

单级特征图融合坐标注意力的视觉位置识别方法*

刘子健 张军 刘元盛 路铭 宋庆鹏

(北京联合大学,北京市信息服务工程重点实验室,北京 100101)

【摘要】针对视角和环境变化的场景中现有视觉位置识别方法存在的匹配遗漏和实时性差的问题,提出基于单级特征图融合坐标注意力的视觉位置识别方法。首先通过坐标注意力捕获特征的相对位置信息,然后利用扩张卷积和局部聚合向量网络(NetVLAD)构造多尺度特征融合的编码器,最后基于三元组损失训练网络。经 Pitts30k 和 Nordland 数据集验证,在位置识别试验中,与同基线的先进方法 Patch-NetVLAD 相比,所提出的方法能够获得同等的召回精度且检索速度提高 19%。在回环检测试验中,所提出的方法达到了合理平衡鲁棒性和检索速度的目标。

关键词:自动驾驶 视觉位置识别 回环检测 坐标注意力 局部聚合向量网络 三元组损失
中图分类号:TP391.4;U463.6 **文献标识码:**A **DOI:** 10.19620/j.cnki.1000-3703.20220668

Visual Place Recognition with One Level Feature Fusing Coordinate Attention

Liu Zijian, Zhang Jun, Liu Yuansheng, Lu Ming, Song Qingpeng

(Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101)

【Abstract】To solve the problems of matching omission and poor real-time performance of existing visual place recognition methods in scenes with changing viewpoints and environments, this paper proposed visual place recognition method based on one level feature fused with coordinate attention. Firstly, the relative place information of features was captured by coordinate attention. Secondly, an encoder for multi-scale feature fusion was constructed using dilated convolution and NetVLAD. Finally, the network was trained based on triplet loss. Validated by Pitts30k and Nordland datasets, the proposed method achieves the same recall accuracy and 19% faster retrieval speed compared with the state-of-the-art method Patch-NetVLAD of the same baseline in the test of position recognition. In the test of loop detection, the proposed method achieves a reasonable balance between robustness and retrieval speed.

Key words: Autopilot, Visual place recognition, Loop detection, Coordinate attention, NetVLAD, Triplet loss

【引用格式】刘子健,张军,刘元盛,等.单级特征图融合坐标注意力的视觉位置识别方法[J].汽车技术,2023(3):19-25.

LIU Z J, ZHANG J, LIU Y S, et al. Visual Place Recognition with One Level Feature Fusing Coordinate Attention[J]. Automobile Technology, 2023(3): 19-25.

1 前言

自动驾驶汽车基于同步定位与建图(Simultaneous Localization and Mapping, SLAM)^[1]构建自主导航系统。在长期运行中,定位系统会产生累积误差,导致定位失效。实践中,自动驾驶汽车先存储已访问过场景的表征向量,再在汽车运动时匹配当前环境的表征向量解决定位问题,即回环检测与重定位。随着自动驾驶汽车的发

展,以视觉位置识别(Visual Place Recognition, VPR)^[2]技术为代表的回环检测与重定位方法受到关注。VPR以环境中的视觉特征为基础,计算形成环境表征向量^[3]。目前应用于VPR技术的视觉特征主要分为基于规则设计的手工特征和基于数据驱动的学习特征。

基于规则设计的角点和描述子组合的ORB(Oriented FAST and Rotated BRIEF)^[4]局部特征具有尺度和旋转不变性,对图像噪声及透视变换具有鲁棒性。因此,ORB

*基金项目:国家重点研发计划项目(2021YFC3001300);国家自然科学基金重点项目合作项目(61931012);

北京市自然科学基金面上项目(4222025);北京联合大学高水平孵化项目和新进博士孵化项目(ZK10202208)。

通讯作者:张军(1971—),女,山东临沂人,副教授,博士,研究方向为自动驾驶感知与决策技术,xxtzhangjun@buu.edu.cn。

特征能准确地描述和检索稳定不变的环境信息。但此类方法可能受到视角转换、光照和季节变换等因素影响,产生当前观测的环境信息与数据库中已知的环境信息存在较大差异的现象,最终导致位置识别结果的准确率降低等问题。

基于数据驱动的卷积神经网络(Convolutional Neural Networks, CNN)^[5]具有强大的图像表示能力。为在现有VPR方法的识别精度下提高定位系统的鲁棒性和实时性,本文针对环境中的视觉特征空间关系学习和高层特征学习两方面进行改进:在骨干网络中融合坐标注意力,获得视觉特征的空间位置关系和显著的空间位

置信息;利用多重扩张卷积和局部聚合向量网络,构造多尺度特征融合的特征编码器,获得单级特征图在不同尺度下的视觉特征向量。最后,将该方法在公开数据集上进行验证。

2 本文方法

本文基于残差网络(Residual Networks, ResNet)^[6]提出一种弱监督训练的VPR方法,实现高效的视觉位置识别:输入图像集或图像序列,通过本文搭建的神经网络得到特征匹配结果,最终输出检索结果。系统流程如图1所示。

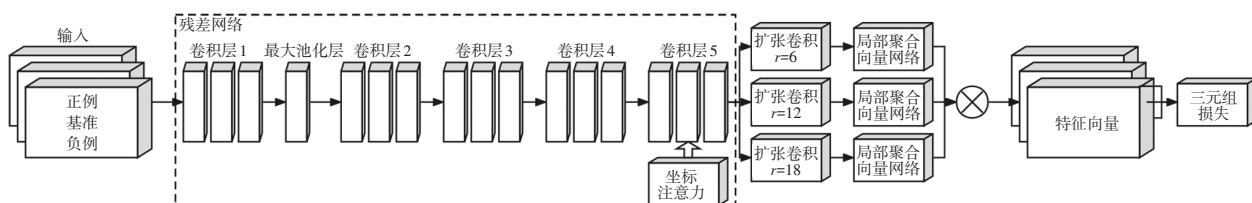


图1 系统流程

2.1 特征提取

2.1.1 残差网络

深层CNN存在参数爆炸、梯度消失等问题。ResNet通过短路连接解决深层网络梯度消失带来的学习退化问题。当理论残差为零时,残差结构只做恒等映射以确保实际残差不会为零,同时保持当前网络性能,结构如图2所示。

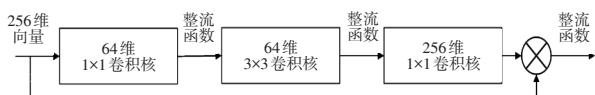


图2 残差结构

本文采用ResNet-50作为提取图像特征的骨干网络。自动驾驶汽车采集的视觉信息在ResNet中经过卷积(Convolutional, Conv)层、批归一化(Batch Normalization, BN)层、激活函数和池化层完成预处理,随后再经过4组瓶颈(Bottleneck)层提取特征信息。

2.1.2 坐标注意力网络

坐标注意力(Coordinate Attention, CA)^[7]网络将输入的二维全局池化转化为单个特征向量的过程,分解为水平方向和垂直方向一维特征编码过程,将聚合的方向信息嵌入特征图中。通过级联捕获的2个空间方向的依赖关系和位置信息,经BN层和非线性变换生成空间信息的特征图。沿着水平和垂直方向切分特征图得到注意力权重,使网络捕获到特征的位置关系。网络结构如图3所示。

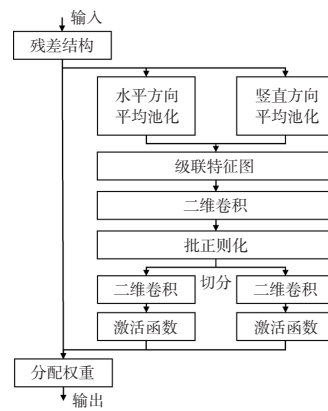


图3 坐标注意力网络结构

2.1.3 融合CA的ResNet特征提取网络

视觉特征的相对位置关系对于位置识别结果至关重要。为了在开放环境中保持定位系统的鲁棒性,需要网络模型具备筛选重要的环境特征和获取视觉特征坐标的能力。为弥补ResNet对图像特征的空间关系不敏感的缺陷,本文在其残差结构中引入CA,构建新的特征提取网络。通过捕获环境特征的空间位置关系,获得特征间的依赖关系和精确的位置信息。

CNN的浅层输出边缘和色彩等具体的图像特征,随着网络层数的加深,输出的特征图尺寸逐渐减小,视觉信息逐渐抽象和语义化。ResNet的Conv5特征图包含对图像变化足够鲁棒的视觉特征,同时具有丰富的语义信息。因此,本文在Conv5的短路连接和瓶颈层之间引入CA,将生成的特征图编码为一组方向感知和位置

敏感的注意力热图,促使网络学习到特征间隐含的空间结构关系,帮助模型准确定位感兴趣的目标,增强特征图表达能力。改进后的网络利用CNN提取图像的局部信息,发挥了CA对全局信息建模的能力,有效提高视觉特征的表达能力。改进后的网络结构如图4所示。

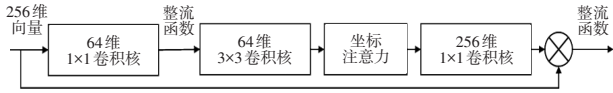


图4 改进的坐标-残差结构

2.2 特征编码

对环境的局部视觉特征进行建模,不足以区分局部特征相似的场景,需要联系上下文信息和环境的整体信息对场景进行同一性判断。特征金字塔(Feature Pyramid Network, FPN)是获取多尺度特征的经典网络,但车载设备的算力普遍较低,导致基于FPN的VPR方法移植到车载设备存在实时性差的问题。扩张卷积(Dilated Convolution, DC)^[8]具有仅增加1个参数即可获得更广视野范围特征图的特点。因此,本文基于DC和局部聚合向量网络(Vector of Locally Aggregated Descriptors Network, NetVLAD)构造图像编码器,利用DC-NetVLAD对Conv5卷积层输出的单级特征图进行视野扩张,然后对不同视野下的图像特征进行聚合得到图像的描述向量,在不失精度的前提下节约计算成本,提高系统的实时性。

2.2.1 局部聚合向量网络

基于局部特征解决VPR问题,需要对特征矩阵进行聚类处理以压缩数据维度。相比视觉词袋(Bag of Word, BoW)编码,局部聚合描述符向量(Vector of Locally Aggregated Descriptors, VLAD)不仅保存每个特征点到离它最近的聚类中心的距离,还记录了特征点在不同维度的取值,避免了信息损失问题。VLAD的计算公式为:

$$V(J, K) = \sum_{i=1}^N a_k(\mathbf{X}_i)(\mathbf{x}_i(j) - \mathbf{c}_k(j)) \quad (1)$$

式中, \mathbf{X}_i 为输入的数量为 N 的局部特征向量; $V(J, K)$ 为输出的聚类后的特征向量; K 为聚类中心的数量; J 为第 k 个特征向量的维度; $\mathbf{x}_i(j)$ 为第 i 个局部特征的第 j 个特征值; $\mathbf{c}_k(j)$ 为第 k 个聚类中心的第 j 个特征值; $a_k(\mathbf{X}_i)$ 为二值的符号函数,对每个属于聚类中心 \mathbf{c}_k 的特征向量 \mathbf{x}_i , a_k 的值取为 1, 否则为 0。

原始的 VLAD 为不可导函数,为得到端到端的特征聚类结果,Arandjelovic 等^[9]在 VLAD 算法的基础上,用归一化指数函数(Softmax)代替原始方法中的最近邻二值

函数 a_k , 提出 NetVLAD 算法。NetVLAD 设定的聚类的数量为 K , 计算局部特征在这些聚类的差值分布得到全局特征,使 VLAD 算法中需要手工聚类获得的参数改为通过网络训练获取。系数 $\bar{a}_k(\mathbf{X}_i)$ 的计算公式为:

$$\bar{a}_k(\mathbf{X}_i) = \frac{e^{2\alpha \bar{c}_k \cdot \mathbf{x}_i - \alpha \|\bar{c}_k\|^2}}{\sum_{k=1}^K e^{2\alpha \bar{c}_k \cdot \mathbf{x}_i - \alpha \|\bar{c}_k\|^2}} \quad (2)$$

式中, \bar{a}_k 为基于 Softmax 计算每个局部特征到聚类中心的距离时的权重,其取值与距离成反比,特征离聚类簇中心越近,则其权重分配越高,远则权重分配越低,即当该特征属于聚类中心时,把局部特征描述子 \mathbf{X}_i 在每个聚类簇下的权重调整为 1, 否则为 0; \mathbf{x}_i 为特征向量; \bar{c}_k 为网络学习的聚类中心; α 为权重系数。

2.2.2 基于 DC 和 NetVLAD 构造特征编码器

首先去掉 ResNet-50 网络最后的全局平均池化和全连接层,利用 1×1 卷积对残差网络输出的特征图进行通道压缩减少计算量,然后经过 3 个并联的具有不同扩张率参数的扩张卷积,得到不同视野下的图像特征。为避免扩张卷积中心点不连续,本文的扩张率参数分别设定为 6、12 和 18。在 NetVLAD 中,经过激活函数 Softmax 和局部聚合层的池化获得 VLAD 特征,通过类内正则化对每个聚类中心的所有残差分别进行正则化,最后将所有 VLAD 特征共同进行欧式范数正则化得到图像全局表示的特征向量。改进后的 DC-NetVLAD 编码器结构如图 5 所示。

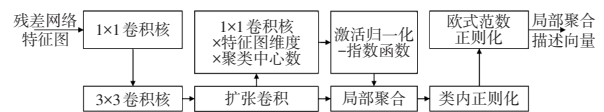


图5 本文改进的扩张卷积-局部聚合向量网络编码器

2.3 基于三元组损失训练网络

NetVLAD 算法首先将图像的局部特征聚类压缩,得到以聚类中心表示的全局特征图,然后将待查询图像的特征向量与数据库中的图像向量进行相似度计算,以余弦相似度为度量,距离越小的图像之间越相似。余弦距离的计算公式为:

$$d(\mathbf{G}, \mathbf{H}) = 1 - \frac{\sum_{i=1}^n (g_i \times h_i)}{\sqrt{\sum_{i=1}^n (g_i)^2} \times \sqrt{\sum_{i=1}^n (h_i)^2}} \quad (3)$$

式中, \mathbf{G}, \mathbf{H} 分别为输入的 2 个特征向量; n 为向量维度; g_i, h_i 分别为 \mathbf{G}, \mathbf{H} 第 i 个维度的特征值。

针对环境中普遍存在着视觉特征相似场景的情况,本文使用三元组损失(Triplet Loss, TL)^[10]训练网络来提

高系统的查准能力。TL的输入是一组编码后的图像特征,包括正例、负例和基准图像。基于TL训练的网络能够优化图像与特征向量之间的映射关系,在新的特征空间中使负例到基准图像的距离尽可能大于正例到基准图像的距离,从而提升网络识别正例的能力,效果如图6所示。

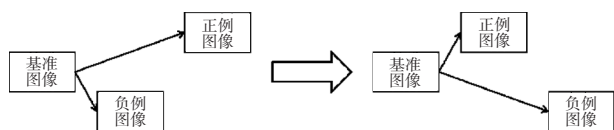


图6 三元组损失

根据模型训练的需要设定边缘阈值 β 控制正、负样本的距离。TL目标函数的表达式为:

$$L = \sum_i^N \left[\left\| f(x_i^+) - f(x_i^+) \right\|_2^2 - \left\| f(x_i^+) - f(x_i^-) \right\|_2^2 + \beta \right] \quad (4)$$

式中, $f(x_i^+)$ 为输入的数量为 N 的三元组中的第 i 个基准样本的特征值; $f(x_i^+)$ 为第 i 个正例的特征值; $f(x_i^-)$ 为第 i 个负例的特征值。

超参数 β 控制损失函数的结果: $L>0$ 时取该值作为损失; $L<0$ 时,表示负例与基准样本的距离大于正例与基准样本的距离,因此损失函数的结果记为0。

3 试验设置与结果分析

3.1 试验环境和数据集说明

本文的试验环境为Ubuntu 18.04 LTS操作系统,硬件平台采用11代英特尔i7处理器、32 GB内存和英伟达RTX 2080Ti显卡。

采用Pitts30k和Nordland^[12]2个公开数据集进行测试。Pitts30k是谷歌公司制作的都市街景数据集,该数据集基于地理位置进行划分,包括训练、验证和测试部分,分别含有约7 000张查询图像和10 000张参考图像。Nordland数据集采集自挪威北部地区,包含4个帧率为25帧/s的视频影像,展示长约728 km的环路场景在四季变化下不同的视觉外观。

3.2 试验设计与评价指标

针对Nordland数据集,通过插值GPS数据实现时间同步,保证每个视频中任意帧的车辆位置都能够对应到其他3个视频中的同一帧。

以Pitts30k的训练集作为训练样本,设定10 m以内的图像为正例,25 m以外的图像为负例。训练模型的输入图像组包括基准图像、1个正例和10个负例。

本文以召回率-查准率作为评价指标,与典型的VPR方法进行比较,包括手工聚类特征的算法DBoW2、本文工作的基线方法NetVLAD和同基线方法的先进成

果Patch-NetVLAD^[12]。具体规则如下:

a. 基于Pitts30k数据集的验证,规定如果召回的前 Q 张图像中,至少存在1张图像距离用于测试的基准图像在地理范围的10 m内,则认为图像已被正确检索。

b. 基于Nordland数据集的验证,规定如果召回的前 Q 张图像中,至少存在1张图像距离用于测试的基准图像在10帧内,则认为图像已被正确检索。

c. 试验的评价指标为召回率(Recall)和查准率(Precision):

$$R = T_p / (T_p + F_N) \quad (5)$$

$$P = T_p / (T_p + F_p) \quad (6)$$

式中, R 为召回率; P 为查准率; T_p 为真正例数量; F_p 为假正例数量; F_N 为假负例数量。

3.3 试验结果与讨论

3.3.1 特征降维试验

首先对本文方法在位置识别任务中的召回率进行测试, Q 的不同取值条件下试验结果如表1所示。

表1 本文方法的召回率

| 数据集 | 特征维度 | 召回率/% | | |
|----------|--------|-------|-------|--------|
| | | $Q=1$ | $Q=5$ | $Q=10$ |
| Nordland | 原始特征 | 37.41 | 44.42 | 47.54 |
| | 512维特征 | 43.17 | 47.71 | 50.63 |
| Pitts30k | 原始特征 | 81.63 | 91.16 | 93.22 |
| | 512维特征 | 86.73 | 94.18 | 94.93 |

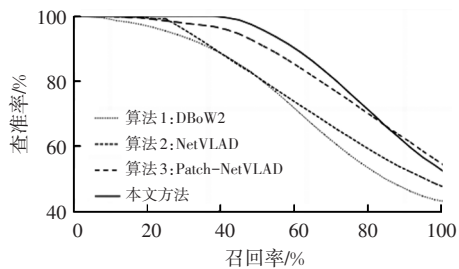
相比原始特征向量,经过主成分分析(Principal Components Analysis, PCA)降维的特征向量在位置识别任务中的召回率表现更好。在Nordland数据集上的平均召回率提高4.05个百分点,在Pitts30k数据集上的平均召回率提高3.28个百分点。提高平均召回率的主要原因是降维后的图像特征忽略了没有辨识度的特征区域,突出了图像中具有区分度的部分,产生更好的召回结果。

3.3.2 位置识别试验

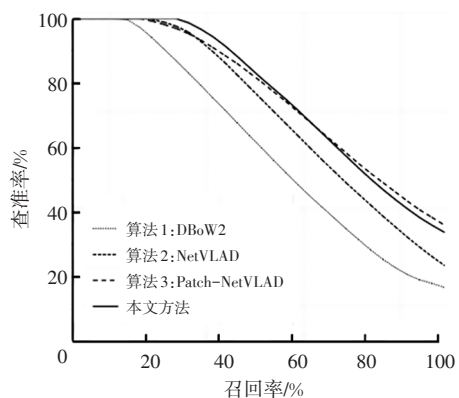
为验证本文改进方法的有效性,将本文方法与近年来具有代表性的VPR方法和同基线的方法进行对比。对不同方法得到的特征向量均降维到512维,使用召回率-查准率评价指标进行分析。本文方法和对比方法在Pitts30k数据集和Nordland数据集上的试验结果如图7所示。

在侧重视角变化的Pitts30k数据集上,NetVLAD算法在召回率为30%时开始出现精度下降,DBoW2和Patch-NetVLAD算法在召回率为40%时开始出现精度的快速下降,本文方法在召回率为50%时才出现明显的精度下降。与DBoW2和NetVLAD算法相比,本文方法

更快地检索到了所有正例图像;仅在召回率达到80%后,本文方法的召回精度与Patch-NetVLAD算法相比略有不足。



(a)Pitts30k数据集



(b)Nordland数据集

图7 不同方法在2种数据集上位置识别任务中的召回率-查准率曲线

在具有季节变化和光照变化的Nordland数据集上,本文方法始终优于DBoW2和NetVLAD算法,与先进的Patch-NetVLAD算法相比,在召回率达到70%前,本文方法也保持着更好的召回精度。

可以看出,相比DBoW2和NetVLAD算法,本文方法在试验中始终保持着更好的召回精度。这是由于融合的坐标注意力能够记录明显的空间特征位置关系,在视角变化的情况下,特征间的相对位置关系保持稳定,而不同视野特征图的组合编码能够覆盖不同视距下的观测结果。同时也可以观察到,本文方法在高召回率的场景中召回精度相比同基线的Patch-NetVLAD算法略低。这是由于Patch-NetVLAD算法采用基于区块尺度的图像特征和基于一致性评分的重新排序策略实现了召回精度的提升。相应地,此策略降低了Patch-NetVLAD算法的检索速度,因此本文方法的检索速度与Patch-NetVLAD算法相比有19%的提升。视觉SLAM的实际测试中,常常需要快速实现场景重识别和回环检测,很少产生需要全部召回的现象,因此本文方法的实用性更强。

经上述试验验证,本文方法在召回率-查准率的评

价体系下明显优于DBoW2和NetVLAD算法,与先进的Patch-NetVLAD相比,本文方法牺牲较少的召回精度获得了检索速度的大幅提升,证明本文改进方法的有效性。

3.3.3 回环检测试验

回环检测是视觉定位中校正全局地图累积误差的重要手段。为验证本文方法在回环检测任务中的表现,设计一组基于旋风智能车^[13]试验平台的应用测试。试验平台如图8所示,其处理器为Jetson TX2,包含6核CPU、256核帕斯卡(Pascal)架构GPU和8 GB内存。

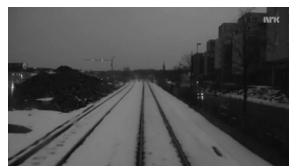


图8 旋风智能车试验平台

利用ORB-SLAM2^[14]算法在Nordland数据集的原始影像上提取用于建立回环地图的关键帧序列,以关键帧序列为数据集进行回环测试。将本文方法与DBoW2、NetVLAD和Patch-NetVLAD算法在回环检测任务中的表现进行对比,试验结果如图9所示。



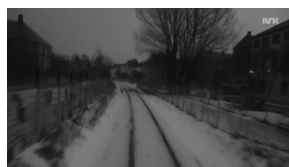
(a)基准图像



(b)DBoW2算法试验结果



(c)NetVLAD算法试验结果



(d)Patch-NetVLAD算法试验结果



(e)本文方法试验结果

图9 季节变化下的回环检测

由图9可以看出,Nordland数据集夏季序列和冬季序列之间存在较大的外观差异。相比DBoW2和NetVLAD算法,本文方法和Patch-NetVLAD算法用于回环检测任务的结果更接近真实的回环位置,进一步说明

本文改进方法的鲁棒性。统计几种方法在回环检测任务中平均闭环准确率和检测时间,结果如表2所示。

表2 对比方法在回环检测任务中的性能

| 算法 | 平均闭环准确率/% | 平均闭环检测时间/ms |
|---------------|-----------|-------------|
| DBoW2 | 71.34 | 2 634 |
| NetVLAD | 82.66 | 2 753 |
| Patch-NetVLAD | 90.11 | 4 746 |
| 本文方法 | 87.42 | 2 832 |

由表2可知:相比DBoW2和NetVLAD算法,本文方法在回环检测任务中的平均闭环准确率有所提高;与Patch-NetVLAD相比,本文方法的平均闭环准确率略低,但是在时间性能上具有明显优势。

3.4 消融试验

为探究本文所提出的改进点对试验结果的影响,在Pitts30k数据集上通过消融试验对两方面的改进进行验证,如表3所示。

表3 消融试验结果

| 网络结构 | 平均查准率/% | 检索时间/ms |
|----------------------|---------|---------|
| ResNet+NetVLAD(基线) | 90.18 | 30.7 |
| ResNet+CA+NetVLAD | 94.05 | 33.8 |
| ResNet+DC-NetVLAD | 93.52 | 34.4 |
| ResNet+CA+DC-NetVLAD | 96.22 | 37.5 |

相比基线方法,本文改进方法获得了更好的平均精度,检索时间仅延长6.8 ms。其中,融合CA对网络的精度提升3.87百分点,使用DC-NetVLAD编码器对网络的精度提升3.34百分点。通过融合CA使特征间的位置关系得到利用,利用并联的DC-NetVLAD编码器发挥不同视野下特征图信息的互补性。最终的试验结果表明,同时采用2个模块对网络进行改进,平均精度提升6.04百分点,进一步提升了网络的位置识别能力,证明本文的2个改进点具有一定的互补性。

4 结束语

针对开放场景中的自动驾驶汽车的视觉定位和位置识别问题,本文在现有方法的基础上提出两点改进:基于残差网络融合坐标注意力提取图像特征,提高原始网络对图像特征位置关系的捕获能力;通过多重扩张卷积和NetVLAD实现轻量的多尺度特征融合,编码图像的特征向量。试验结果表明:在季节、光照和视角等变化导致环境外观发生改变的场景中,本文提出的方法取得了96.22%的召回精度,单帧检索时间仅需37.5 ms,与同基线方法相比,显著提高了检索速度和召回精度。

针对视觉传感器捕获的信息模态较为单一,在视觉受限的场景容易产生系统失效的问题,后续将开展多源传感器融合定位的研究以及通过云端计算的方式缓解边缘设备的算力问题等相关工作,实现自动驾驶车辆的高精度定位。

参 考 文 献

- [1] 高翔,张涛,刘毅,等.视觉SLAM十四讲:从理论到实践[M].2版.北京:电子工业出版社,2019.
GAO X, ZHANG T, LIU Y, et al. 14 Lectures on Visual SLAM: From Theory to Practice[M]. Beijing: Publishing House of Electronics Industry, 2019.
- [2] LOWRY S, SÜNDERHAUF N, NEWMAN P, et al. Visual Place Recognition: A Survey[J]. IEEE Transactions on Robotics, 2016, 32(1): 1-19.
- [3] 刘强,段富海,桑勇,等.复杂环境下视觉SLAM闭环检测方法综述[J].机器人,2019,41(1):112-123+136.
LIU Q, DUAN F H, SANG Y, et al. A Survey of Loop-Closure Detection Method of Visual SLAM in Complex Environments[J]. Robot, 2019, 41(1): 112-123+136.
- [4] 李兆凯,李龙勇,李泽晖,等.基于RGB-D SLAM的智能车自主避障与路径规划试验研究[J].汽车技术,2021(9):55-62.
LI Z K, LI L Y, LI Z H, et al. Experimental Research on Autonomous Obstacle Avoidance and Path Planning of Intelligent Vehicle Based on RGB-D SLAM Technology[J]. Automobile Technology, 2021(9): 55-62.
- [5] 陈涛,范林坤,李旭川,等.基于深度学习的智能车辆视觉里程计技术发展综述[J].汽车技术,2021(1):1-10.
CHEN T, FAN L K, LI X C, et al. Review on the Development of Deep Learning-Based Vision Odometer Technologies for Intelligent Vehicles[J]. Automobile Technology, 2021(1): 1-10.
- [6] 段立娟,孙启超,乔元华,等.基于注意力感知和语义感知的RGB-D室内图像语义分割算法[J].计算机学报,2021,44(2):275-291.
DUAN L J, SUN Q C, QIAO Y H, et al. Attention-Aware and Semantic-Aware Network for RGB-D Indoor Semantic Segmentation[J]. Chinese Journal of Computers, 2021, 44(2): 275-291.
- [7] HOU Q B, ZHOU D Q, FENG J S. Coordinate Attention for Efficient Mobile Network Design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN, USA: IEEE, 2021: 13713-13722.
- [8] 李琳辉,张溪桐,连静,等.结合道路结构化特征的语义SLAM算法[J].哈尔滨工业大学学报,2021,53(2):175-183.
LI L H, ZHANG X T, LIAN J, et al. Semantic SLAM
汽车技术

- Algorithm Combined with Road Structured Features[J]. Journal of Harbin Institute of Technology, 2021, 53(2): 175-183.
- [9] ARANDJELOVIC R, GRONAT P, TORII A, et al. NetVLAD: CNN Architecture for Weakly Supervised Place Recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 5297-5307.
- [10] 姚足, 龚勋, 陈锐, 等. 面向行人重识别的局部特征研究进展、挑战与展望[J]. 自动化学报, 2021, 47(12): 2742-2760.
YAO Z, GONG X, CHEN R, et al. Research Progress, Challenge and Prospect of Local Features for Person Re-Identification[J]. Acta Automatica Sinica, 2021, 47(12): 2742-2760.
- [11] SÜNDERHAUF N, NEUBERT P, PROTZEL P. Are We There Yet? Challenging SeqSLAM on a 3000 km Journey Across all Four Seasons[C]// Proceedings of Workshop on Long-Term Autonomy, IEEE International Conference on Robotics and Automation (ICRA). Karlsruhe, Germany: IEEE, 2013.
- [12] HAUSLER S, GARG S, XU M, et al. Patch-Netvlad: Multi-Scale Fusion of Locally-Global Descriptors for Place Recognition[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN, USA: IEEE, 2021: 14141-14152.
- [13] 常飞翔, 刘元盛, 李中道, 等. 一种基于TX2平台的改进YOLOv3红绿灯检测算法[J]. 汽车技术, 2021(8): 1-6.
CHANG F X, LIU Y S, LI Z D, et al. An Improved YOLOv3 Traffic Light Detection Algorithm Based on TX2 Platform[J]. Automobile Technology, 2021(8): 1-6.
- [14] MURARTAL R, TARDOS J D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras[J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.

(责任编辑 斛 畔)

修改稿收到日期为2022年7月19日。

《汽车工艺与材料》投稿须知

《汽车工艺与材料》杂志创刊于1986年,是由中国第一汽车集团有限公司主办的国内外公开发行的汽车制造类月刊、中国汽车工程学会会刊,是全面报道汽车先进制造工艺与材料应用的专业性杂志。

《汽车工艺与材料》杂志始终致力于汽车制造新技术的开发(如汽车制造领域中的四大工艺)、新材料的研究应用、轻量化技术、智能制造、绿色制造技术、加工技术、生产装备、检测技术、数字化模拟技术及生产管理等方面的报道。努力为汽车及零部件生产中各门类的决策者及研究开发人员提供更好、更快、更节约成本的汽车行业发展动态信息,为其适应国内与国际激烈的汽车生产市场竞争提供有效的帮助。

《汽车工艺与材料》杂志关注领先的整车及零部件企业和材料、装备等供应商,及时报道汽车行业最新的设计、制造、材料、加工技术、生产装备、检测技术等方面的成功案例。

主要栏目:

AT&M视界、生产现场、材料应用、生产装备、检测技术、数字化园地、行业动态等。

投稿要求:

(1)来稿须具有独创性并与实践相结合,文章字数最好控制在6000字以内。

(2)来稿不能在国内、外公开杂志上发表过,严禁一稿多投,对同时一稿多投者,编辑部将视对本刊声誉、工作上造成的影响程度,采取相应处理措施。

(3)来稿的试验方法、试验数据、试验结论必须准确、可靠。

(4)来稿须包括以下项目:题名、作者姓名、作者单位、中文摘要(200字左右)、参考文献等。来稿采用word文档的格式。

(5)来稿文章格式应符合一般科技论文格式,或参考近期本刊所刊登文章格式。

(6)稿件图片请发清晰原图,图形尺寸应适中,并配上图名和图号,且文中要有相应体现(例如:如图1所示、见图2等)。图中英文需要翻译成中文。

(7)计量单位必须采用(或换算成)中华人民共和国法定计量单位。文章中文字字母大小写、上下角标应书写正确,避免同一符号代表2种及以上物理量的情况出现。

(8)文章必须附有公开发表的参考文献,且在文中应标注文献引用处。参考文献注册项目应齐全,包括作者姓名、题名、刊名(或书名,图书应注明出版地与出版社)、出版年、卷、期。

(9)请在来稿时注明作者联系方式(包含手机、办公室电话、E-mail等)。

(10)本刊投稿网址:<http://qcgycj.cbpt.cnki.net>,咨询电话:0431-82028069。

竭诚欢迎汽车行业及相关各界的专家学者积极向本刊投稿。