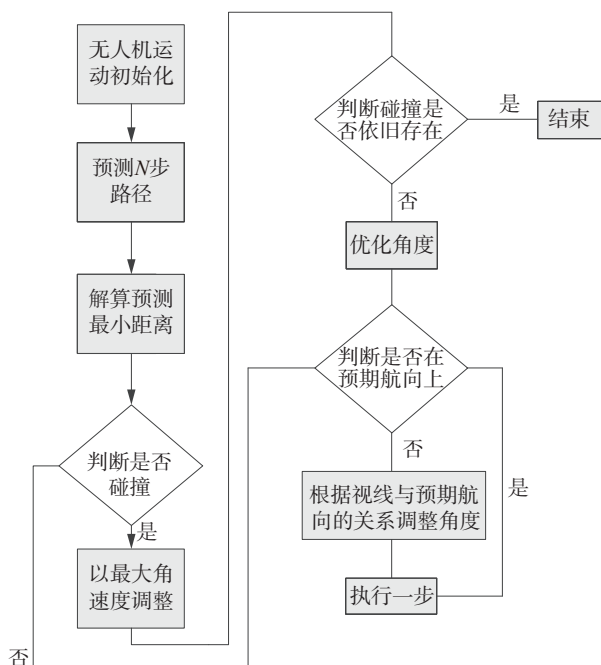


图3 预测控制模型流程图

Fig.3 Flow chart of predictive control model

4 无人机编队避障与控制技术发展趋势

4.1 密集编队的避障与控制技术

目前对于密集无人机或蜂群无人机的避障与控制技术研究较少，主要原因是由于密集编队碰撞很容易产生链式效应，由于每架无人机可机动的空间很小，一旦某无人机发生碰撞，将很快影响邻近的无人机，但邻近无人机为了防止碰撞向周边作出机动时可能又会碰到其他的无人机，而

其他的无人机又作规避机动，又碰到其邻近无人机，这种碰撞就像链式反应一样，或称为链式效应。因此，避障与控制算法的主要难点是：如何确定无人机周围可容许机动的自由空间，以及如果机动空间不够时如何处理确保链式反应不要进行下去。目前，基于鸟类透光避障原理模型已经应用到密集编队无人机避障与控制中，但对于密集编队规避静态或动态障碍物碰撞的技术尚未见报道。可以考虑将现有复杂网络的分割方法、各种协同避障算法与透光性避障模型相结合，利用复杂网络的分割方法将密集编队无人机划分成一些子群组。首先以子群组为单元利用规划调度的方法进行避障子群组控制设计；然后基于透光性模型解决无人机可机动的自由空间问题；再采用虚拟力或调整高度的方法设计密集编队在可机动自由空间约束条件下的子群组成员避障控制算法。

4.2 干扰条件下的避障与控制技术

当无人机编队接近障碍物时，各种地理环境的自然干扰和人为干扰对无人机编队传感器探测、编队间的通信以及网络拓扑结构都产生很大的影响，从而造成网络拓扑结构时变、通信产生延时和丢包以及传感器测量误差增大等问题。目前基于干扰环境下无人机编队避障与控制技术的研究较少，可以考虑将障碍物也看作无人机编队的一个节点构成新的无人机编队，同时在干扰条件下的多智能体模型中加入各种虚拟力函数解决碰撞问题。

4.3 密集障碍物的避障与控制技术

当无人机编队在城市或山区飞行时，有可能遇到密集障碍物的避障问题，其主要难点一方面是各种干扰不可避免，另一方面可能在短时间内需要进行编队拆分、编队形成等过程。当基于多智能体一致性原理设计避障控制时，一方面考虑网络拓扑结构时变、通信产生延时和丢包以及信息精度不高的问题，另一方面要考虑多智能体一致性收敛时间的问题，即如果基于干扰条件下多

智能体通过收敛达到一致性的时间大于飞行两个相邻两个障碍物需要的时间时，多智能体一致性条件就无法满足。因此，可以考虑采用将有限时间收敛的控制方法与干扰条件下多智能体避障控制算法结合，给出密集障碍物的避障与控制算法。

5 结束语

本文首先针对人工势场法存在易陷入局部最优、航迹抖动、对动态障碍物规避效果不佳的问题，介绍了几种改进的人工势场法。提出将静态环境下应用成功的速度障碍法推广到动态环境的思路，以及改进Dubins算法，推广为在多未知障碍环境下的实时避障控制。在单无人机控制基础上，进一步分析了多无人机编队避障与控制技术在不同方法下的优缺点，并介绍了多无人机机间防碰撞的常用方法。最后分别提出了避障与控制技术在密集编队、干扰环境、密集障碍物条件下的技术难点，并提出了可行的解决思路，为研究无人机避障和控制技术的科研人员提供借鉴和参考。

参考文献

- [1] 沈林成, 张庆杰, 王林, 等. 美国及欧洲无人作战飞机计划发展 [J]. 火力与指挥控制, 2009, 34 (10): 8-14.
- [2] 刘重阳. 国外无人机技术的发展 [J]. 舰船电子工程, 2010, 30 (1): 19-23.
- [3] 朱华勇, 张庆杰, 沈林成. 提高无人作战飞机协同作战能力的关键技术 [J]. 控制工程, 2010 (s1): 52-55.
- [4] 王京. 多智能体系统编队避障算法研究 [D]. 上海: 华东理工大学, 2013.
- [5] Khatib O. Real-Time Obstacle Avoidance for Manipulators and Mobile Robots [J]. 1986.
- [6] Chakravarthy A, Ghose D. Obstacle avoidance in adynamic environment: a collision cone approach [J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 1998, 28 (5): 562-574.
- [7] 熊超, 解武杰, 董文瀚. 基于碰撞锥改进人工势场的无人机避障路径规划 [J]. 计算机工程, 2018, 44 (9):

- 314-320.
- [8] 姚远, 周兴社, 张凯龙, 等. 基于稀疏 A* 搜索和改进人工势场的无人机动态航迹规划 [J]. 控制理论与应用, 2010, 27 (7): 953-959.
- [9] 曹建福, 凌志浩, 高冲, 等. 基于群集思想的多智能体编队避障算法研究 [J]. 系统仿真学报, 2014, 26 (3): 562-566.
- [10] 张佳龙, 闫建国, 张普, 等. 基于改进人工势场的无人机编队避障控制研究 [J]. 西安交通大学学报, 2018, 52 (11).
- [11] 代冀阳, 王村松, 殷林飞, 等. 飞行器分层势场路径规划算法 [J]. 控制理论与应用, 2015, 32 (11): 1505-1510.
- [12] 樊晓平, 李双艳, 陈特放. 基于新人工势场函数的机器人动态避障规划 [J]. 控制理论与应用, 2005, 22 (5): 703-707.
- [13] Ge S S . Dynamic motion planning for mobile robots using potential field method [J]. Autonomous Robots, 2002, 13 (3): 207-222.
- [14] Seder M , Petrović I . Dynamic window based approach to mobile robot motion control in the presence of moving obstacles [C]. CD ROM Proceedings of the 2007 IEEE International Conference on Robotics and Automation Roma, Italy, 10-14 April 2007.
- [15] Chakravarthy A , Ghose D . Obstacle avoidance in a dynamic environment: a collision cone approach [J]. Systems Man & Cybernetics Part A Systems & Humans IEEE Transactions on, 1998, 28 (5): 562-574.
- [16] Belkhouche F , Belkhouche B . Kinematics-Based Characterization of the Collision Course [J]. International Journal of Robotics & Automation, 2008, 23 (2): 127-136.
- [17] Large F , Sckhavat S , Shiller Z , et al . Using non-linear velocity obstacles to plan motions in a dynamic environment [C]. International Conference on Control. IEEE, 2003.
- [18] 杨秀霞, 刘小伟, 张毅. 基于比例导引角度线性化的无人机避障研究 [J]. 导航定位与授时, 2014, 1 (3): 1-5.
- [19] 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.
- [20] 关震宇, 杨东晓, 李杰, 等. 基于Dubins路径的无人机避障规划算法 [J]. 北京理工大学学报, 2014, 34 (6): 570-575.
- [21] Lain T M . Cooperative rendezvous of multiple unmannedair vehicles [C]. AIAA GNC, 2000. and Exhibit, Hawaii, 2008.
- [22] Bilimoria K D , Sridhar B , Chatterji G B , et al . FACET: FUTURE ATM CONCEPTS EVALUATION TOOL [J]. Proceedings of the 3rd USA/Europe ATM 2001 R&D Seminar, 2000, 9.
- [23] 宋雪倩, 胡士强. 基于Dubins路径的A*算法的多无人机路径规划 [J]. 电光与控制, 2018, 25 (11): 29-33.
- [24] 朱旭. 基于信息一致性的多无人机编队控制方法研究 [D]. 西安: 西北工业大学, 2014.
- [25] 王超瑞. 基于信息一致性理论的无人机编队控制算法研究 [D]. 哈尔滨: 哈尔滨工业大学, 2017.
- [26] 成成. 多无人机协同编队飞行控制关键技术研究 [D]. 北京: 中国科学院大学, 2018.
- [27] 吴森堂. 飞行控制系统 [M]. 北京: 北京航空航天大学出版社, 2013.
- [28] 杨勇, 丁勇, 黄鑫城. 改进APF与Bezier相结合的多无人机协同避障航路规划 [J]. 电光与控制, 2018, 25 (11): 40-45+51.
- [29] 张思宇. 多无人机协同航迹规划及其控制方法研究 [D]. 北京: 北京理工大学, 2016.
- [30] 魏健, 李相民, 代进进. 基于预测控制的无人机编队内部避障 [J]. 海军航空工程学院学报, 2015, 30 (4): 387-391.

作者简介:

方洋旺(1966-), 男, 博士, 教授, 主要研究方向为随机最优控制、导航制导与控制、非线性控制、非线性信号处理及智能信号处理等。

欧阳楚月(1994-), 女, 硕士研究生, 主要研究方向为飞行器协同制导与控制、编队避障技术。

符文星(1973-), 男, 博士, 教授, 主要研究方向为飞行器智能化、制导控制系统设计与仿真等。

程昊宇(1990-), 男, 博士, 助理教授, 主要研究方向为飞行器制导控制系统设计、飞行智能控制算法、分布式智能空战等。

基于Voronoi图质心的多无人机协同区域搜索算法

朱利¹, 符小卫²

(1. 中国电子科技集团第十研究所, 成都 610036; 2. 西北工业大学电子信息学院, 西安 710129)

摘要: 为了在最短时间内降低整个区域的环境不确定度, 提高多无人机协同区域搜索效率, 提出了一种基于 Voronoi 图 (V图) 质心的多无人机协同区域搜索算法。首先, 建立了相关数学模型, 并设计搜索效能评估指标; 然后, 针对不断变化的环境信息, 提出了 Discrete Combined Partition and Search (DCPS) 策略, 该策略将 V图质心用于引导无人机运动; 该策略下, 任务区域在每一步搜索时都会被重新进行 V图划分, 从而提升了搜索效率; 接着, 研究了传感器探测半径约束对于搜索算法的影响; 最后, 仿真不同情况下多无人机协同区域搜索效率的变化。通过对比仿真说明, DCPS策略能够保证多无人机协同区域搜索的鲁棒性, 并尽快地降低整个区域的环境不确定度。

关键词: 多无人机; 协同搜索; V图质心; DCPS策略; 传感器半径约束; 环境不确定度

中图分类号: V279+.3 **文献标识码:** A **文章编号:** 2096-5915 (2019) 02-0417-13

Multiple UAVs Cooperative Area Search Algorithm Based on Centroid of Voronoi Diagram

ZHU Li¹, FU Xiaowei²

(1. China Electronics Technology Group Corporation No.10 Research Institute, Chengdu 610036, China;

2. Northwestern Polytechnical University, School of Electronics and Information, Xi'an 710129, China)

Abstract: In order to reduce the environmental uncertainty of the whole mission area in the shortest time and improve the efficiency of multi-UAV cooperative area search, a new cooperative area search algorithm based on the centroid of Voronoi Diagram is proposed. Firstly, mathematical models are established and search effectiveness evaluation indexes are designed. Then, according to the changing environment information, a Discrete Combined Partition and Search (DCPS) strategy is proposed, which uses the centroid of Voronoi Diagram to guide UAVs movement. Under this strategy, the mission area will be re-divided into Voronoi Diagram in each step of search, which improves the search efficiency. Then, the influence of sensor detection radius constraint on search algorithm is studied. The simulation validates the change of search efficiency of multi-UAV cooperative area under different conditions. The simulation results show that the DCPS strategy can guarantee the robustness of multi-UAV cooperative area search and reduce the environmental uncertainty of the whole region as soon as possible.

Keywords: Multiple UAVs; Cooperative Area Search; Centroid of Voronoi Diagram; DCPS Mechanism; Sensors Radius Constraint; Environmental Uncertainty

收稿日期: 2019-02-28; 修回日期: 2019-03-07

基金项目: 国家自然科学基金 (61573285)

1 引言

无人机是一种由动力驱动，机上为无人驾驶，可以重复使用的飞行器^[1]。在无人机承担的各种作战任务中，搜索和侦察是目前无人机系统的首要任务^[2]，对任务区域执行侦察和搜索并发现目标是进行情报搜集、目标定位乃至打击等一系列作战活动的前提。这种无人机通常被称为无人侦察机，典型的包括中国的翔龙和美军的全球鹰。

组织多架无人机协同执行任务是未来网络化战场上的一种重要作战方式，而如何有效控制多架无人机在复杂环境中协同执行搜索任务也逐渐成为近年来协同控制领域的研究热点^[3]。

在搜索环境信息描述与建模方面，最主要的方法是基于搜索区域网格化的建模，例如占用图^[4-6]、概率图^[7-9]、不确定图^[10-12]等。总的来说，这类方法在本质上都是构造一个二维离散的网格地图，该网格地图可以反映无人机对当前目标和环境信息的认知情况，这样的二维离散的网格地图统称为环境感知地图。随着搜索任务的执行，无人机对环境的不断感知，环境感知地图则按照合理的规则动态更新，以便反映最新的环境和目标信息，从而使得无人机可以有效利用实时探测信息，适用于动态搜索过程。

在多无人机搜索决策方面，主要是区域分割方法^[13-15]，这类方法主要将区域根据不同的指标划分给各架无人机，再由各架无人机对各自的搜索区域进行搜索，从而达到协同搜索。滚动时域优化(Receding Horizon Optimization)方法^[16-18]，同时考虑了无人机短期和长期搜索代价，能够从宏观上提升无人机的搜索效率。

Voronoi图(V图)在搜索路径中已经有了广泛的应用，如果障碍物可以近似成质点，那么机器人沿着障碍物的V图的边行走是最安全的。如果障碍物不能用质点来近似，那么就该应用扩展的V图(站点为线段、多边形或多面体等)。同时，考虑V图本身的特殊性，黄长强等^[19]提出基于质心

V图划分的无人机搜索决策方法对多架无人机进行搜索空间的分配，建立了质心V图的数学模型，并在此基础上提出了基于质心V图的目标分配算法，结果表明提出的搜索策略能够有效进行随机目标搜索且算法具有良好的自适应能力。

总的来说，国内外学者在多无人机协同搜索、传感器信息感知、V图路径规划方面已经做了大量的研究工作，并且取得了丰富的研究成果，为接下来的研究奠定了坚实的基础。由于问题的复杂性，目前还存在一些问题有待进一步研究，主要表现为以下几个方面：对多机协同的控制；对传感器进行更符合现实规律的建模；对V图动态划分的研究。

针对上述问题，本文将V图引入多机协同搜索中来。首先，针对多机协同搜索问题，以最短时间内降低环境不确定度为目标，结合环境建模，根据各架无人机位置将环境进行V图划分，各架无人机在自身所在的V图区域内进行搜索，随着搜索的不断进行，各自V图区域内的环境不确定度不断降低。且通过无人机之间的通信，各无人机能够掌握当前整个环境的不确定度，从而根据当前环境信息的重新规划路径。针对单无人机在V图区域内如何运动的问题，将环境不确定度作为网格质量求解各V图区域的质心，并证明得到无人机在各自的V图质心处搜索时，环境不确定度降低最大，同进提出了Discrete Combined Partition and Search(DCPS)策略，并考虑在传感器半径约束下搜索策略的有效性。该策略下，每搜索一次都会进行一次V图划分，所有无人机在每个时刻都处于搜索状态，保证了所有无人机的充分利用。因此，DCPS算法能够在保证协同的基础上同时保证搜索的有效性。

2 问题描述

如图1所示，假设 n 架无人机在未知环境 $E = \{(i, j) | i = 1, 2, \dots, L_x, j = 1, 2, \dots, L_y\}$ 中搜索，为降低整个环境的不确定度，问题可以描述为：

(1) $E = \{(i, j) | i = 1, 2, \dots, L_x, j = 1, 2, \dots, L_y\}$ 为一个矩形待搜索区域;

(2) $\phi(i, j) \rightarrow [0, 1]$ 表示网格 (i, j) 处的不确定度;

(3) n 架无人机, 配备了相同的传感器和通讯设备, n 架无人机在区域 E 中搜索, 收集信息, 从而降低环境的不确定度;

(4) $P(t) = \{p_1(t), p_2(t), \dots, p_N(t)\}$ 表示无人机 t 时刻在搜索区域 E 中的位置。 $p_i(t)$ 表示第 i 架无人机在 t 时刻的位置, 为了方便起见, 在后文中我们直接使用 p_i 表示 $p_i(t)$;

(5) 传感器对位置 (i, j) 的搜索效率随传感器到 (i, j) 的距离变大而降低;

(6) 在理想情况下, 我们可以寻找一种利用无人机获取关于搜索区域 E 的完整信息的最佳方法, 从而达到 $\phi(i, j) = 0, \forall (i, j) \in E$.

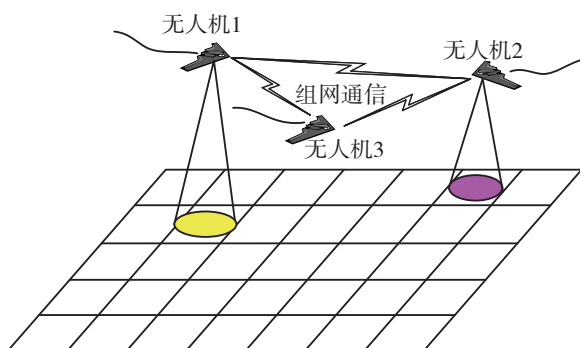


图1 多无人机协同区域搜索

Fig.1 Illustration of multiple UAVs cooperative area search

3 协同区域搜索建模

3.1 无人机平台模型

假设无人机 i 在时刻 k 的状态为 $state_i(k)$, $state_i(k)$ 由以下三个方面构成:

$$state_i(k) = [\lambda_i(k), o_i(k), \delta_i(k)] \quad (1)$$

其中 $\lambda_i(k) = (x_i(k), y_i(k)) \in \{1, 2, \dots, L_x\} \times \{1, 2, \dots, L_y\}$ 表示第 i 架无人机在 k 时刻的位置; $o_i(k) \in \{0, 1, 2, \dots, 7\}$ 表示第 i 架无人机在 k 时刻的方位且分别

定义为 $\{0$ (东), 1 (东北), 2 (北), 3 (西北), 4 (西), 5 (西南), 6 (南), 7 (东南) $\}$; $\delta_i(k) \in \{0, 1\}$ 表示在 k 时刻第 i 架无人机是否已经被摧毁, 或者是由于故障不能执行搜索任务, $\delta_i(k) = 0$ 表示无人机由于被摧毁或者是故障不能正常执行搜索任务, $\delta_i(k) = 1$ 表示无人机状态良好, 可正常执行搜索任务。

由于无人机的动力学模型服从转弯半径的限制, 假设无人机的最大转弯角度为 45° , 在这种转弯半径的限制下, 各架无人机在下一个时间步长内有三种可选位置, 分别为左转、直行和右转, 可以表示为 $\{l(left), f(front), r(right)\}$, 无人机当前航向与下一时间步长内可选航向关系如图2所示。

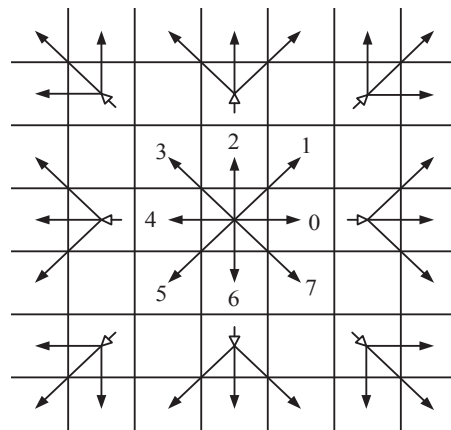


图2 无人机当前航向与下一时间步长内可选航向

Fig.2 Current heading and next step heading of an UAV

3.2 战场环境建模

在多无人机协同搜索过程中, 由于环境是动态变化的, 因此建立良好的环境模型至关重要, 需要其能够基于多个传感器信息和多架不同无人机进行构建和更新, 环境模型的有效性将最终影响无人机的决策行为。基于上述讨论, 本文采用环境不确定度模型来描述任务区域各个单元格。首先将搜索区域划分为 $L_x \times L_y$ 的离散单元, 并且称所有矩形单元的集合 $E = \{(i, j) | i = 1, 2, \dots, L_x, j = 1, 2, \dots, L_y\}$ 为搜索环境, (i, j) 表示位于第 i 行第 j 列的单元, 如图3所示。

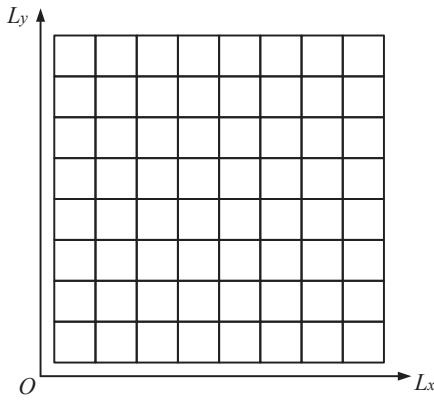


图3 搜索环境栅格化
Fig3 Discreted search area

假设有 n 架完全相同的无人机在区域内执行搜索任务，无人机使用传感器感知环境，收集环境中的信息和定位目标。假设无人机之间有一定的通信能力，能保证无人机之间信息交换的正确性，在该模型中，无人机使用地图存储环境中每个单元的状态信息，将每架无人机视为一个质点，在每一个步长都可以满足无人机操纵限制的前提下，从一个单元 (x_1, y_1) 的中心点移动到与之相邻的单元 (x_2, y_2) 的中心点，并在此处执行搜索任务。

每一个单元 (i, j) 都有一个属性值 ϕ 表示此单元的不确定度，也就是无人机对该网络的确定程度， $\phi(i, j)$ 是指某单元 (i, j) 的不确定度， $\phi(i, j) = 1$ 表示无人机对该处的环境信息一无所知， $\phi(i, j) = 0$ 则表示无人机完全了解该处的环境信息，地图中全部单元的 ϕ 值的集合为 Φ 。由于受到通信距离和通信角度的约束，导致各无人机之间不能实时交换不确定地图信息，因此对于不同的无人机来说，不确定地图是不同的。

因此第 i 架无人机 k 时刻的不确定地图为：

$$\Phi_i(k) = \{\phi_i(i, j, k) \mid i \in 1, 2, \dots, L_x, j \in 1, 2, \dots, L_y\} \quad (2)$$

3.3 传感器模型

用函数 f 表示传感器搜索效率：

$$f(r) = ke^{-\alpha r^2}, k \in (0, 1) \text{ and } \alpha > 0 \quad (3)$$

其中， r 为待搜索位置到传感器的距离，当 $r = 0$ 时， f 最大， $r \rightarrow \infty$ 时， f 最小且趋近于零。 k

和 α 为两个可调参数，可以为多种传感器建模。在 $k = 0.8, \alpha = 0.1$ 时，传感器的探测效率如图 4 所示。

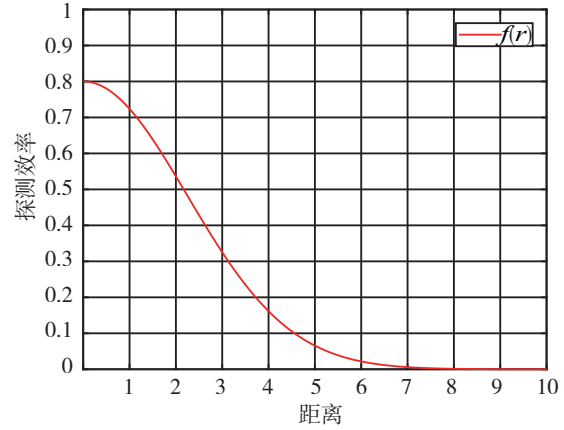


图4 传感器效率更新曲线
Fig.4 Sensor detection efficiency

本文将传感器探测半径进行离散处理，假设在某个区域内的传感器探测效率是相同的，在 $k = 0.8, \alpha = 0.1$ 时，传感器的探测效率如图 5 所示。

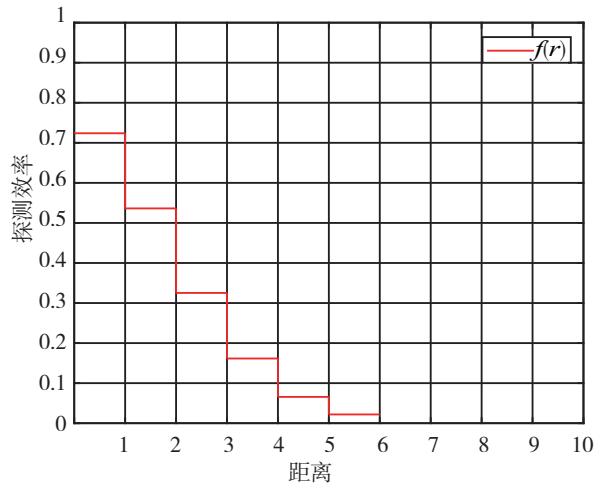


图5 离散传感器效率更新曲线
Fig.5 Discreted sensor detection efficiency

栅格化地图后传感器的探测效率如图 6 所示，搜索区域距离无人机位置可被离散为 $r = 1, 2, \dots, n$ ，假设无人机位于 O 处，颜色相同的区域表示无人机的探测效率是相同的。

如图 6 所示，灰白色区域表示传感器探测半径为 1 的所有网格区域，红色区域表示传感器探测半

径为2的所有网格区域，蓝色区域表示传感器探测半径为3的所有网格区域。也就是说无人机对灰白色区域的探测效率为 $f(1)=ke^{-a^1}$ ，无人机对红色区域的探测效率为 $f(2)=ke^{-a^2}$ ，无人机对蓝色区域的探测效率为 $f(3)=ke^{-a^3}$ 。假设网格 (i, j) 在图6中红色区域，当前网格的不确定度为 $\phi(i, j)$ ，经过传感器探测之后，当前网格的不确定度变为 $\phi(i, j) \cdot (1 - ke^{-a^2})$ 。

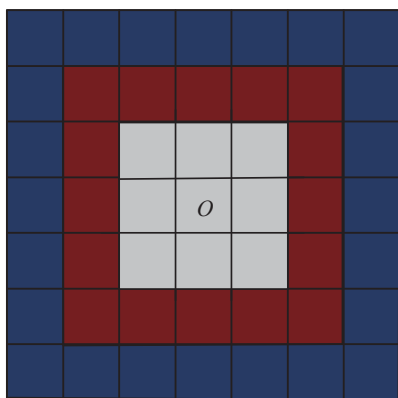


图6 传感器探测半径模型

Fig.6 Sensor detection radius model

3.4 搜索效能评估指标

多无人机协同区域搜索的主要目标为在保证各架无人机飞行安全的前提下，以最短的时间降低整个区域的平均不确定度，在本文中我们采用相同时间步长下整个区域平均不确定度作为无人机系统协同搜索的评价指标。

整个区域的平均不确定度可以表示为：

$$\text{Average Uncertainty} = \frac{\sum_{i=1,2,\dots,L_x, j=1,2,\dots,L_y} \phi(i, j)}{L_x \cdot L_y} \quad (4)$$

当无人机每步搜索都能最大的降低整个区域的不确定度，那么区域的平均不确定度当然也是最低的，因此评价指标可以变为区域不确定度降低程度。

无人机每步搜索的不确定度降低的最大值如式(5)表示：

$$H = \max_{i=1,2,\dots,L_x, j=1,2,\dots,L_y} \Delta\phi(i, j) \quad (5)$$

因此本文需要研究无人机如何运动从而使其每步运动都能最大限度的降低整个区域的不确定度。

4 基于V图质心的区域搜索策略

针对多无人机协同区域搜索的主要目标，根据各架无人机当前位置对搜索区域进行V图划分。在对环境建模时将搜索区域划分为 $L_x \times L_y$ 的离散网格，每个网格 (i, j) 都有一个属性值 $\phi(i, j)$ 代表该网格的不确定度，将当前环境不确定度作为网格质量求解各无人机V图的质心。将V图质心用于引导无人机运动。在文献[20]搜索策略的基础上，考虑无人机的运动特性，将环境地图、无人机运动、传感器探测半径都进行了离散处理，提出DCPS策略。该策略下，能够根据无人机当前位置重新划分V图区域，通过当前的环境不确定度重新求解质心，保证了无人机的搜索效率。但在DCPS策略下，每个搜索时间步长都会重新划分V图区域，当无人机数量太多时，V图划分的计算量会变大，因此，当无人机数量很大时，DCPS策略将不再适用，本文考虑无人机数量较少(10个以内)的情况，搜索策略计算量不大，可以保证在线实现，因此本节在计算量上不做过多研究，将研究重点放在无人机的搜索效率上。

4.1 栅格化的V图质心求解

将区域栅格化后再根据生成元位置进行V图划分可以得到栅格化的V图划分^[21]，如图7所示。

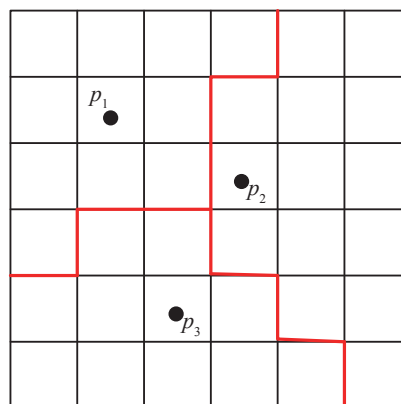


图7 栅格化V图划分

Fig.7 Discretized Voronoi Diagram

图7中， p_1, p_2, p_3 为栅形区域中的三个点，红色折线以及边界将栅形区域划分为三个Voronoi

区域，Voronoi边由各个小网格的边构成。假设 P_i 点的V图区域内有 n 个网格，且其网格中心的坐标分别为 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ ，本文用网格不确定度表示网格的质量，设各个网格的质量分别为 m_1, m_2, \dots, m_n 。假设质心的坐标为 (\bar{x}, \bar{y}) ，那么质心公式为：

$$\begin{cases} \bar{x} = \frac{m_1 x_1 + m_2 x_2 + \dots + m_n x_n}{m_1 + m_2 + \dots + m_n} \\ \bar{y} = \frac{m_1 y_1 + m_2 y_2 + \dots + m_n y_n}{m_1 + m_2 + \dots + m_n} \end{cases} \quad (6)$$

特殊情况下，当各个网格质量相同时，质心为：

$$\begin{cases} \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \\ \bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n} \end{cases} \quad (7)$$

4.2 目标函数

为了寻找一种无人机在区域 E 的最优搜索位置使得每次搜索都能最大化的降低整个区域的不确定度。建立目标函数描述如下：

$$\begin{aligned} H_n &= \sum_{(i,j) \in E} \Delta \phi_n(i, j) \\ &= \sum_{(i,j) \in E} \max_{m \in \{1, 2, \dots, N\}} \{ \phi_n(i, j) - \beta (\| p_m - (i, j) \|) \phi_n(i, j) \} \\ &= \sum_{(i,j) \in E} \phi_n(i, j) - \min_{m \in \{1, 2, \dots, N\}} \{ \beta (\| p_m - (i, j) \|) \} \phi_n(i, j) \\ &= \sum_{m \in \{1, 2, \dots, N\}} \sum_{(i,j) \in V_m} \phi_n(i, j) (1 - \beta (\| p_m - (i, j) \|)) \end{aligned} \quad (8)$$

式(8)为区域离散化下的目标函数， n 表示第 n 次搜索， $\phi_n(i, j)$ 表示第 n 次搜索时网格 (i, j) 处的不确定度， $\phi_n(i, j)$ 表示在第 n 步搜索中，网格 (i, j) 处不确定度的降低值， $\phi_n(i, j)$ 表示无人机 m 的V图区域， $p_m \in E$ 表示第 n 架无人机的位置， $\| p_m - (i, j) \|^2$ 表示无人机位置到网格 (i, j) 的欧式距离。

4.3 最优解

为了求得目标函数的最优解，我们先讨论未进行区域栅格化的情况。

将目标函数采取下面形式：

$$H_n = \sum_m \int_{V_m} \phi_n(q) k e^{-\alpha r^2} dQ \quad (9)$$

式(9)中 $m \in \{1, 2, \dots, N\}$ 表示第 m 架无人机， Q 表示未栅格化之前的整个搜索区域， $q \in Q$ 表示区域中的任意一点， $\phi_n(q)$ 表示 q 处的环境不确定度，式(9)关于 p_m 的梯度可以表示为：

$$\begin{aligned} \frac{\partial H_n}{\partial p_m} &= \sum_m \int_{V_m} \phi_n(q) k e^{-\alpha (\| p_m - q \|^2)} (-2\alpha) (p_m - q) dQ \\ &= -2\alpha M_{V_m} (p_m - C_{V_m}) \end{aligned} \quad (10)$$

式(10)中， M_{V_m} 和 C_{V_m} 分别代表 V_m 的质量和质心， $\phi_n(q) k e^{-\alpha r^2}$ 表示传感器探测后的不确定度降低大小。因此，最优性的必要条件是：

$$p_m = C_{V_m} \quad (11)$$

根据式(11)可以看出，在未栅格化的情况下，能够对整个环境区域进行效率最高的搜索位置是每个V图的质心处，因此，在栅格化以后，同样是质心处能够最大的降低区域的不确定度。

4.4 搜索代价函数

无人机在向质心运动的过程中，每架无人机需要满足以下准则：

- (1) 考虑到任务的完成效率，无人机将飞向距离质心最近的地方；
- (2) 考虑到无人机的过载限制以及燃油经济性，无人机有沿当前航向直飞的趋势；
- (3) 考虑到无人机的作用区域，无人机的下一步航路点需在自身V图区域内。

基于以上的分析，可以建立无人机运动到质心的过程中各个飞行点的代价函数公式：

$$cost_k(i, j) = q_k(i, j) \times (w_1 \cdot \Delta \varphi(i, j) + w_2 \cdot D_k(i, j)) \quad (12)$$

式(12)中， $cost_k(i, j)$ 表示无人机 k 飞向 (i, j) 的代价； $q_k(i, j)$ 表示无人机的可行性，如果 $(i, j) \in V_k$ ，那么 $q_k(i, j) = 1$ ，反之，如果 $(i, j) \notin V_k$ ，那么 $q_k(i, j) = \infty$ ； w_1, w_2 分别表示航向变化以及距离的权重值且 $w_1 + w_2 = 1$ ； $\Delta \varphi(i, j)$ 表示无人机从当前位置运动到 (i, j) 时所

需要转过的角度，该值越大则表示飞向 (i, j) 所需的过载也就越大； $D_k(i, j)$ 表示 (i, j) 到质心的距离，无人机飞向距离质心更近的位置。

无人机在执行搜索任务的过程中，通过上述代价函数，运动到各自的质心位置执行搜索，能够最大的降低环境的不确定度。

4.5 不确定度更新

无人机对环境信息的获取主要包括三个方面：初始先验信息、探测信息和通信信息。

(1) 初始先验信息

初始先验信息每架无人机都是相同的，对环境的掌握程度都为零，环境的不确定度都为1。

(2) 探测信息

探测信息是无人机通过传感器的探测对环境信息有了一定了解。

通过对传感器的建模分析，我们知道探测效率为 $f(r) = ke^{-\alpha r^2}$, $k \in (0, 1)$ and $\alpha > 0$ ，也就是说，假设网格区域 (i, j) 时刻 k 不确定度为 $\phi_k(i, j)$ ，那么经过一次传感器的搜索以后，该网格区域环境不确定度降低了：

$$\Delta\phi_k(i, j) = \phi_k(i, j) \cdot ke^{-\alpha r^2} \quad (13)$$

$k+1$ 时刻网格 (i, j) 的不确定度为：

$$\phi_{k+1}(i, j) = \phi_k(i, j) - \phi_k(i, j) \cdot ke^{-\alpha r^2} \quad (14)$$

假设：

$$\beta = 1 - ke^{-\alpha r^2} \quad (15)$$

在没有传感器半径约束的情况下，无人机能探测整个搜索区域，由于是多机协同搜索，在同一位置有多架无人机搜索，但由于传感器探测效率随距离增加而降低，因此，仅考虑对该位置探测效率最高的无人机。

在搜索过程中，传感器在执行完第 $k+1$ 次搜索步骤后，环境的不确定度可以表示为：

$$\phi_{k+1}(i, j) = \phi_k(i, j) \min_m \{\beta(\|p_m - (i, j)\|)\} \quad (16)$$

其中， $\phi_k(q)$ 表示在第 k 次搜索时环境的不确定度。 p_m 表示第 m 架无人机的位置。 $\min_i \{\beta(\|p_i - (i, j)\|)\}$ 表

示对网格 (i, j) 搜索效率最高的传感器对网格不确定度的影响。

由于V图本身是一个根据距离划分的区域，也就是说，各无人机只搜索其V图范围内的区域。图8中，P1, P2, P3, P4, P5表示五架无人机，划分的V图区域表示各无人机传感器的有效探测范围。

如果任意无人机在其自身Voronoi单元内搜索，根据V图的势力范围特性，环境不确定度更新可由式(17)表示。

$$\phi_{k+1}(i, j) = \phi_k(i, j) \cdot \beta(\|p_m - (i, j)\|) \quad (17)$$

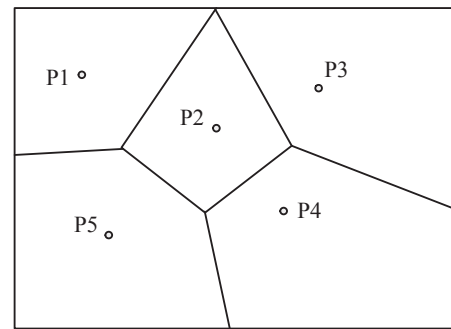


图8 无人机传感器探测距离

Fig.8 The detection ranges of UAVs sensors

(3) 通信信息

无人机传感器在探测过程中只降低自身V图区域内的环境不确定度，各无人机之间的通信使得无人机能掌握自身V图以外区域的环境信息。如图9所示，阴影部分表示无人机1掌握的区域信息，假设无人机2在无人机1通信半径内，也就是无人机1与无人机2之间能够进行通信，那么这两架无人机能够同时掌握各自所在的V图区域内的不确定度。

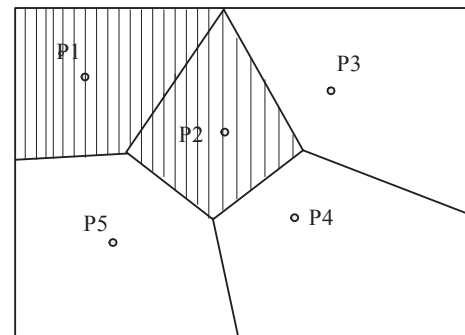


图9 两架无人机通过通信更新不确定度

Fig.9 Two UAVs update uncertainty by communication

因此，如果各架无人机之间不存在通信约束，即无人机能跟所有的无人机进行通信，那么无人机则能够掌握整个环境区域的不确定度信息。

4.6 传感器半径约束

在实际中，传感器不可能有无限的探测范围，假设传感器存在探测半径约束 R ，对距离大于 R 网格区域，传感器对其没有探测能力。如式(18)所示。

$$\tilde{f}(r) = \begin{cases} f(r), & \text{if } r < R \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

考虑下面的目标函数：

$$\tilde{H} = \sum_{m \in \{1, 2, \dots, N\}} \sum_{(V_m \cap \tilde{B}(p_m, R))} \phi_n(q) \cdot \tilde{f}(\|p_m - q\|) \quad (19)$$

根据 $\tilde{\beta}$ 得到的目标函数的梯度由式(10)表示：

$$\frac{\partial(\tilde{H})}{\partial p_m} = 2\tilde{M}_{(V_m \cap \tilde{B}(p_m, R))} (\tilde{C}_{(V_m \cap \tilde{B}(p_m, R))} - p_m) \quad (20)$$

其中，质量 \tilde{M} 和质心 \tilde{C} 是根据区域 $(V_m \cap \tilde{B}(p_m, R))$ 计算得到的，这种情况下的极值点为：

$$p_m = \tilde{C}_{(V_m \cap \tilde{B}(p_m, R))} \quad (21)$$

在上述的表达式中 $V_m \cap \tilde{B}(p_m, R)$ ，可以更为直观表示如图10所示。五边形表示无人机 m 所属

的Voronoi区域，四边形表示具有探测距离约束的传感器的最远探测边界，阴影部分则表示新的计算质量 \tilde{M} 和质心 \tilde{C} 的区域。因此在没有传感器距离约束的情况下，进行V图划分后的各个无人机所属的那一个区域是无人机能都探测到的区域，属于无人机的有效区，如图10(a)所示。而由于现在传感器有了距离约束，在探测范围外的区域实际上属于无人机的无效区，那些区域对降低区域的不确定性是没有作用的，如图10(b)所示。很多情况下，无人机的探测距离和Voronoi图并非属于包含或被包含的情况，而是互相重叠，这种情况下则需要计算出重叠区域，并根据重叠区域重新计算出质心，如图10(c)。

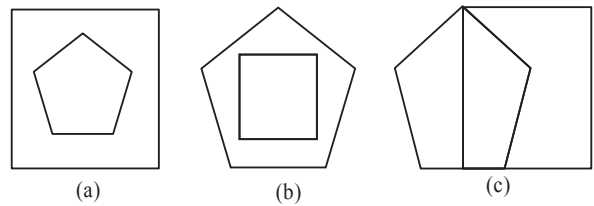


图10 传感器探测半径约束下的无人机搜索范围

Fig.10 The UAV search range under the constraint of sensor's detection radius

传感器半径约束下的DCPS策略流程图如图11所示。

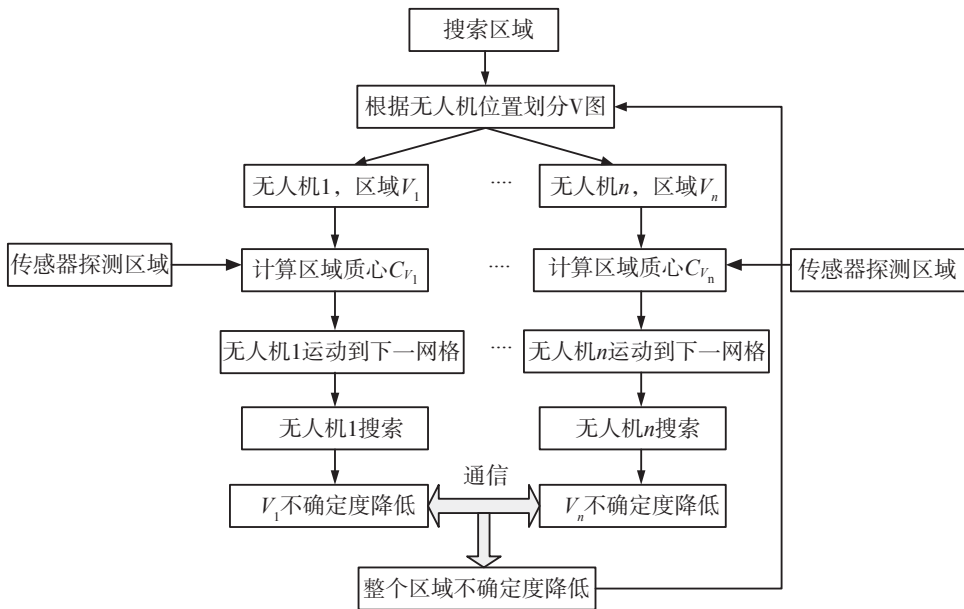


图11 基于DCPS策略的无人机协同区域搜索流程图

Fig.11 The flowchart of multiple UAVs cooperative area search based on DCPS

定理 本文提出的DCPS策略可以将平均不确定性降低为任意小值。

证明: 首先看任意点 $(i, j) \in E$ 的不确定度更新函数:

$$\begin{aligned}\phi_n(i, j) &= (1 - ke^{-\alpha r_m^i})\phi_{n-1}(i, j) \\ &= \gamma_{n-1}\phi_{n-1}(i, j)\end{aligned}\quad (22)$$

式(22)中, r_m 表示点 $(i, j) \in E$ 到无人机 m 的距离, $(i, j) \in V_m$ 对应于无人机 m 的Voronoi分区, $\gamma_{n-1} := (1 - ke^{-\alpha r_m^i})$ 。递归地应用上述更新规则, 可以得到:

$$\phi_n(i, j) = \gamma_{n-1}\gamma_{n-2}\dots\gamma_1\gamma_0\phi_0(i, j) \quad (23)$$

假设 $D(E) := \max_{p, q \in E} (\|p - q\|)$ 表示搜索区域 E 中两点的最远距离, 应该指出的是:

- (1) $0 \leq k < 1$;
- (2) $0 \leq r_i \leq D(Q)$, 且由于集合 E 有界, $D(E)$ 也是有界的。
- (3) $0 \leq \gamma_j \leq 1 - ke^{-\alpha\{D(E)\}^2} = l, j \in \mathbb{N}$ and $l < 1$ 。

考虑序列:

$$\Gamma = \{\Gamma_0, \Gamma_1, \Gamma_2, \dots\} \quad (24)$$

且:

$$\Gamma_n := \gamma_n\gamma_{n-1}\dots\gamma_1\gamma_0 \leq l^{n+1} \quad (25)$$

当取 $n \rightarrow \infty$ 时:

$$\lim_{n \rightarrow \infty} \Gamma_n \leq \lim_{n \rightarrow \infty} l^{n+1} = 0 \quad (26)$$

因此:

$$\lim_{n \rightarrow \infty} \phi_n(i, j) = \lim_{n \rightarrow \infty} \Gamma_{n-1}(\phi_0(i, j)) = 0 \quad (27)$$

取极限时, 每一点 $(i, j) \in E$ 的不确定度 ϕ 都会趋于零, 因此当 $n \rightarrow \infty$ 时, E 的平均不确定度也会趋于零。

5 多无人机协同搜索仿真

5.1 仿真参数设置

仿真设定了一块 40×40 的矩形网格区域作为多无人机协同的任务区域, 小网格长度为1。设定了五架携带不同传感器的无人机对该区域进行搜索。无

人机状态主要包括: 无人机位置信息 (x 坐标, y 坐标)、无人机方位信息 $\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$ 、无人机状态信息 (1 表示无人机状态良好可以正常工作, 0 表示无人机损毁不能正常工作) 以及所携带的传感器探测半径。五架无人机在地图中的位置分布如表1所示:

表1 无人机初始状态

Table 1 The initial states of UAVs

	x 坐标	y 坐标	偏航角	无人机状态
无人机 1	10.5	10.5	3	1
无人机 2	30.5	15.5	3	1
无人机 3	30.5	30.5	3	1
无人机 4	20.5	35.5	3	1
无人机 5	5.5	20.5	3	1

设无人机搜索之前区域的初始不确定度都为1, 如图12所示:

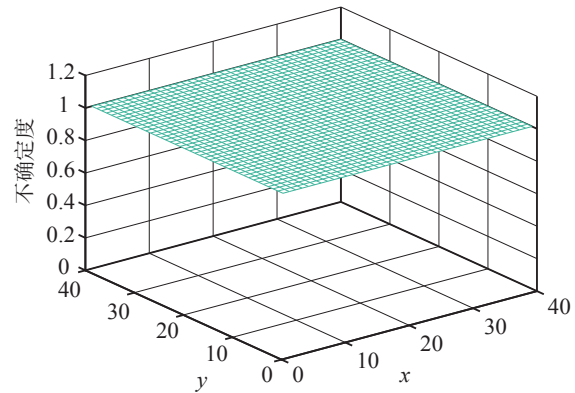


图12 区域不确定度初始状态图

Fig.12 The initial area uncertainty

随着无人机对区域不断搜索, 区域的不确定度不断降低, 无人机搜索某个网格区域时, 将该网格的不确定度根据离散传感器模型进行降低。设传感器模型为 $ke^{-\alpha r}$, $k = 0.8, i = 0.1$, 通过对 k, i 的确定, 也就确定了传感器的探测效能。

5.2 DCPS策略仿真

图13表示区域平均不确定度随搜索时间步长降低的变化曲线, 随着无人机对区域的搜索, 区

域不确定度逐渐降低最后趋近于0.05。区域的最终不确定度如图14所示。

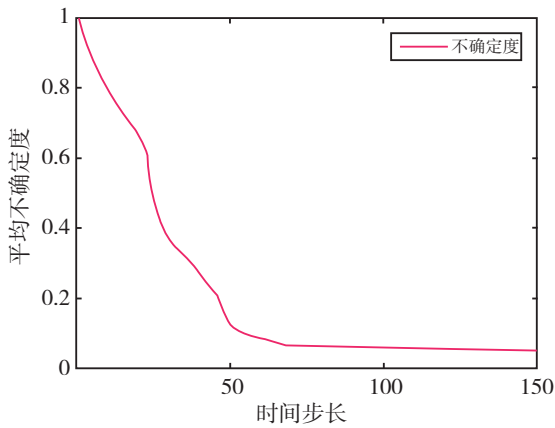


图13 区域不确定度随时间步长变化
Fig.13 The reduction of area uncertainty

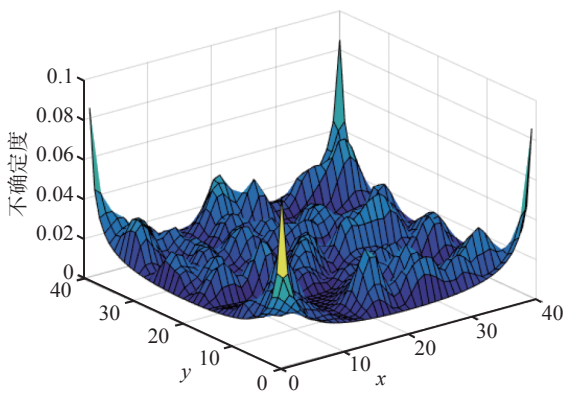
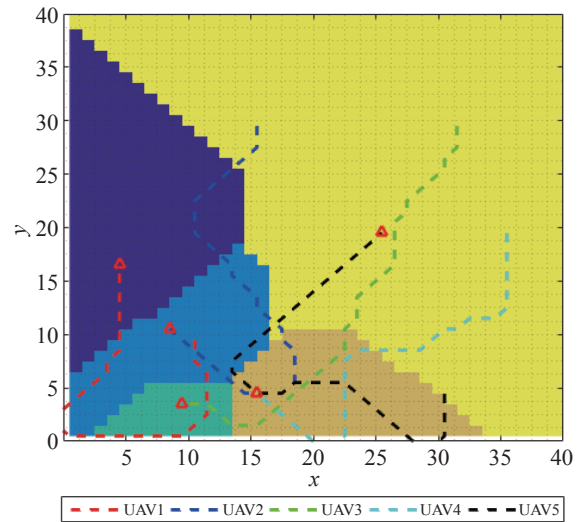


图14 区域最终不确定度
Fig.14 The final area uncertainty

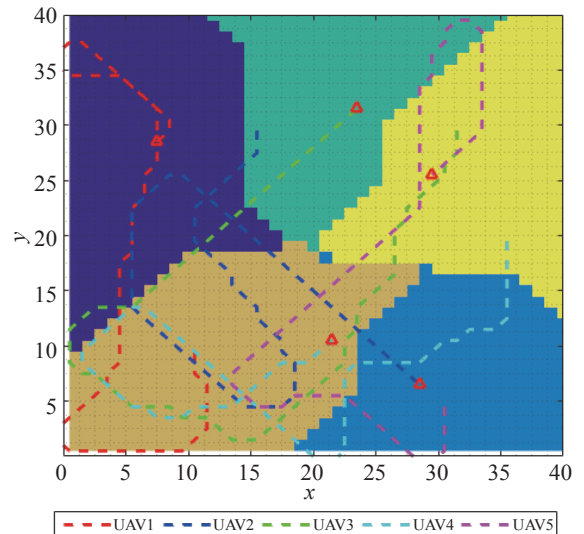
图15为五架无人机在没有传感器半径约束的情况下，搜索步长分别为35、70、100时DCPS策略的搜索航迹图，和根据当前时刻无人机位置划分的V图。图中红色三角形代表当前时刻无人机的位置，颜色相同的区域表示同一架无人机相对应的V图区域。

5.3 无人机新增或毁损对搜索效率的影响

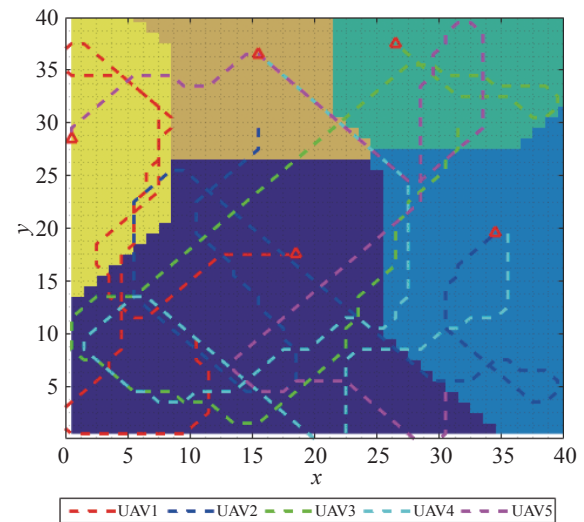
在没有传感器半径约束的情况下，假设在运行第50步时出现无人机毁损和无人机新增，图16分别表示出现毁损和新增时刻的V图重划分，以及将不确定度降低到0.05以下的整个搜索航迹。



(a) 35步搜索航路及V图划分



(b) 70步搜索航路及V图划分

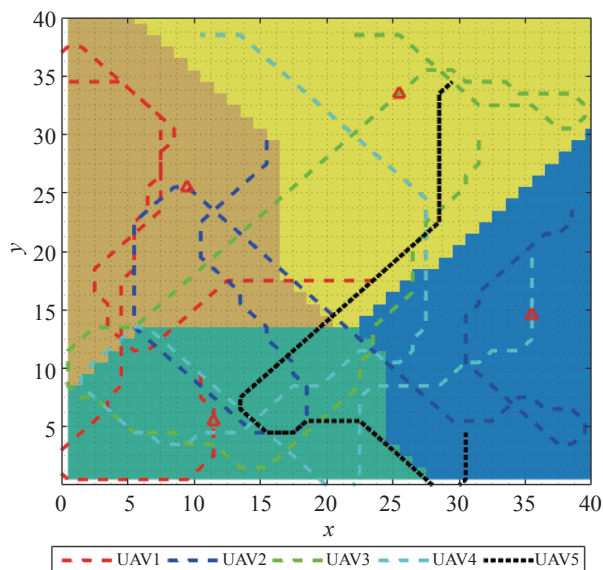


(c) 100步搜索航路及V图划分

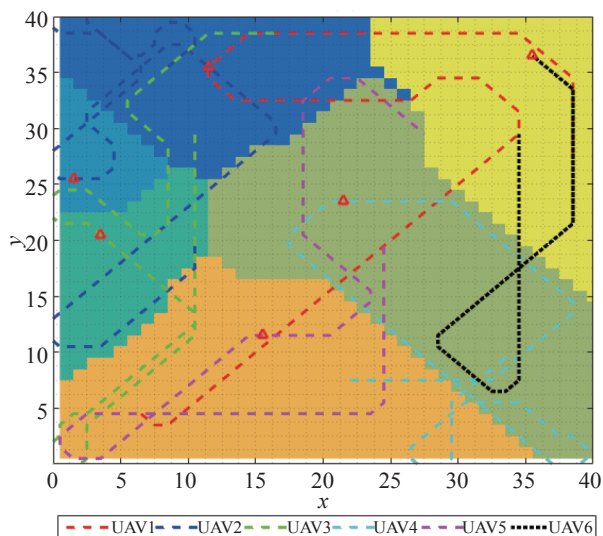
图15 无人机搜索航迹图

Fig.15 The search trajectories of UAVs

图16(a)中,红色三角形表示当前时刻无人机位置。黑色虚线表示在前50步无人机5的搜索路径,在第50步时,无人机5毁损,不再工作,只剩下四架无人机继续搜索直到不确定度降低到0.05以下。图16(b)中,黑色虚线表示在第50步时,新增的无人机6的搜索路径,在前50步,只有五架无人机在区域内搜索,50步时,新增无人机6加入与各架无人机一起协同搜索,直到区域不确定度降低到0.05以下。



(a) 无人机毁损瞬间的V图重划分



(b) 无人机新增瞬间的V图重划分

图16 无人机毁损和新增瞬间的V图重划分

Fig.16 Voronoi diagrams reconstruction

5.4 传感器探测半径约束对搜索效率的影响

根据本文所给出的传感器探测模型可得,在半径大于6时,传感器的探测效率已经非常低了,所以我们这里给出传感器半径小于6时对搜索策略的影响,并且对比无人机数量分别为5和10时传感器半径对搜索效率的影响程度,如图17所示。

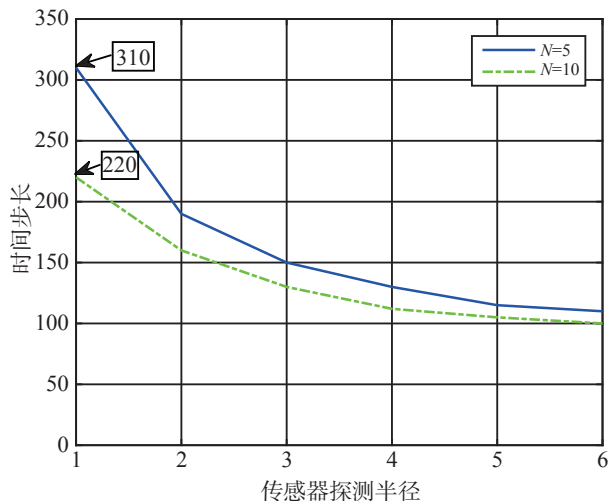


图17 将区域不确定度降低到0.05以下所需时间步长与传感器探测半径关系曲线

Fig.17 The relation between the time step and the sensor detection radius when area uncertainty is 0.05

由图17可以看出,不管是无人机数量为5还是10,随着传感器探测半径的增大,探测整个区域所需要的时间都是逐渐减少的,且在传感器探测半径为1时,无人机数量的增加能够很明显的降低搜索该区域的时间,但是随着传感器半径的逐渐增加,无人机数量对于该区域的搜索已经没有那么明显的优势,一方面由于搜索区域是固定的,无人机数量太多反而会导致重复性的搜索,另一方面,在无人机数量较少而传感器的探测范围又相对较大时就已经能够满足搜索效率了。

根据图17可以看出,在传感器半径为1时,要将区域不确定度降低到0.05以下,五架无人机需要搜索的时间步长为310步左右。图18为传感器探测半径为1时五架无人机搜索310步的路径图以及最后的V图划分,红色三角形表示当前时刻无人机位置。

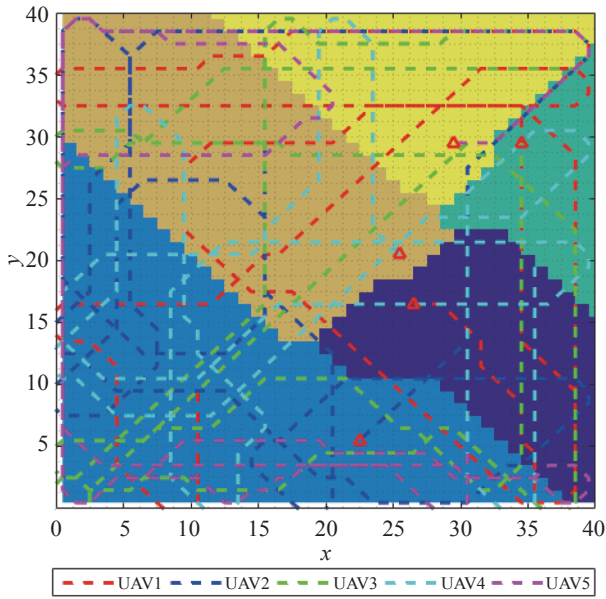


图18 传感器半径为1时五架无人机的搜索航迹图
Fig.18 The search trajectories when sensors radius is 1

5.5 DCPS策略与贪婪搜索、随机搜索策略对比

接下来，我们比较DCPS策略与贪婪搜索、随机搜索的性能。贪婪搜索和随机搜索是适用于任何情况下的搜索策略。

(1) 贪婪搜索

贪婪搜索与DCPS的主要区别在于：在控制无人机运动时，DCPS策略中，第 m 架无人机向 $V_m \cap \bar{B}(p_m, R)$ 的质心运动，在进行Voronoi分区的过程中，考虑了相邻无人机之间的协同。在贪婪搜索中，只让第 m 架无人机向着 $\bar{B}(p_m, R)$ 的质心运动。

考虑之前所给的不确定度更新函数 $\phi_{n+1}(q) = \phi_n(q) \min_i \{\beta(\|p_i - q\|)\}$ ，只有对某个网格搜索效率最高的无人机才会对该网格进行搜索，这也就导致了各架无人机都只会在其V图区域内执行搜索操作，这种贪婪搜索策略被称为Voronoi Greedy Search (VGS)策略，在该策略中，只有控制无人机的运动是贪婪的，而搜索是以协作的方式执行的。

不确定度更新规则采用如下形式：

$$\phi_{n+1}(q) = \phi_n(q) \prod_{\{i|p_i \in \bar{B}(q,R)\}} \beta(\|p_i - q\|) \quad (28)$$

(2) 随机搜索

随机搜索 (RS) 是最简单的搜索策略。假设无人机以恒定的速度移动，无人机的方向是随机产生的。

和DCPS策略一样，VGS策略和RS策略无人机在每个时间步长都执行搜索操作，仿真设定五架无人机协同搜索，传感器探测半径 $R = 6$ 搜索过程中，没有出现新增或者毁损。各搜索算法对整个区域不确定度的降低效率随着时间步长的变化关系如图19所示。

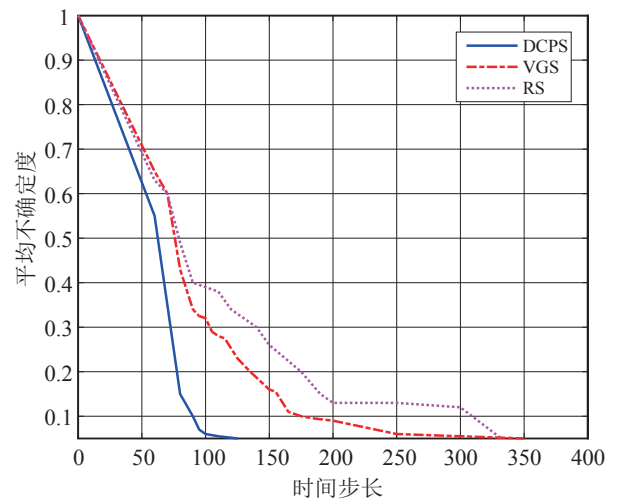


图19 算法效果对比图

Fig.19 The uncertainty comparison of different search strategies

可以看出，DCPS降低整个区域不确定度所需时间最短，值得注意的是，尽管所有这些策略在每个时间步中都执行搜索任务，但是由于各架无人机之间通过Voronoi分区进行协作，DCPS具有更好的搜索空间覆盖率，因此性能更好。在大多数情况下，对于相同数量的搜索，DCPS比其他策略（如VGS和RS）能更快降低区域平均不确定度。

6 结论

为了尽快降低不确定区域的环境不确定度，提出了一种基于V图质心的多无人机协同区域搜索算法，得到如下结论：

(1) 本文提出的DCPS策略相对于随机搜索和贪婪搜索能更快的将环境不确定度降低到指定范

围以下；且具有良好的鲁棒性，能够应对无人机的突发情况，包括无人机的新增和毁损；

(2) 在传感器半径的约束下，随着传感器探测半径不断减少，对整个区域的搜索时间会不断增加，无人机的搜索路径也就越密集，且在固定的搜索区域内，无人机的增多会在一定程度上减少搜索时间，但由于区域是固定的，无人机数量太多反而会导致重复性的搜索，影响搜索效率。

本文在多机协同搜索过程中，没有考虑各架无人机之间的通信约束，区域搜索指标仅限于不确定度的降低，还需要进一步考虑通信约束，以及更多的搜索指标，对多无人机协同区域搜索算法进一步完善。

参 考 文 献

- [1] 巴恩哈特. 无人机系统导论 [M]. 北京: 国防工业出版社, 2014.
- [2] Chandler P R, Swaroop D, Howlett J K, et al. Complexity in UAV cooperative control [C]. American Control Conference. 2002.
- [3] Schumacher C, Chandler P R, Rasmussen S R. Task allocation for wide area search munitions via iterative network flow [C]. American Control Conference. 2003.
- [4] Lum C W, Rysdyk R T, Pongpunwattana A. Occupancy based map searching using heterogeneous teams of autonomous vehicles [J]. Aiaa Journal, 2006.
- [5] Lum C, Vagners J. A Modular algorithm for exhaustive map searching using occupancy based maps [J]. Aiaa Journal, 2013.
- [6] Hussein I I, Wang Y, Erwin R S. Building a probabilistic occupancy map for space situational awareness [C]. Aiaa Guidance, Navigation & Control Conference. 2006.
- [7] Khan A, Yanmaz E, Rinner B. Information merging in multi-UAV cooperative search [C]. IEEE International Conference on Robotics & Automation. 2014.
- [8] 田菁, 陈岩, 沈林成. 不确定环境中多无人机协同搜索算法 [J]. 电子与信息学报, 2007, 29 (10): 2325-2328.
- [9] 彭辉, 苏菲, 沈林成. 用于多无人机广域目标搜索的扩展搜索图方法 [J]. 系统工程与电子技术, 2010, 32 (4): 795-798.
- [10] 刘重, 高晓光, 符小卫. 带信息素回访机制的多无人机分布式协同目标搜索 [J]. 系统工程与电子技术, 2017, 39 (9): 1998-2011.
- [11] 符小卫, 魏广伟, 高晓光. 不确定环境下多无人机协同区域搜索算法 [J]. 系统工程与电子技术, 2016, 38 (4): 821-827.
- [12] Yang Y, Minai A A, Polycarpou M M. Decentralized cooperative search by networked UAVs in an uncertain environment [C]. American Control Conference, 2005.
- [13] Cortes J, Martinez S, Karatas T, et al. Coverage control for mobile sensing networks [C]. IEEE International Conference on Robotics & Automation. 2003.
- [14] Hu J, Xie L, Lum K Y, et al. Multiagent information fusion and cooperative control in target search [J]. IEEE Transactions on Control Systems Technology, 2013, 21 (4): 1223-1235.
- [15] 于驷男. 多无人机协同搜索区域分割与覆盖 [J]. 北京航空航天大学学报: 自然版, 2015, 41 (1): 167-173.
- [16] 沈东, 魏瑞轩, 祁晓明. 基于MTPM和DPM的多无人机协同广域目标搜索滚动时域决策 [J]. 自动化学报, 2014, 40 (7): 1391-1403.
- [17] Di B, Zhou R, Duan H. Potential field based receding horizon motion planning for centrality-aware multiple UAV cooperative surveillance [J]. Aerospace Science and Technology, 2015, 46: 386-397.
- [18] 彭辉, 沈林成, 朱华勇. 基于分布式模型预测控制的多UAV协同区域搜索 [J]. 航空学报, 2010, 31 (3): 593-601.
- [19] 黄长强. 多无人机协同搜索随机目标决策 [J]. 控制与决策, 2013, 28 (5): 711-715.
- [20] Guruprasad K R, Ghose D. Performance of a class of multi-robot deploy and search strategies based on centroidal Voronoi configurations [J]. International Journal of Systems Science, 2013, 44 (4): 680-699.
- [21] 李佳田, 罗富丽, 余莉, 等. 梯度Voronoi图及其构建算法 [J]. 武汉大学学报 (信息科学版), 2016, 41 (2): 163-170.

作者简介:

朱利 (1995-), 女, 硕士, 工程师, 主要研究方向为无人机协同控制。

符小卫 (1976-), 男, 博士, 副教授, 主要研究方向为无人机协同控制与效能评估。本文通信作者。

无人潜航器仿生推进技术专利分析

张鸣宇, 李 峰, 周 鹏

(北京海鹰科技情报研究所, 北京 100074)

摘 要: 针对无人潜航器仿生推进技术的全球专利, 结合科技文献等其他信息, 分析主要国家该领域专利的申请态势、技术领域、专利布局、研发机构等内容, 从专利的视角分析了无人潜航器仿生推进技术的发展, 以期为相关机构提供参考。结果显示该领域处于快速发展期, 专利数量增长迅速, 中国申请人的专利数量较多, 但专利质量同美国、日本存在差距。且美、日等国已在中国进行了专利布局, 而中国申请人缺少国外布局。高校与科研机构应重视柔性驱动技术与智能材料应用的研究, 提前进行专利布局, 扩大无人潜航器仿生推进技术研发的产学研合作, 弥补同发达国家的产业化差距。

关键词: 无人潜航器; 仿生推进; 柔性驱动; 机器鱼; 仿鱼机器人; 专利分析; 专利申请; 专利布局
中图分类号: U674.941 **文献标识码:** A **文章编号:** 2096-5915 (2019) 02-0430-08

Patent Analysis of Bionic Propulsion Technology of Unmanned Underwater Vehicle

ZHANG Mingyu, LI Feng, ZHOU Peng

(Beijing HIWING Scientific and Technological Information Institute, Beijing 100074, China)

Abstract: Aiming at the global patents of bionic propulsion technology for unmanned underwater vehicle, combining with other information such as scientific and technological literature. This paper analyses the patent application situation, technology field, patent layout, research and development institutions of major countries in this field. Meanwhile, it analyses the development of bionic propulsion technology for unmanned underwater vehicle from the perspective of patent, with a view to providing reference for relevant institutions. The results show that this field is in a period of rapid development, the number of patents is growing rapidly, the number of patents from Chinese applicants is larger, but the quality of patents is far from that of the United States and Japan. The United States, Japan and other countries have patent layout in China, in contrast, Chinese applicants neglect foreign layout. Universities and research institutes should attach importance to the research of flexible driving and intelligent materials, carry out patent layout ahead of time. It is vital to expand cooperation between industry, University and Research on this field, then eliminate the industrialization gap from China and developed countries.

Keywords: Unmanned Underwater Vehicle; Bionic Propulsion; Flexible Drive; Robotic Fish; Fishlike Robot; Patent Analysis; Patent Application; Patent Layout

收稿日期: 2019-01-11; 修回日期: 2019-02-20

1 引言

仿生推进无人潜航器，是以水生生物为模仿对象，综合了流体动力学、仿生学、导航控制、机械结构、电子、拓扑学和图像处理等多个学科，实现游动方式仿生的无人潜航器。

相比常规无人潜航器采用的螺旋桨或叶轮推进，仿海洋生物游动推进的潜航器，具有明显优势：第一，推进效率较高、能耗较低，克服了螺旋桨推进在产生推力的同时还产生的无用扭矩；第二，噪音较小，不易被发现；第三，结构尺寸和重量小，机动性强。因此，在海洋资源勘探、考古、污染源追踪等民用领域和水下攻击、侦察和扫雷等军用领域均可发挥重要的作用^[1-3]。

鉴于无人潜航器仿生推进技术的重要作用，国内外相关机构开展了多年研究，做了众多仿蝠鲼、仿乌贼、仿龙虾、仿海龟等不同潜航器样机试验，同时也在全球范围内布局了部分专利。为全面掌握无人潜航器仿生推进技术的发展趋势和专利布局情况，本文基于Patsnap专利数据库，开展了全面检索和逐一筛选，共筛选出625项直接相关专利。

2 专利申请趋势

无人潜航器仿生推进技术的专利申请状况可以分为三个阶段：第一阶段是20世纪90年代中期以前，主要侧重于仿生推进基础理论方面的研究和纯机械式游动设计；第二阶段是90年代中后期至2010年，伴随着自动化、电机驱动、新型材料技术的进步，人们开始了真正意义上的仿生推进无人潜航器的研制，专利申请缓慢增长，年申请量大多不超过20件；第三阶段是2011年至今，专利申请呈迅速增长态势，特别是由2014年的58件、2015年的59件，到2016年暴发性增长为123件，如图1所示。由于近两年的部分专利申请尚未公开，统计不完全，实际申请量更多。

无人潜航器仿生推进最早可追溯于1933年，由美国专利商标局受理的专利US1928418，申请人为Charles Garland，但受限于当时科技发展，技术方案仅用纯机械传动，未采用电机或柔性致动等现代驱动控制手段，仅用于浅海短距离航行。

第一件真正意义的仿生推进无人潜航器专利US5740750由麻省理工学院水下机器人研究所

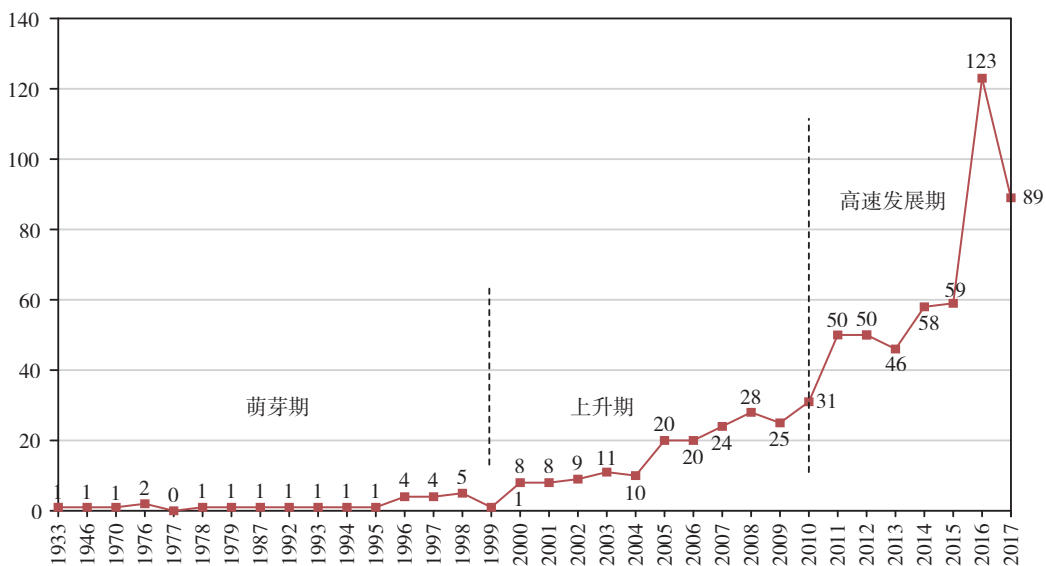


图1 全球专利申请趋势图

Fig.1 Global patent application trend chart

表1 专利申请情况表

Table 1 Table of patent application situation

年代	发展阶段	代表性技术	重要申请人	代表性专利
20 世纪 90 年代 中期以前	萌芽期	仿生推进机理及基础理论； 机械手工驱动	Charles Garland	US1928418
			麻省理工学院 三菱重工	US5401196 JP1978000594A
90 年代中后期至 2010 年	新技术推动 上升期	仿生推进模式设计； 电机驱动轮系机构与传动杆连杆操纵； 液压带动齿轮、曲柄滑块等刚性机构驱动； PVC 塑料制作仿生胸鳍； 微机中央控制	麻省理工学院	US5740750
			德国 FESTO 公司	DE202006006115U1
			三菱重工	DE102008019454A1
			日本宝生公司	JP2002137791A
			中科院自动化所	JP2002136776A
			哈尔滨工程大学	JP1999152085A
2011 年至今	高速发展期	应用智能材料的柔性驱动； 形状记忆合金、电致动聚合物、 压电材料、离子交换聚合金属材料制作 胸鳍； 气动肌腱、人工肌肉驱动； 中枢模式发生器控制； 模糊控制	北京航空航天大学	JP2001070667A
			上海劲钻实业公司	CN1631620A
			北京大学	CN101143618A
			麻省理工学院	US8733099
			波士顿工程公司	US20160257386A1
			美国东北大学	US20140109821
			美国海军	US20100262283A1
			佐治亚技术研究公司	US20130305978A1
			美国查尔斯塔克德雷珀实验室	US20150239539A1
			美国 Irobot 公司	US7007626
三菱重工	JP6001085B2			
哈尔滨工业大学	CN105292425B			
中科院自动化所	CN104627342B			
哈尔滨工程大学	CN1785747A			
北京航空航天大学	CN103085955B			
国防科技大学	CN103625625B			
杭州畅动智能公司	CN106184674A			

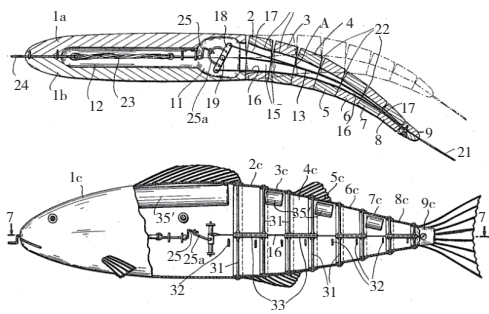


图2 第一件仿生机器鱼专利US1928418附图
Fig.2 First bionic robotic fish patent



图3 仿金枪鱼潜航器实物图
Fig.3 Real object of biomimetic tuna unmanned underwater vehicle

Triantafyllou Michael团队于1996年申请。通过多年仿生机理基础研究和金枪鱼的游动实验，成功研制了世界上第一条真正意义上的仿生机器金枪鱼 (RoboTuna) 潜航器，长约1.2米，由2843个零件组成，具有高级推进系统，包括关节式铝合金脊柱、真空聚苯乙烯肋骨、网状泡沫组织，并用聚氨酯甲酸酯弹性纤维表皮包裹，装有多台2马力的无刷直流伺服电动机（实际上只使用每台电机额定功率的十分之一）、精密轴承及电路等。机器鱼在多处处理器控制下，通过摆动躯体和尾鳍，能像真鱼一样游动，速度可达2m/s^[4]。

3 专利技术领域分布

水游生物推进模式和潜航器仿生推进机理是研究的理论基础，包括对摆动波动、划动、水翼

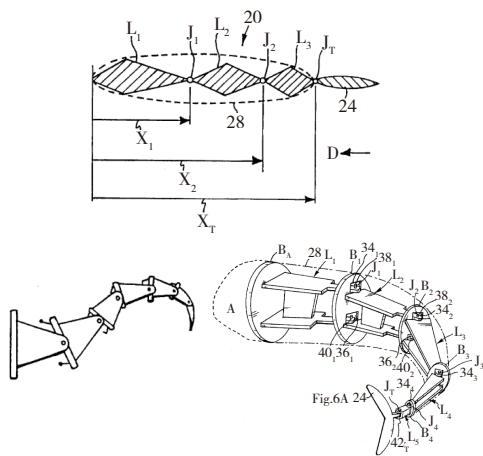


图4 专利US5740750附图

Fig.4 Attached figure of patent US5740750

推进、射流反冲等运动形式的数学描述和运动机理解析计算^[5-6]。在仿生推进机理研究的基础上,转化为工程实际过程中还涉及仿生推进模式设计、结构与机构设计、仿生推进驱动技术、控制及建模技术,专利检索和逐一判读筛选出的直接相关专利也几乎集中在这些领域。但同时,仿生机器鱼、仿生乌贼等现实产品的研制也依托了潜艇、鱼雷及常规潜航器等多数水下装备较为通用的技术,包括能源供给、浮力材料、导航通信技术等,甚至未来向金枪鱼高速游动技术的发展可能会涉及的超空泡技术等,但目前鲜有运用于仿生推进的专利申请,故不作为专利分析重点。

根据仿生推进技术公开专利的实际情况,可分为仿生推进总体设计、结构与执行机构设计、控制与建模、仿生推进机理、仿生试验测试五大

领域。全球仿生推进技术相关专利的技术领域分布如图5所示,涉及仿生推进总体设计的专利占73%,涉及结构与执行机构设计的专利占12%,建模与控制技术专利占11%,仿生推进基础理论的专利占4%。在仿生推进总体设计中,采用波动摆动推进技术方案的专利占绝大部分,喷射推进有69件,水翼推进和划动推进共22件。

无人潜航器采用仿生推进方式在水下游动时,其展向截面形状为多项式曲线,通常有三种方法拟合曲线运动规律:刚性机构、柔性机构、智能材料。目前刚性机构的专利数量居多,无人潜航器仿生推进采用电机作为驱动源,依靠机械传动装置驱动推进器运动,因而存在整体结构复杂、传动装置易出现机械故障、难以实现大潜深等问题,近几年来出现的智能材料与柔性驱动仿生推进的专利在逐步克服原有缺陷。

以仿生推进无人潜航器的波动鳍为例^[7],目前大多由有滑动关节的刚性段构件连接组成,采用液压、电磁或电机驱动,相较于运用柔性机构或智能材料的柔性驱动增加了复杂性和潜在故障点,且滑动接头需要适应不同距离的驱动元件,这是仿生推进系统中的一个常见问题。因此,无人潜航器仿生推进越来越离不开用于其航行体的柔性变形与运动的驱动器技术,要求其驱动变形大、响应快速、控制简单、能耗低且重量轻。随着仿生无人潜航器的微型化发展,智能材料的优势愈发明^[8]。目前专利申请较多并应用在仿生

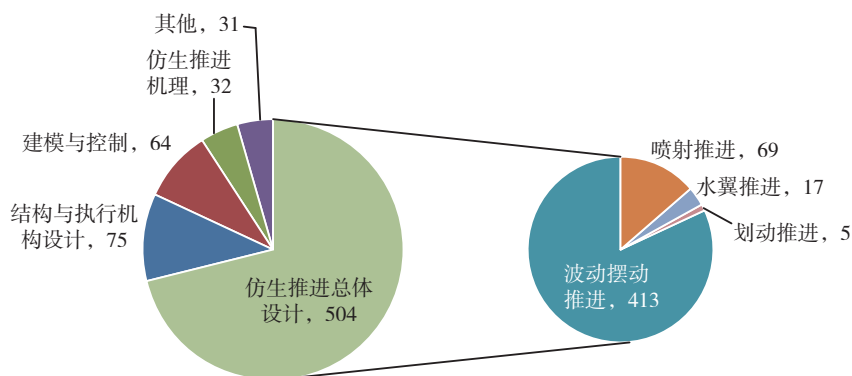


图5 全球专利技术领域分布图

Fig.5 Distribution map of global patent technology field

无人潜航器上的智能材料主要包括形状记忆合金、电致动聚合物、压电材料等。与传电机、液压带动齿轮、曲柄滑块等刚性驱动不同，柔性驱动摒弃了复杂的机械传动装置，通常以功能聚合物或者聚合物基复合材料为基础实现复杂终端运动。

4 申请人国别分布

专利申请人的国别构成如图6所示，其中专利申请量的国别排名依次是：中国(430)、美国(81)、日本(25)、韩国(24)和德国(14)。上述排名与无人潜航器仿生推进技术的研究热点区域基本吻合，美国和中国是目前世界上这一领域研究和专利申请的热点区域。但美国、日本、德国的申请人专利布局区域更广，如德国FESTO公司的同一项专利在美、日、中、欧和世界知识产权组织均有布局。而国内申请人的专利几乎都只在中国布局，截至目前，国内只有江苏科技大学和广东尚鱼电子有国际专利申请，前者关于一种仿生胸鳍推进装置基于行星齿轮机构的专利US20170152011A1在美国布局申请，广东尚鱼电子的创始人卢小平关于一种仿生鱼的驱动与控制方法及仿生鱼的专利CN102556310B在近40个国家/地区布局。

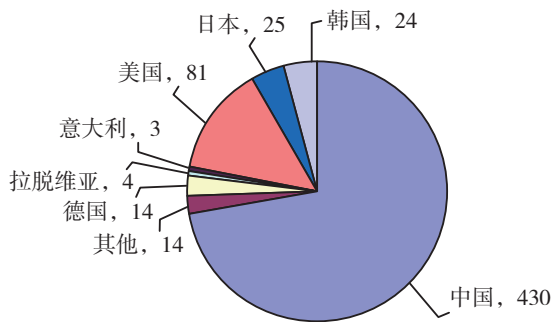


图6 专利申请人国别分布图
Fig.6 Country distribution of patent applicants

5 国内相关研究分析

国内无人潜航器仿生推进的相关论文最早出

现于1995年，是东南大学陆金伟发表的介绍麻省理工学院研究成果的科普文章。国内论文发表数量从2002年开始大幅增长，2013年达到70篇。依托863计划机器人主题技术、国家自然科学基金、国防基础科研等大量项目的支撑，目前处于平稳增长期。

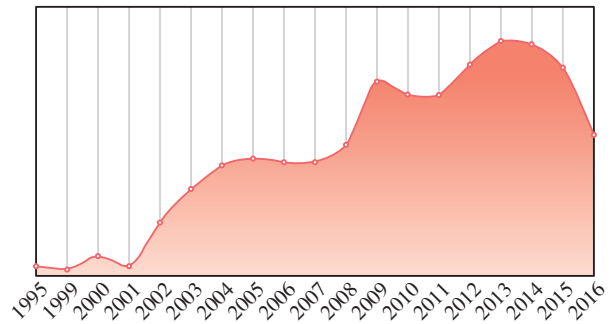


图7 中国论文发表量趋势
Fig.7 Chinese publication trend

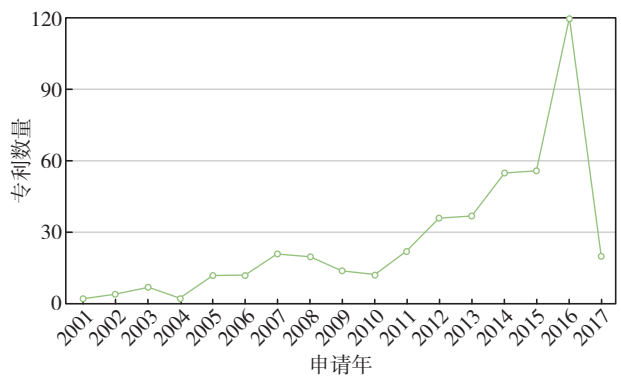


图8 中国专利申请趋势图
Fig.8 Chinese patent application trend

2001年，我国出现了第一件仿生潜航器领域的专利申请：新型单关节仿生机器鱼，由北京航空航天大学梁建宏老师申请。中国无人潜航器仿生推进技术专利申请量自2001年后逐年上升。国内该领域专利申请经历了低速增长和高速增长两个阶段。2002年至2013年，专利总体量较少，但数量平稳上升，最多时，年申请32项专利。但是在2013年后专利申请数量直线上升，2016年申请120项专利。一方面由于国内高校、中科院等科研机构获批的自然基金等项目较多，国有船舶、航天、电子、军工等集团在仿生水下装备领域的介

人，另一方面国内有多家专注于消费级仿生机器人产品的初创企业成立。例如，杭州畅动智能科技有限公司2016年4月注册成立后，立即申请了44项专利，博雅工道机器人公司的专利也大多为2016年申请。我国相关专利由之前全部为高校院所申请向民营初创企业大面积申请转变，且初创企业同高校院所合作紧密。

6 重要申请人分析

麻省理工学院水下机器人研究所M.Trantafyllou团队在无人潜航器仿生推进研究方面首屈一指，自20世纪80年代就有多篇海洋生物游动机理方面的论文，在90年代初申请的专利US5401196中就证明水翼在一定的频率和振幅等参数组合下推进效率可接近90%，采用水翼法推进双臂协同克服了鱼尾单鳍摆动过程中的侧向力和重浮心偏移问题。1995年成功研制出世界上第一个真正意

义的仿生推进无人潜航器，并在20多年来不断改进升级，申请的专利数量不算多，但引用量很大，每篇专利涉及的内容很丰富。

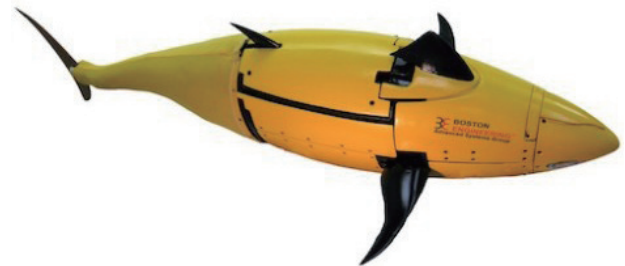


图9 波士顿动力公司仿生鱼
Fig.9 Aquatic vehicle of boston

耐顿公司一直在生物力学界的研究浪潮中引领风骚，从自然界中撷取灵感设计水下机器人，专利US7007626公开了其四鳍驱动的仿生水下航行器Pilot Fish的技术方案。2008年耐顿公司被iRobot公司收购，成为iRobot公司海洋系统分部的初创和

表2 国际主要申请人

Table 2 Main international applicants

序号	研究机构	典型专利	数量
1	麻省理工学院	US7865268 Mechanical fish robot exploiting vibration modes for locomotion US20160257386A1 Underwater Vehicle Design And Control Methods US9032900 Marine Vehicle Systems And Methods	6
2	佐治亚理工大学	US61638138P0 Robotic System for Ocean Exploration and Monitoring: A Buoyancy-Propelled, Autonomous Platform with Ray-like Mobility	9
3	波士顿动力公司	US9090320 Aquatic Vehicle US20140109821A1 Aquatic Vehicle US61716236P0 Underwater Vehicle	5
4	iRobot 公司	US7495999 Underwater guidance systems, unmanned underwater vehicles and methods US7007626 Amphibious robot devices	5
5	美国海军	US6835108 Oscillating appendage for fin propulsion US9045211 采用 ABS (丙烯腈族) 材料柔性鳍骨构成的潜航器 WO2015051383A3 采用隆头鱼胸鳍柔性鳍条推进模式的仿生装置	5
6	日本东海大学	JP1997301274 Aunderwater Sailing Device JP2001191985A 水中航走装置 JP4399526B2 水中航走装置	3
7	三菱重工	JP2002253873A 魚ロボット JP2002136776A 魚ロボット 及び水中通信装置 JP2003231495A 振動翼付き水中航走体	14
8	日本宝生公司	US20050235899A1 Fish-Shaped Underwater Navigating Body, Control System Thereof, And Aquarium JP2001070667A OPTICAL UNDERWATER TOY	9

表3 中国主要申请人
Table 3 Main Chinese applicants

机构	专利数量	主要领域或代表性成果	重点专利
中国科学院自动化研究所	46	建模控制、主动视觉跟踪方法、两关节仿生机器鱼、四关节仿生海豚	CN103332282B CN106873632A CN104298996A
杭州畅动智能科技有限公司	44	玩具机器鱼	CN106143851A
哈尔滨工程大学	41	总体设计、智能材料、建模控制、仿生机理、弹性组件提高驱动效率的实验平台，基于人工肌肉的仿生机器章鱼，IPMC 驱动的自主仿生机器鱼	CN107097921A CN107031806A
北京航空航天大学	24	总体设计、智能材料、建模控制、机器鳗鱼、机器牛鼻鲛、机器海豚	CN105022269B CN104477357A
哈尔滨工业大学	23	总体设计、智能材料、结构与机构、仿鱼鳍推进机理研究，建立了利用弹性组件提高驱动效率的实验平台	CN104787283B
北京大学	16	水下感知、总体设计、多仿生机器鱼协同控制、双自由度仿生尾鳍推进机构	CN106043644A
博雅工道（北京）机器人科技有限公司	16	消费级水下摄像设备	CN205801473U CN205801479U
南京航空航天大学	10	总体设计、柔性驱动、仿乌贼水下探测器、离子型人工肌肉驱动的机器鱼	CN102514697B
浙江大学	10	总体设计、建模控制、空潜两栖机器人、用于鱼群养殖的自主追赶鱼群机器鱼	CN205770079U CN105151301A
中国科学技术大学	9	建模控制、仿生机理、三维波动板理论	CN104875868A

主体力量，专注于用于海洋勘探和军事侦察的自主水下潜航器研发。

中国科学院自动化所先进机器人与制造系统团队、水下装备研究团队自 2001 年以来聚焦水下仿生领域，从最初跟踪模仿到如今已具有大量自主知识产权。哈尔滨工程大学和江苏科技大学依托海军和海洋装备的传统优势拓展水下仿生领域。北京航空航天大学依托机器人研究所的专业积累，公开了智能材料、建模控制研究和机器鳗鱼、机器牛鼻鲛、机器海豚等样机试验产品的专利。中国科学技术大学自 20 世纪 80 年代就有多篇海洋生物游动机理、数学计算的论文，后依托信息控制的学科优势开展仿生潜航器样机的研制，专利数量不多，但系统性、延续性较强。

尚鱼电子科技有限公司和博雅工道机器人科技有限公司均为国内初创企业。尚鱼电子是一支小而强的团队，专注于消费级但具有较高技术含量的仿生机器鱼，通过申请专利，形成保护壁

垒，再以知识产权与一些业内具有一定规模的公司合作，迅速将产品推向市场，通过收取专利授权费实现回报。例如研发出的微型机器鱼授权给世界最大的玩具商之一 Zuru 公司，2014 年销售超过 1000 万件，产品销售到 80 多个国家，2015 年还取得了迪斯尼的授权，2016 年产品与电影同步开售。相应的，尚鱼电子也在全球 40 多个国家布局了大量专利。博雅工道成立于 2015 年，由北京大学工学院博士团队创立，致力于推动水下潜航器从工业级走向消费级，为用户提供水下运动搭载平台以及水下摄影应用，目前拥有独立知识产权机器鱼产品两款，专利 16 项。

7 结束语

国内研究机构近两年的专利申请数量激增，但同美、日等国相比，专利几乎仅在中国本土进行布局。从专利撰写质量、创新点保护、重点

专利对比来看,中国专利的权利要求数、引用数量、同族数量明显少于美国和日本的专利,例如查尔斯塔克德雷珀实验室申请的远海自主潜航器专利US6138604引用达51次,权利要求数24个,而国内引用次数最多的专利CN101913419A被引用11次,权利要求数7个,独权2301个字,其余从权平均10多个字。建议我国相关研究机构提高无人潜航器仿生推进技术研发过程中的专利撰写质量,更好的运用于武器装备的立项和未来的国际化竞争。

柔性变形与运动的驱动技术对于提高水下航行体大变形驱动、快速响应、灵敏控制、降低能耗和减轻重量仍是关键问题。各国在柔性驱动技术的专利申请量逐渐增多,但柔性驱动器、记忆合金等柔性材料应用于仿生推进领域的专利申请量还很少,柔性驱动未来的工程化应用仍将持续。举例来说,关于复杂曲面驱动器的研究尚处于探索阶段,现有的基于IPMC(离子交换聚合金属材料)型柔性驱动器中,有的仅利用了IPMC的单自由度弯曲变形,有的虽然在研究多自由度弯曲变形但却未用到传感功能,而对于兼备传感功能和致动功能的阵列式柔性驱动器的分布设计及研究几乎空白,专利申请尚未发现。建议国内研发机构可以寻求电致动聚合物、压电材料、记忆合金、离子交换聚合金属等智能材料应用于水下航行体仿生推进设计中,解决无人潜航器仿生推进技术发展的瓶颈。

国内在无人潜航器领域的专利申请主体主要是高校和科研院所,且大多处于工程化样机研发阶段。而美国和日本创新主体大多是企业,与大学通过产学研结合实现了商业化运营。例如擅长仿生警用无人潜航器和两栖仿生潜航器的美国

iRobot公司于1990年由美国麻省理工学院教授罗德尼布鲁克斯及其同事创立。建议国内高校加强潜航器技术孵化,同创新型微小企业紧密合作,弥补同发达国家存在的产业化差距。

参 考 文 献

- [1] 刘军考. 仿鱼水下推进器理论与实验研究[D]. 哈尔滨: 哈尔滨工业大学. 2001.
- [2] 喻俊志, 陈尔奎, 王硕, 等. 仿生机器鱼研究的进展与分析[J]. 控制理论与应用, 2003, 20(4): 485-491.
- [3] 王松, 王田苗, 梁建宏, 等. 机器鱼水下考古实验研究[J]. 机器人, 2005, 27(2): 147-151.
- [4] Massachusetts Institute of Technology. Method And Apparatus For Reducing Drag On A Moving Body: US5740750[P]. 1996-05-28.
- [5] 刘晓白. 仿海龟柔性水翼推进技术研究[D]. 哈尔滨: 哈尔滨工程大学, 2012.
- [6] 詹榆菲. 仿生鱼类推进系统的理论建模及分析[J]. 中国科技纵横, 2017(5): 201-202.
- [7] 闫勇程, 王扬威, 兰博文, 等. 基于环形长鳍波动推进的仿生水下机器人设计[J]. 机械制造与自动化, 2018(1): 166-169.
- [8] 王扬威, 王振龙, 李健, 等. 形状记忆合金驱动仿生蝠鲼机器鱼的设计[J]. 机器人. 2010(2): 256-261.

作者简介:

张鸣宇(1990-), 男, 硕士, 工程师, 主要研究方向为专利分析与科技咨询。

李 峰(1986-), 男, 硕士, 工程师, 主要研究方向为知识产权管理与专利分析。

周 鹏(1985-), 男, 硕士, 工程师, 主要研究方向为专利数据分析与信息系统开发。

边缘计算的发展趋势综述

葛悦涛, 尹晓桐

(北京海鹰科技情报研究所, 北京 100074)

摘要: 边缘计算是继分布式计算、网格计算、云计算之后, 在网络边缘执行计算的一种新型计算范式, 因其能够就近提供智能互联服务, 所以可以满足各行业在数字化变革过程中的关键需求。主要对近期边缘计算的相关研究与成果进行了综合评述, 并对未来发展趋势进行了展望。首先, 概述了边缘计算的定义, 并讨论这一技术在数据安全、交互延迟、带宽成本等方面的优势, 边缘计算的飞速发展与多项先进技术及应用存在协同共进、相互促进的关系。最后从三个角度分析边缘计算发展趋势: 边缘计算与云计算协同发展, 成为实现智能物联网的关键; 发展边缘计算需要人工智能芯片铺路; 5G商用提速为边缘计算发展提供新的机遇。

关键词: 边缘计算; 云计算; 人工智能; 人工智能芯片; 5G; 物联网

中图分类号: TP391 **文献标识码:** A **文章编号:** 2096-5915 (2019) 02-0438-05

A Survey on Edge Computing

GE Yuetao, YIN Xiaotong

(Beijing HIWING Scientific and Technological Information Institute, Beijing 100074, China)

Abstract: Edge computing is a novel computing model after distributed computing, grid computing and cloud computing. This computing paradigm operates computing on the edge of the network. Because it provides intelligent and interconnection services nearby, edge computing could meet the key demand of the industry in the digital revolution process. This paper reviews the recent research and achievements of edge computing, and the future development tendency is then prospected. Firstly, the definition of edge computing is summarized, and its advantages in data security, interaction delay, bandwidth cost are discussed. The rapid development of edge computing has a synergistic and mutually reinforcing relationship with many frontier technologies and applications. Then this paper analyses the development tendency of edge computing from three perspectives: edge computing twines together with cloud computing, becoming the key to the realization of intelligence Internet of Things (IoT); the further development of edge computing needs the support of artificial intelligence chip; the commercialization and acceleration of 5G provides new opportunities for the development of edge computing.

Keywords: Edge Computing; Cloud Computing; Artificial Intelligence; Artificial Intelligence Chip; 5G; Internet of Things

1 引言

当前, 互联网和物联网数据量呈井喷式成倍增加, 为提升效率, 追求低延迟性已经成为一种趋势。然而, 数据从终端设备上传到云端, 计算后再回传至终端设备, 这种传统云计算技术已经无法满足人们对计算效率的高要求, 为此, 边缘计算(Edge Computing)技术应运而生^[1]。边缘计算是指在靠近物(如智能移动终端等)或数据源头

的一侧, 采用集“网络-计算-存储-应用”核心能力为一体的开放平台, 就近提供“最近端”服务。边缘计算通常处于物理实体和工业连接之间或物理实体的顶端, 其应用程序在边缘侧发起需求, 产生更快的网络服务响应, 满足行业在实时业务、应用智能、安全与隐私保护等方面的基本需求。对于智能制造而言, 边缘计算技术的不断突破与普及应用, 意味着许多控制、分析将通过本地设备实现, 而无需交由云端处理, 处理

收稿日期: 2019-01-28; 修回日期: 2019-02-25

过程也将在本地边缘计算层完成，这无疑将大大提升处理效率、减轻云端的负荷。

根据CB Insights公司的市场规模量化工具，到2022年，全球边缘计算市场规模预计将达到67.2亿美元；根据TrendForce公司预测，2018年至2022年全球边缘计算相关市场规模的年均复合增长率将超过30%；国际数据公司IDC统计数据显示，到2020年将有超过500亿个终端和设备接入网络，其中超过50%的数据需要在网络边缘侧分析、处理与存储，边缘计算技术将为未来的百亿终端提供人工智能能力；市场研究公司Tractica预测，到2025年，人工智能边缘设备的出货量将达到每年26亿个（边缘设备包括智能手机、智能家居扬声器、物联网摄像头、无人机、自动驾驶汽车和制造机器人等）。

2 边缘计算的概念

边缘计算是继分布式计算、网格计算、云计算之后，被认为是在网络的边缘执行计算的一种新型计算范式，其核心理念是使计算更靠近数据的源头^[2-3]。按照中国边缘计算产业联盟的定义，边缘计算是指靠近物或者数据源头的网络边缘侧，融合了网络、计算、存储、应用等核心能力的开放平台，就近提供边缘智能服务，满足行业数字在敏捷联接、实时业务、数据优化、应用智能、安全隐私等方面的关键需求^[4]。通俗来说，边缘计算就是将云端的计算存储能力下沉到网络边缘，用分布式的计算与存储在本地直接处理或解决特定的业务需求，从而满足不断出现的新业态对于网络高带宽、低延迟的硬性要求。

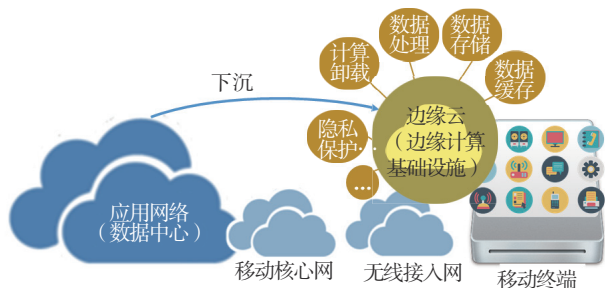


图1 边缘计算概念示意图

Fig.1 The concept sketch of edge computing

3 边缘计算的优势

由于数据只在源数据设备和边缘设备之间交换，不再全部上传至云计算平台，这也意味着边缘计算在下述问题上有着明显的先天优势。

第一，数据安全。传统云计算模型中，用户端的一切数据都需要上传到位于云端的统一数据中心，而在这个过程中，数据安全性就成了一个非常重要的问题。从电子金融账户密码，到搜索历史再到智能摄像头监控记录，这些个人隐私数据或者商业机密在上传到数据中心的过程中，都暗藏了数据泄露的风险。而在边缘计算架构中，由于数据的收集和计算都是基于本地，数据也不再被传输到云端，因此重要的敏感信息可以不经网络传输，能够有效避免传输过程中的泄漏。

第二，交互延迟。物联网应用面对的数据量极大，已经不再适合直接上传到云计算中心进行处理，不仅网络带宽压力大，对海量数据的搜索耗时通常也是无法接受的。边缘计算更靠近数据源，能够提供更实时、更快速的数据处理能力，由于减少了中间传输的过程，数据处理的速度也被加快。边缘计算的及时性和即时性对响应速度有苛刻时间要求的应用至关重要，比如自动驾驶应用、视频监控应用等。

第三，带宽成本。随着联网设备的增多，网络传输压力会越来越大，而边缘计算的过程中，与云端服务器的数据交换并不多，也不需要占用太多网络带宽。因此，边缘计算处理的数据是“小数据”，从数据计算、存储上都具有成本优势。这对基于互联网或者跨多个域数据转发的应用尤为关键。边缘计算既可以通过减少网络传输数据量来消减传输成本，还可以进一步提高云计算中心和数据中心的计算效率。

4 边缘计算的发展趋势分析

边缘计算的兴起、发展与普及，离不开新一轮技术及产业变革的重大机遇，为诸多边缘计算

技术愿景的实现与落地提供了助力。反过来，边缘计算技术与标准化体系的成熟，也为诸多技术的进一步跨越式发展提供了契机与支撑。下文将从边缘计算与云计算的协同发展关系、边缘计算与人工智能芯片的协同发展关系、边缘计算与5G的协同发展关系三个方面，分析边缘计算的发展趋势。



图2 与边缘计算协同发展的相关技术

Fig.2 The technologies coordinating with edge computing

4.1 边缘计算与云计算的协同发展分析

在科技飞速发展的今天，物联网已成为在公共云上运行的关键工作负载之一。虽然现在云端的物联网的平台化要比定制开发的线下平台能提供更大的价值，但是数据的延迟与宽带的消耗是目前面临的巨大难题之一，大多数企业难以承受延迟时间带来的巨大损耗与开支，边缘计算就是为了解决这个问题的。边缘设备充当“端”和“云”之间的中介，通过云的物联网的控制平面进行集中化的管理^[5]。因此，边缘计算被认为是继物联网和人工智能之后下一个热门技术。

边缘计算的发展对云计算有一定冲击，但它与云计算也有很强的协同。据国际数据公司IDC测算，到2021年，全球云计算市场的规模将达到5650亿美元，这其中约有20%为边缘云，可达到1130亿美元。全球云服务商为了守住原本的市场空间，纷纷提前布局边缘计算从而避免被吞噬。例如，2018年1月，亚马逊全球用户大会推出三款非云端产品，标志着传统云端服务巨头亚马逊开始发力边缘计算；2018年4月，亚马逊以机器学习推理支持的形式改版升级其边缘计算平台AWS Greengrass，凭借该平台对机器学习和深

度学习的最新支持，用户将能够构建自己的深度学习摄像机，并在边缘侧进行推理和分析。

2018年1月，卡内基梅隆大学启动普适感知、认知和网络基础设施的计算(Computing On Network Infrastructure for Pervasive Perception, Cognition and Action, CONIX)项目，该项目获得了2750万美元的资金，在未来五年中将重点探索创建位于边缘设备和云之间的网络计算架构，为边缘计算的深入发展与军事应用做准备。

2018年4月，微软宣布计划在未来四年内，向云操作系统、智能终端和智能物联网相关项目投资50亿美元，其中对边缘计算技术与产品的研发被摆在突出位置。2018年6月，微软进一步推动云业务与边缘计算业务的融合，宣布2017年公开预览的Azure IoT Edge边缘计算服务正式进入官方版，并通过GitHub共享平台将其开源。Azure IoT Edge的核心功能是将基于云的分析 and 定制的业务逻辑转移到边缘设备，对这些边缘设备进行即时数据处理，从而使用户能够专注于洞察商业机会而非数据管理。

对于戴尔等计算机服务器制造商而言，边缘计算被视为是一个反超的机会。2018年2月，戴尔子公司VMware在全球移动通信大会(Mobile World Congress, MWC)上发布一系列新的边缘计算解决方案，在边缘侧对由物联网设备收集的传感器数据进行实时分析，以解决资产管理和监控中的使用案例，这些边缘计算解决方案能够让企业用户以更轻松的方式，安全且有效地利用物联网产生的数据。2018年8月，VMware公布了用于边缘计算用例的设备和新软件Project Dimension，以及用于管理物联网连接设备的Pulse IoT Center软件的新版本，该套件将公司的旗舰vSphere服务器虚拟化平台与其存储和网络管理软件相结合。这表明作为传统数据中心领导者的戴尔正在接受和布局边缘计算。

为了满足物联网设备的需求，边缘计算和云计算需要协同工作：边缘计算是云计算的延伸，它与云计算各有所长，云计算擅长全局性、非实

时、长周期的大数据处理与分析；边缘计算擅长现场级、实时和短周期智能分析等。来自智能设备和传感器的所有数据仍然需要在云上进行汇总，这需要更深入的分析，以便从中获取有意义的见解，所以云计算仍然在使物联网设备更智能和更好的过程中发挥关键作用。云计算和边缘计算正在塑造智能物联网的未来^[6]。这种组合为物联网网络中连接的设备带来了稳定性，并通过处理更接近源头的数据来解决延迟问题。

4.2 边缘计算与人工智能芯片的协同发展分析

在近年来人工智能芯片不断革新的大背景下，边缘计算与人工智能芯片已成协同共进之势。随着智能移动终端及物联网应用的不断普及，将数量巨大的智能移动终端所产生的数据悉数上传至云端是不现实也是不可取的，其中很大比例数据需要在“边缘侧”完成对数据的处理与分析。在这种应用需求下，“低延迟”的处理将会成为主要难题——如何实现在毫秒级时间甚至微秒量级时间内完成处理，这便需要依靠边缘计算技术。然而发展边缘计算，人工智能芯片是必不可缺的。

2018年2月，美国半导体公司AMD推出了两款面向边缘计算的嵌入式处理器，分别是EPYC 3000系列和Ryzen V1000系列，均采用代号为ZEN的CPU架构。2018年2月，英特尔推出了最新一代的“至强 D (Xeon D)”系列人工智能处理器，该系列处理器基于Skylake架构，重点瞄准的是边缘及其他一些受限的环境，这些环境对硬件密度、电力消耗、智能支撑的问题更为敏感。2018年2月，ARM公司公布延龄草(Trillium)项目，旨在通过优化的人工智能芯片以运行那些利用了张量流(以TensorFlow系统为代表)和卷积神经网络框架(以Caffe框架为代表)等神经网络框架的应用程序和软件，来驱动位于边缘侧设备的机器学习、人工智能和目标检测能力。

2018年7月，谷歌发布张量处理器(TPU)的轻量级版本——Edge TPU，专为在边缘运行TensorFlow Lite ML模型而设计，它能在很小的物

理占用和低功耗的限制下提供高性能运算，可以充当传感器或网关设备中的标准芯片或微控制器，使得在边缘部署高精度的人工智能成为可能。这标志着谷歌不仅在自己的数据中心开发人工智能芯片，还将其设计的Edge TPU应用在其他公司生产的产品中。2019年3月，谷歌发布基于Edge TPU的边缘计算专用芯片Coral。2018年11月，英特尔人工智能大会推出英特尔神经计算棒二代(英特尔NCS 2)，该计算棒可以在网络边缘构建更智能的人工智能算法和计算机视觉设备。英特尔认为边缘侧人工智能的最大机会将是视觉领域——机器视觉、视频监控、医疗影像等。英特尔NCS 2基于其视觉处理单元(Vision Processing Unit, VPU)，得到英特尔OpenVINO工具包的支持，与上一代神经计算棒相比性能更优，能够以可负担的成本显著加快深度神经网络推理应用的开发。此外，英特尔正在研发全新专门面向5G无线接入和边缘计算的基于10nm工艺的网络系统芯片，研发代号“Snow Ridge”，计划于2019年下半年交付给合作伙伴，并于2020年初推出产品。

支撑边缘计算的人工智能芯片，需要边缘设备在性能、功耗与尺寸之间进行平衡。从发展趋势看，边缘计算专用人工智能芯片需要在架构复杂度、支持人工智能算法多样化以及多场景适应性上不断创新和提升。此外，还需要“边-云”协同的人工智能体系架构来降低应用开发和部署成本，以便更有效地利用基础设施的资源。

4.3 边缘计算与5G的协同发展分析

第五代移动通信网络(5G网络)的即将商用，为边缘计算的发展提供了新的机遇。5G所具有的延时小、带宽大、容量大等优势，解决了传统通信领域里遇到的很多问题，但是也导致数据量的极速增长，这时候亟需提供可靠、有用、可执行的商业模式。5G的快速处理、低延迟等特点可以在迅速响应方面提供一个新的途径，能够对端、边缘、云上进行联合优化。边缘计算的这种能力，可以从用户体验、功耗、计算负载、性能、成本

等方面，在物联网设备、边缘设备和云设备之间智能配置资源，为联合优化提供了一种新的途径。因此，边缘计算技术的发展与5G有着密切的关系：一方面，边缘计算能够给予5G支持，5G的重要组成部分便是边缘计算；另一方面，因为5G是以软件的形式进行表现，恰好可以灵活运用边缘计算^[8]。

在欧洲市场，边缘计算产业已经形成了产业联盟，以沃达丰、德国电信、西门子等公司为代表的大型科技企业已经加入其中。欧洲电信标准化协会(European Telecommunications Standards Institute, ETSI)已启动标准化移动边缘计算(Moving Edge Computing, MEC)的制定，运营商可以向授权的第三方开放其无线网络边缘，使其能够灵活快速地为移动用户、企业和垂直网段部署创新的应用和服务^[9-10]。移动边缘计算是移动基站迭代和IT与电信网络融合的自然发展的结果，将为消费者和企业客户提供新的垂直业务服务，包括视频分析、位置服务、物联网、增强现实、优化本地内容分发和数据缓存等。2018年2月，ETSI发布两篇白皮书，分别是《云端无线接入网络(Cloud RAN)和移动边缘计算：完美配对》和《4G中的移动边缘计算部署以及向5G演进》，以实现移动边缘计算与5G保持同步。

5 结束语

边缘计算能就近提供智能互联服务，满足行业在数字化变革过程中的关键需求。人工智能与5G的迅速落地，与边缘计算的发展密不可分，边缘计算必将是人工智能之后的下一个热点。随着移动网络向5G演进的速度进一步加快，边缘计算将在大流量业务的普及下发挥更多价值，如降低核心网压力、提升接入网的能力与价值等。随着可穿戴设备及附带传感器的智能设备数量呈爆发式增长，未来更多设备将被接入物联网，边缘侧分析计算能力需求将倍增。边缘计算正在不断融合和完善运营、信息和通讯(OICT)，并有效地推动产业的发展，这也标志着边缘计算技术已经迈进了快速发展的道路。目前，智慧安防、工业互

联网、智能家居等已成为当下边缘计算主要的应用场景，各厂商在这些领域也都有了完整的解决方案。

参考文献

- [1] 施巍松, 孙辉, 曹杰, 等. 边缘计算: 万物互联时代新型计算模型[J]. 计算机研究与发展, 2017, 54(5): 907-924.
- [2] Mao Y, You C, Zhang J, et al. A survey on mobile edge computing: the communication perspective[J]. IEEE Communications Surveys & Tutorials, 2017, 19(4): 2322-2358.
- [3] Satyanarayanan M. The emergence of edge computing[J]. Computer, 2017, 50(1): 30-39.
- [4] 洪学海, 汪洋. 边缘计算技术发展对策研究[J]. 中国工程科学, 2018, 20(2): 28-34.
- [5] 赵梓铭, 刘芳, 蔡志平, 等. 边缘计算: 平台、应用与挑战[J]. 计算机研究与发展, 2018, 55(2): 327-337.
- [6] Jararweh Y, Doulat A, Alqudah O, et al. The future of mobile cloud computing: integrating cloudlets and mobile edge computing[C]. International Conference on Telecommunications. IEEE, 2016.
- [7] Taleb T, Samdanis K, Mada B, et al. On multi-access edge computing: a survey of the emerging 5G network edge architecture & orchestration[J]. IEEE Communications Surveys & Tutorials, 2017, 19(3): 1657-1681.
- [8] Huang S C, Luo Y C, Chen B L, et al. Application-aware traffic redirection: a mobile edge computing implementation toward future 5G networks[C]. IEEE International Symposium on Cloud & Service Computing, 2017.
- [9] 李子姝, 谢人超, 孙礼, 等. 移动边缘计算综述[J]. 电信科学, 2018, 34(1): 87-101.
- [10] 吕华章, 陈丹, 范斌, 等. 边缘计算标准化进展与案例分析[J]. 计算机研究与发展, 2018, 55(3): 487-511.

作者简介:

葛悦涛(1982-), 男, 博士, 高级工程师, 主要研究方向为智能通信、导航、人工智能技术等。

尹晓桐(1989-), 女, 本科, 工程师, 主要研究方向为情报跟踪分析、无人系统技术、科技期刊出版等。

5G在无人驾驶汽车中的应用前景分析

张朝阳, 李星宏

(北京亿欧网盟科技有限公司, 北京 100020)

摘要: 伴随着5G技术以及汽车工业的发展, 基于5G的无人驾驶汽车近年来受到了政府、媒体以及相关企业的关注。首先从两个方面对5G的发展进行了整体概述, 包括5G的优势以及国际电信联盟ITU确定的5G三大应用情景, 即增强移动宽带、超高可靠低时延通信、大规模机器类通信。随后分析了5G如何应用于无人驾驶汽车领域, 并简单阐述了当前企业的研究方向。研究表明, 现阶段5G应用于无人驾驶领域还处于测试阶段, 在未来2~3年内, 有望实现特定场景的大规模商用。

关键词: 5G网络; 无人驾驶; V2X; 智慧交通; 增强型移动宽带; 超高可靠与低时延通信; 大规模机器类通信

中图分类号: TN929.5; U463.6 文献标识码: A 文章编号: 2096-5915 (2019) 02-0443-04

Analysis of the Application Prospect of 5G in Driverless Cars

ZHANG Zhaoyang, LI Xinghong

(EO Network Technology Co., Ltd, Beijing 100020, China)

Abstract: With the development of 5G technology and the automotive industry, driverless cars of 5G-based have recent years, it has received special attention from the government, media and related companies. At first, this paper gives a general overview of the development of 5G from two aspects, including the advantages of 5G and the three major application scenarios defined by the ITU, including enhanced mobile bandwidth, ultra-reliable low-latency communication, and large-scale machines Communication. Then, the second part analyzes how 5G is applied to the driverless car field and briefly explains what the current related companies focus on. The research has shown that the current 5G application in the unmanned field is still in the testing stage, and it is expected to realize large to scale commercial use of specific scenarios in the next 2 to 3 years.

Keywords: 5G Network; Driverless; V2X; Smart Transportation; Enhance Mobile Broadband; Ultra Reliable & Low Latency Communication; Massive Machine Type of Communication

1 引言

移动通信技术到现在为止已经发展到了第五代(5G)。第一代(1G)利用了模拟技术, 只是满足了基本的语音通话功能。从第二代(2G)开始,

移动通信开始使用数字技术, 并增加了短信以及简单的上网功能。而随着第三代(3G)、第四代(4G)网络的发展, 又促进了智能手机、移动视频以及直播平台的兴起, 且移动带宽的增强大大地增加了移动设备的连接数量。前四代移动通信主

要实现的是人与人的连接，完成了互联网到移动互联网的转变。而伴随着第五代通信技术的发展，连接人与人的移动互联网时代将迈向“万物互联”的物联网时代，包括家居、汽车、机器人以及高速公路、城市等一切将被联网。在汽车工业领域，汽车联网的终极形态是实现汽车的无人驾驶，无人驾驶汽车需要具有低时延以及高稳定性的传输，而这正是5G的两大优势，5G的发展势必会推动无人驾驶汽车的进一步发展。本文将着重探讨5G在无人驾驶汽车领域的应用。

2 5G的概述及其发展

2015年6月24日，国际电信联盟 (ITU) 正式公布了5G技术的名称为IMT-2020^[1]，到目前为止，5G仍处于标准制定阶段，据估计，2019年底有望实现5G的商用。5G不仅能够实现人与人的连接，还能实现人与物、物与物的连接，因此，为了满足将来各种各样的应用场景，ITU提出了5G技术应具有8大关键技术指标^[2]，如表1所示。

可以看出，相对比于4G，5G要求的各项指标性能均提高了几倍到上百倍不等。5G不仅是4G的延伸，同时也是第一个全球统一标准的网络，伴随着5G标准的落地完成，5G也将应用在各种各样的实际场景中。ITU还确定了未来5G应具有的三大应用情景，如图1所示。

第一个是增强型移动宽带 (Enhance Mobile

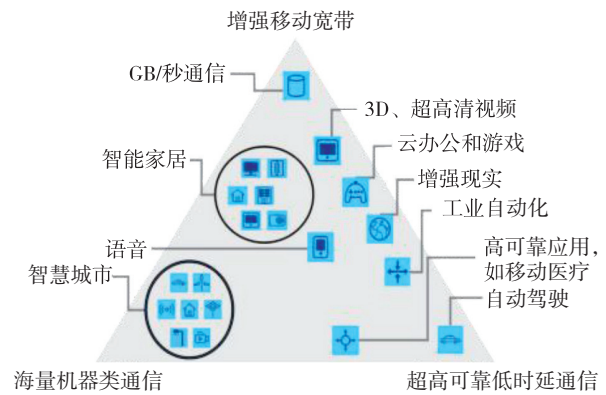


图1 5G的三大应用情景

Fig.1 Three application scenarios of 5G

Broadband, EMBB)，其主要带来的是移动连接速率的大幅提升，现阶段的智能手机能享受到的峰值网络速度大约为几百兆每秒，而到5G时代，可实现千兆级别的连接速度，为智能手机上的AR/VR游戏提供了便捷操作。第二个是超高可靠与低时延通信 (Ultra Reliable & Low Latency Communication, URLLC)，这种类型的应用情景需要非常低的时延，尤其是在无人驾驶领域。5G技术能让系统的延迟降低至1ms，4G的延迟大约为50ms，而人类最快的反应也仅为100ms，5G的要求基本能达到安全的无人驾驶。第三个是大规模 (海量) 机器类通信 (Massive Machine Type of Communication, MMTC)。现有的4G通信技术架构，每平方公里只能支持几千个用户同时在线，而到了5G时代，每平方公里能够支持几万甚至是几十万的物联终端同时在线，连接密度的大幅

表1 5G技术的8大关键技术指标
Table 1 Eight key technical indicators of 5G

技术指标名称	技术指标含义	4G 要求	5G 要求	性能提升情况
用户体验速率	真实网络环境下，用户可获得的最低传输速率	0.01Gbps	0.1~1Gbps	10~100倍
用户峰值速率	单个用户可获得最高传输速率	1Gbps	20Gbps	20倍
移动性	获得指定的服务质量，收发双方间获得的最大相对移动速度	350km/h	500km/h	提升30%
时延	数据从源节点到目的节点的时间间隔	20~30ms	低至1ms	数十倍
连接数密度	单位面积内的连接数量总和	10万台设备/km ²	100万台设备/km ²	10倍
能量效率	单位能量所能传输的比特数	1倍	100倍	100倍
频谱效率	单位带宽数据的传输速率	1倍	3倍	3倍
流量密度	单位面积内的总流量	0.1~0.5Tbps/km ²	数十Tbps/km ²	数百倍

度增加，为今后大规模的物联网发展，即万物互联提供了可能。

这三大应用情景分别代表了大带宽、低时延和广连接，在ITU确定的三大应用情景中，包括了多种类型的实际应用，如智能家居、3D超高清视频、自动驾驶等，而在超高可靠低时延类通信的应用中，自动驾驶又是最重要的应用方向之一（自动驾驶也可被称为无人驾驶，无人驾驶是自动驾驶的高级阶段）。在现阶段，行业内企业布局无人驾驶存在三种不同的形态，第一种形态是汽车厂商自己进行创新和提升，如宝马、大众等；第二种形态是互联网公司进入这个领域，如百度、谷歌等；第三种形态是汽车厂商与互联网公司进行合作。当然，不论是哪一种形态，5G都将是实现无人驾驶最基本的技术之一，通过5G，可以实现交通灯、路灯等城市基础设施以及汽车、行人等的连接。除了5G，也有部分企业利用基于专用短程通信（Dedicated Short Range Communications, DRSC）进行连接，但该连接存在许多不足之处，通信的覆盖距离较短且需要针对路边设施进行大规模改造和资金投入^[3]。只有使用5G技术，才有可能真正的实现无人驾驶。

3 5G在无人驾驶汽车领域的应用

根据美国汽车工程师协会（SAE）的划分标准，

将自动驾驶汽车分为了六个等级，分别为无自动化（L0）、驾驶支援（L1）、部分自动化（L2）、有条件自动化（L3）、高度自动化（L4）和完全自动化（L5）^[4]，详细标准见表2。而实际上，从L3开始，无人驾驶系统才开始完成相应的驾驶操作。L1、L2属于辅助驾驶，L3、L4属于自动驾驶，只有达到L5才算真正的无人驾驶。无人驾驶也是联网汽车、自动驾驶的终极形态。

真正的无人驾驶需要实现V2X（Vehicle To Everything）连接，X代表的是车、路、行人以及周围环境。英特尔原CEO科再奇表示，未来的无人驾驶汽车，每辆车每小时需要处理的数据量将达到100GB以上，且需要延迟低至几毫秒，而目前4G网络的延迟还不足以满足未来无人驾驶的需要。同时在遇到突发状况需要紧急停车时，低延迟也能够保证足够短的制动距离以及行车安全。假如汽车的行驶速度为60km/h，以4G的时延来说，50ms时延的制动距离为0.83m，而以5G的时延1ms来计算，制动距离仅为0.0167m，极大地提升了驾驶的安全性^[5]。同时，4G网络还面临着功耗大、传输带宽不足、稳定性差等问题，而5G特有的优势为无人驾驶的发展提供了技术支持。

无人驾驶需要感知、决策、执行三个层面，和多达几十种零部件高效、稳定地配合才能完成^[6]。感知层主要采用摄像头、激光雷达、毫米波雷达、高精度地图等进行采集数据。决策层和执行层，通

表2 自动驾驶技术分级标准

Table 2 Automated driving technology grading standards

自动驾驶分级	称呼（SAE）	SAE定义	驾驶操作	周边监控	支援	系统作用域
L0	无自动化	由人类驾驶者完全操作，在驾驶过程中可以得到警告和保护系统的辅助	人类驾驶者			无
L1	驾驶支援	通过驾驶环境对方向盘和加减速中的一项操作提供驾驶支援，其他的驾驶动作都由人类驾驶者操作	人类驾驶者和系统	人类驾驶者	人类驾驶者	部分
L2	部分自动化	通过驾驶环境对方向盘和加减速中的多项操作提供驾驶支援，其他的驾驶动作都由人类驾驶者操作				
L3	有条件自动化	由无人驾驶系统完成所有的驾驶操作，根据系统请求，人类驾驶者提供适当的应答	系统			
L4	高度自动化	由无人驾驶系统完成所有的驾驶操作，根据系统请求，人类驾驶者不一定需要对所有的系统请求作出回答，限定道路和环境条件等		系统	系统	
L5	完全自动化	由无人驾驶系统完成所有的驾驶操作，人类驾驶者在可能的情况下接管，在所有的道路和环境条件下驾驶				全程

过运用采集回来的数据，建立相应的模型，分析制定出最适合的执行策略，以实现无人驾驶^[7]。

运用5G技术的无人驾驶汽车一定是最终的发展方向，不过目前来看条件还不成熟，尚处于试验阶段。2018年3月，中兴通讯携手中国电信、百度在河北雄安新区完成了基于5G网络实况环境下的无人驾驶车测试。2018年9月，景驰科技与广东联通合作，测试了基于5G网络下的L4级别的无人驾驶应用场景：远程控制无人车。在展示环节，景驰无人驾驶车平稳地开进了现场，完成了直行、转弯、倒车等相应动作。2018年11月，天津联通联合中国汽车技术研究中心、华为共同打造了5G+V2X融合网络无人驾驶业务试点。2019年1月，由重庆移动、华为、东南大学、法国Easy Mile公司等联合研发测试的5G无人驾驶巴士开始试运行，这辆巴士能基于5G网络，准确的绕过障碍物，并完成自动调整前进速度等操作。

国外也在同一时期进行了5G无人驾驶的测试。在2018年世界移动通信大会上，韩国电信运营商SK Telecom展示了虚拟城市K-City中基于5G+V2X技术的无人驾驶试验成果，测试的两辆车均实现了实时的大量数据传输以及精准控制。2019年3月，日本索尼公司与日本移动通信运营商NTT DoCoMo宣布，双方将合作测试索尼的新概念车：New Concept Cart SC-1，利用5G技术，实现了多种遥控的功能。现阶段基于5G技术的无人驾驶汽车企业大部分处于测试阶段，且仅能实现人与汽车的单项通信，若要实现V2X，还需要汽车制造商、通信设备商以及运营商的共同努力^[8]。

4 结束语

5G的一个重要应用方向就是无人驾驶汽车，但现阶段，由于5G网络存在行业发展以及相应标准制定等问题，离实际商用还有较大距离。尤其是在安全性问题上，当车辆的各个零部件与5G进

行连接传输时，极大地增加了可被攻击的范围。而且现在的无人驾驶汽车均处于单车智能状态，短时间内很难达到完全L5级别的全自动无人驾驶。若要进行实际商用，成本也是一个需要考量的重要因素，以激光雷达为例，谷歌使用的单个激光雷达成本就达到了50~70万元，成本之高，显然超出了普通家庭的能力范围。另一方面，根据华为的整体规划，2020年开始部署5G商用网络，而真正的应用也要在几年之后。2到3年内，无人驾驶汽车还不会大范围的出现，但根据当前企业发展状况来看，大概率会出现一些特定场景的实际商用，如园区、学校等。可以想象的是，在不远的将来，5G应用于无人驾驶汽车领域，使得驾驶员也可以变成“自由自在”的乘客。

参考文献

- [1] 王世宝. 基于5G技术车联网的发展趋势及应用前景分析[J]. 时代汽车, 2018, 297(6):170-171.
- [2] 啜钢, 裴静, 刘倩, 等. 安防系统中5G关键技术分析[J]. 中国安防, 2017(12):71-77.
- [3] 许彩霞. 5G车联网对自动驾驶技术发展的影响[J]. 信息通信, 2018, 186(6):51-52.
- [4] 伦一. 自动驾驶产业发展现状及趋势[J]. 电信网技术, 2017(6):40-43.
- [5] 刘川, 陈金鹰, 朱正模, 等. 5G对无人驾驶汽车的影响分析[J]. 通信与信息技术, 2017(3):43-44.
- [6] 辛煜. 无人驾驶车辆运动障碍物检测、预测和避撞方法研究[D]. 中国科学技术大学, 2014.
- [7] 乔维高, 徐学进. 无人驾驶汽车的发展现状及方向[J]. 上海汽车, 2007(7):40-43.
- [8] 孙宇, 崔娜. 5G让自动驾驶成为现实[J]. 信息与电脑(理论版), 2019(2):183-184.

作者简介

张朝阳(1991-), 男, 硕士, 主要研究方向为5G相关技术及其商业应用。

李星宏(1995-), 男, 本科, 主要研究方向为自动驾驶等新兴技术发展趋势。