



(12) 发明专利申请

(10) 申请公布号 CN 111832484 A

(43) 申请公布日 2020.10.27

(21) 申请号 202010675615.X

G06N 3/04 (2006.01)

(22) 申请日 2020.07.14

G06N 3/08 (2006.01)

(71) 申请人 星际(重庆)智能装备技术研究院有限公司

地址 400030 重庆市沙坪坝区凤凰镇皂角树村临谢家院子组2号1-14室

(72) 发明人 宋永端 方磊 赖俊峰 谭力琿 张杰 譙佩雯 张聪毅 包晗 杨魁龙 杨雅婷 陈颖睿

(74) 专利代理机构 重庆信航知识产权代理有限公司 50218

代理人 吴彬

(51) Int.Cl.

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

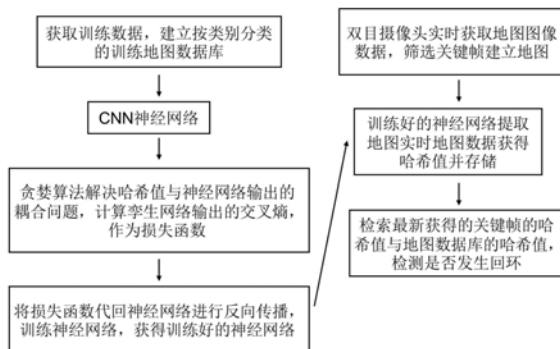
权利要求书1页 说明书5页 附图2页

(54) 发明名称

一种基于卷积感知哈希算法的回环检测方法

(57) 摘要

本发明公开了一种基于卷积感知哈希算法的回环检测方法,其包括构建建立地图所需场景的图像数据集,构建并训练卷积神经网络,将训练好的卷积神经网络模型应用于SLAM系统进行回环检测。本发明基于卷积感知哈希算法的回环检测方法通过贪婪算法解决卷积神经网络的输出与哈希值之间的耦合问题,使用孪生网络输出的交叉熵作为损失函数训练卷积神经网络获得训练好的卷积神经网络。本发明具有较高的准确率和召回率,根据不同场景训练的网络,在应对不同场景时具备优秀的适应能力,能够很好的完成回环检测的功能,辅助SLAM构建正确的高精度地图。



1. 一种基于卷积感知哈希算法的回环检测方法,其特征在于,包括以下步骤:

1) 构建建立地图所需场景的图像数据集,图像数据集由两组图片组成,第一组图片中全为相似的图片,第二组图片中全为不相似的图片,将图像数据集划分为训练集和测试集,训练集中包含第一组图片的部分和第二组图片的部分,训练集用于卷积神经网络的训练,测试集用于检测训练好的神经网络是否有效;

2) 构建并训练卷积神经网络,其包括以下步骤:

a) 构建两个结构相同的卷积神经网络模块,每个卷积神经网络模块包括卷积层、池化层以及全连接层;

b) 从训练集的两组图片中随机选取两张图片构成一组输入图片同时输入所构建的两个卷积神经网络模块,通过贪婪算法分别将两个卷积神经网络模块的输出变换成两个哈希码,卷积神经网络模块的输出和哈希码的连接方式如下:

$$\begin{cases} H^{t+1} = \text{sgn}(O^{t+1}) \\ O^{t+1} = H^t - lr * \frac{\partial L}{\partial H^t} \end{cases}$$

其中,H是哈希码,O是卷积神经网络的输出,lr是超参数,L是交叉熵;再计算两张图像的最大似然估计,两张图像对应的两个哈希码的最大似然函数如下:

$$\log L(S_{ij}|H) = \sum_{ij}^n \log L(s_{ij}|h_i, h_j)$$

通过最大似然函数计算损失函数,损失函数设置如下:

$$\min \sum_{i,j=0}^n \log(1 + \exp(\alpha(h_i, h_j))) - \alpha s_{i,j} \langle h_i, h_j \rangle$$

其中,n是指哈希码的位数,i,j分别指两张图片, α 是超参数, h_i 是第i张图像对应的哈希码, h_j 是第j张图像对应的哈希码, $s_{i,j}$ 指输入的两张图片是否归属为一类图片;

$$\begin{cases} s_{i,j} = 1, & \text{相似} \\ s_{i,j} = 0 & \text{不相似} \end{cases}$$

c) 将损失函数代回卷积神经网络进行反向传播,训练卷积神经网络并对训练好的神经网络进行测试,得到训练好的卷积神经网络模型;

3) 将训练好的卷积神经网络模型应用于SLAM系统,不断的将构建地图的关键帧输入训练好的两个卷积神经网络模型中的其中一个,并计算关键帧的哈希码,再将计算得到哈希码的关键帧存储无序图中;

4) 回环检测:

将当前获得的关键帧的哈希码和之前存储的每个哈希码根据如下公式分别计算汉明距离:

$$\text{Dist}_H(h_i, h_j) = \sum_{i,j}^n h_i \oplus h_j$$

其中 h_i, h_j 表示两幅图的哈希码, \oplus 表示异或操作;将计算得到的最小汉明距离与设定的阈值进行比较,若小于设定的阈值,则认为两帧图像为相似场景或者为同一场景,发生回环。

一种基于卷积感知哈希算法的回环检测方法

技术领域

[0001] 本发明涉及即时定位与建图 (SLAM) 中的回环检测技术领域,更具体地,涉及一种基于深度卷积感知哈希的回环检测方法。

背景技术

[0002] 在视觉SLAM(simultaneous localization and mapping,即时定位与地图构建)中,视觉里程计仅考虑相邻时间上的关键帧,如果上一时刻产生误差,那么误差必定会累积到下一时刻,使得整个SLAM出现累积误差,长期的估计结果必然是不可靠的。SLAM后端虽然可以估计最大后验误差,但在只存在相邻关键帧数据时,同样无法消除累计误差。

[0003] 回环检测的意义就是给出除了相邻关键帧的约束外,时间更久远的约束。当机器人察觉相机经过了同一个地方,采集到了相似的数据这一情况,我们就可以提供更好的估计,得到一个全局已知的估计。回环检测关系到SLAM估计的轨迹和地图在长时间下的正确性,对机器人进行了重定位,能够非常有效的提升SLAM建图的精度。回环检测目前采用的技术一般为词袋法,通过识别图片上的特征并保存为单词,通过比较两帧图片上的特征数量来获得两帧图像的相似度,当相似度达到一定阈值,认为两帧是同一地点,认为发生回环。

[0004] 词袋的回环检测算法完全依赖于外观,没有利用图中任何的几何信息,这容易导致外观相似的图像容易被当成回环。并且,由于词袋不在乎单词顺序,只在意单词有无的表达方式,更容易引发感知偏差。

发明内容

[0005] 本发明的目的在于提供一种能够克服上述技术问题的卷积感知哈希算法的回环检测方法,本发明所述方法的特征在于基于深度学习技术,使用一种图像编码的技术来解决回环检测精度低、效率低的问题。

[0006] 本发明基于卷积感知哈希算法的回环检测方法,包括以下步骤:

[0007] 1) 构建建立地图所需场景的图像数据集,图像数据集由两组图片组成,第一组图片中全为相似的图片,第二组图片中全为不相似的图片,将图像数据集划分为训练集和测试集,训练集中包含第一组图片的部分和第二组图片的部分,训练集用于卷积神经网络的训练,测试集用于检测训练好的神经网络是否有效;

[0008] 2) 构建并训练卷积神经网络,其包括以下步骤:

[0009] a) 构建两个结构相同的卷积神经网络模块,每个卷积神经网络模块包括卷积层、池化层以及全连接层;

[0010] b) 从训练集的两组图片中随机选取两张图片构成一组输入图片同时输入所构建的两个卷积神经网络模块,通过贪婪算法分别将两个卷积神经网络模块的输出变换成两个哈希码,卷积神经网络模块的输出和哈希码的连接方式如下:

$$[0011] \quad \begin{cases} H^{t+1} = \text{sgn}(O^{t+1}) \\ O^{t+1} = H^t - lr * \frac{\partial L}{\partial H^t} \end{cases}$$

[0012] 其中,H是哈希码,O是卷积神经网络的输出,lr是超参数,L是交叉熵;

[0013] 再计算两张图像的最大似然估计,两张图像对应的两个哈希码的最大似然函数如下:

$$[0014] \quad \log L(S_{ij}|H) = \sum_{ij} \log L(s_{ij}|h_i, h_j)$$

[0015] 通过最大似然函数计算损失函数,损失函数设置如下:

$$[0016] \quad \min \sum_{i,j=0}^n \log(1 + \exp(\alpha \langle h_i, h_j \rangle)) - \alpha s_{ij} \langle h_i, h_j \rangle$$

[0017] 其中,n是指哈希码的位数,i,j分别指两张图片, α 是超参数, h_i 是第i张图像对应的哈希码, h_j 是第j张图像对应的哈希码, $s_{i,j}$ 指输入的两张图片是否归属为一类图片;

$$[0018] \quad \begin{cases} s_{i,j} = 1, & \text{相似} \\ s_{i,j} = 0 & \text{不相似} \end{cases}$$

[0019] c) 将损失函数代回卷积神经网络进行反向传播,训练卷积神经网络并对训练好的神经网络进行测试,得到训练好的卷积神经网络模型;

[0020] 3) 将训练好的卷积神经网络模型应用于SLAM系统,不断的将构建地图的关键帧输入训练好的两个卷积神经网络模型中的其中一个,并计算关键帧的哈希码,再将计算得到哈希码的关键帧存储无序图中;

[0021] 4) 回环检测:

[0022] 将当前获得的关键帧的哈希码和之前存储的每个哈希码根据如下公式分别计算汉明距离:

$$[0023] \quad \text{Dist}_H(h_i, h_j) = \sum_{i,j} h_i \oplus h_j$$

[0024] 其中 h_i, h_j 表示两幅图的哈希码, \oplus 表示异或操作;将计算得到的最小汉明距离与设定的阈值进行比较,若小于设定的阈值,则认为两帧图像为相似场景或者为同一场景,发生回环。

[0025] 本发明的有益效果:

[0026] 本发明基于卷积感知哈希算法的回环检测方法,通过贪婪算法解决卷积神经网络的输出与哈希值之间的耦合问题,使用孪生网络输出的交叉熵作为损失函数训练卷积神经网络获得训练好的卷积神经网络。通过视觉SLAM实时建立地图,并存储关键帧作为地图数据库,将关键帧输入训练好的卷积神经网络,计算获得对应关键帧的哈希值,与地图数据库进行对比,检测是否发生回环。本发明具有较高的准确率和召回率,根据不同场景训练的网络,在应对不同场景时具备优秀的适应能力,能够很好的完成回环检测的功能,辅助SLAM构建正确的高精度地图。

附图说明

- [0027] 图1是基于卷积感知哈希算法的回环检测方法的流程图。
- [0028] 图2是卷积神经网络模块的结构图。
- [0029] 图3为卷积网络哈希算法的学习流程示意图。
- [0030] 图4为回环检测的流程示意图。

具体实施方式

[0031] 下面结合附图对本发明的实施方式进行详细描述。

[0032] 本实施例中基于卷积感知哈希算法的回环检测方法,包括以下步骤:

[0033] 1) 构建建立地图所需场景的图像数据集,图像数据集由两组图片组成,第一组图片中全为相似的图片,第二组图片中全为不相似的图片,将图像数据集划分为训练集和测试集,训练集中包含第一组图片的部分和第二组图片的部分,训练集用于卷积神经网络的训练,测试集用于检测训练好的神经网络是否有效。

[0034] 本实施例中图像数据集包含2万个类别,每个类中包含类似的图片。将数据集按照8:2的比例分成训练集和测试集,训练集用于卷积神经网络的训练,测试集用于检测训练好的神经网络是否有效。

[0035] 2) 构建并训练卷积神经网络,其包括以下步骤:

[0036] a) 构建两个结构相同的卷积神经网络模块,每个卷积神经网络模块包括卷积层、池化层以及全连接层。

[0037] 具体实施中可随场景复杂程度改变卷积层结构的卷积核大小和卷积层数量。卷积神经网络模块输入的图像尺寸为224x224x3像素的图像,可根据场景的不同人为选取5层至22层卷积层。高阶池化层均设计为均方差池化层,将卷积层输出的特征转换成样本协方差矩阵。

[0038] 在场景简单,特征少的情况下可构建5层卷积网络,选择11x11的卷积核;在场景复杂,特征较多的情况下,构建深层卷积网络,随层数的加深,增加3x3卷积核的比例,提取更多的特征。同时采取“dropout”技术,以0.2的概率将每个隐层神经元的输出设置为0,避免过拟合。

[0039] 本实施例中,每个卷积神经网络模块包括5层,第一个稠密块采用6个1x1卷积核和6个3x3卷积核,稠密块的内部每一层网络的输出由前一层网络的直接输出和经过卷积核卷积的输出组成,每个稠密块内部都经历这种运算。第一个稠密块的输出尺寸为56x56,经过一层池化层,输出尺寸变为28x28。第二个稠密块采用12个1x1的卷积核和12个3x3的卷积核输出,稠密块内部同样经历上述输出。第二个稠密块的输出尺寸为28x28,经过池化层变为14x14。第三个稠密块采用24个1x1的卷积核和24个3x3的卷积核,经历同样的操作输出7x7,第四个稠密块采用16个1x1的卷积核和16个3x3的卷积核,最终与全连接层连接。

[0040] b) 从训练集的两组图片中随机选取两张图片构成一组输入图片同时输入所构建的两个卷积神经网络模块,通过贪婪算法分别将两个卷积神经网络模块的输出变换成两个哈希码,卷积神经网络模块的输出和哈希码的连接方式如下:

$$[0041] \quad \begin{cases} H^{t+1} = \text{sgn}(O^{t+1}) \\ O^{t+1} = H^t - lr * \frac{\partial L}{\partial H^t} \end{cases}$$

[0042] 其中,H是哈希码,O是卷积神经网络的输出,lr是超参数,L是交叉熵;最终输出为哈希码。

[0043] 再计算两张图像的最大似然估计,两张图像对应的两个哈希码的最大似然函数如下:

$$[0044] \quad \log L(S_{ij}|H) = \sum_{ij}^n \log L(s_{ij}|h_i, h_j)$$

[0045] 通过最大似然函数计算损失函数,损失函数设置如下:

$$[0046] \quad \min \sum_{i,j=0}^n \log(1 + \exp(\alpha \langle h_i, h_j \rangle)) - \alpha s_{ij} \langle h_i, h_j \rangle$$

[0047] 其中,n是指哈希码的位数,i,j分别指两张图片, α 是超参数, h_i 是第i张图像对应的哈希码, h_j 是第j张图像对应的哈希码, $s_{i,j}$ 指输入的两张图片是否归属为一类图片;

$$[0048] \quad \begin{cases} s_{i,j} = 1, & \text{相似} \\ s_{i,j} = 0 & \text{不相似} \end{cases}$$

[0049] c) 将损失函数代回卷积神经网络进行反向传播,训练卷积神经网络并对训练好的神经网络进行测试,得到训练好的卷积神经网络模型。

[0050] 当卷积网络模块的输出使得相似的图片的哈希码的汉明距离足够小,不相似的图片的汉明距离足够大时证明网络训练完毕,保存网络模型。将测试集的图片输入卷积网络,观察结果,是否依旧能保持相似的图片的哈希码的汉明距离足够小,不相似的图片的哈希码的汉明距离足够大,如果能够实现,说明模型训练完成,保存模型。

[0051] 3) 将训练好的卷积神经网络模型应用于SLAM系统,不断的将构建地图的关键帧输入训练好的两个卷积神经网络模型中的其中一个,并计算关键帧的哈希码,再将计算得到哈希码的关键帧存储无序图中。

[0052] 具体的,启动机器人开始构建SLAM地图,机器人的双目摄像头左右相机拍摄照片,根据一定条件适当的选取一些图片作为构建地图的关键帧。将构建地图的关键帧输入训练好的一个卷积网络模块,计算得到关键帧的哈希码,并将计算得到哈希码的关键帧存储在无序图中,无序图是一种根据哈希码排序的容器。

[0053] 在机器人不断的运行构建地图的过程中,如果经过以前经过的场景,这时拍摄的照片同样会被根据一定的条件选取为关键帧。

[0054] 4) 回环检测:

[0055] 将当前获得的关键帧的哈希码和之前存储的每个哈希码根据如下公式分别计算汉明距离:

$$[0056] \quad \text{Dist}_H(h_i, h_j) = \sum_{i,j}^n h_i \oplus h_j$$

[0057] 其中 h_i, h_j 表示两幅图的哈希码, \oplus 表示异或操作;将计算得到的最小汉明距离与设定的阈值进行比较,若小于设定的阈值,则认为两帧图像为相似场景或者为同一场景,发

生回环。机器人就可以以此消除累计误差,减少SLAM构建的地图的漂移和尺寸变化。

[0058] 最后说明的是,以上实施例仅用以说明本发明的技术方案而非限制,尽管参照较佳实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,可以对本发明的技术方案进行修改或者等同替换,而不脱离本发明技术方案的宗旨和范围,其均应涵盖在本发明的权利要求范围当中。

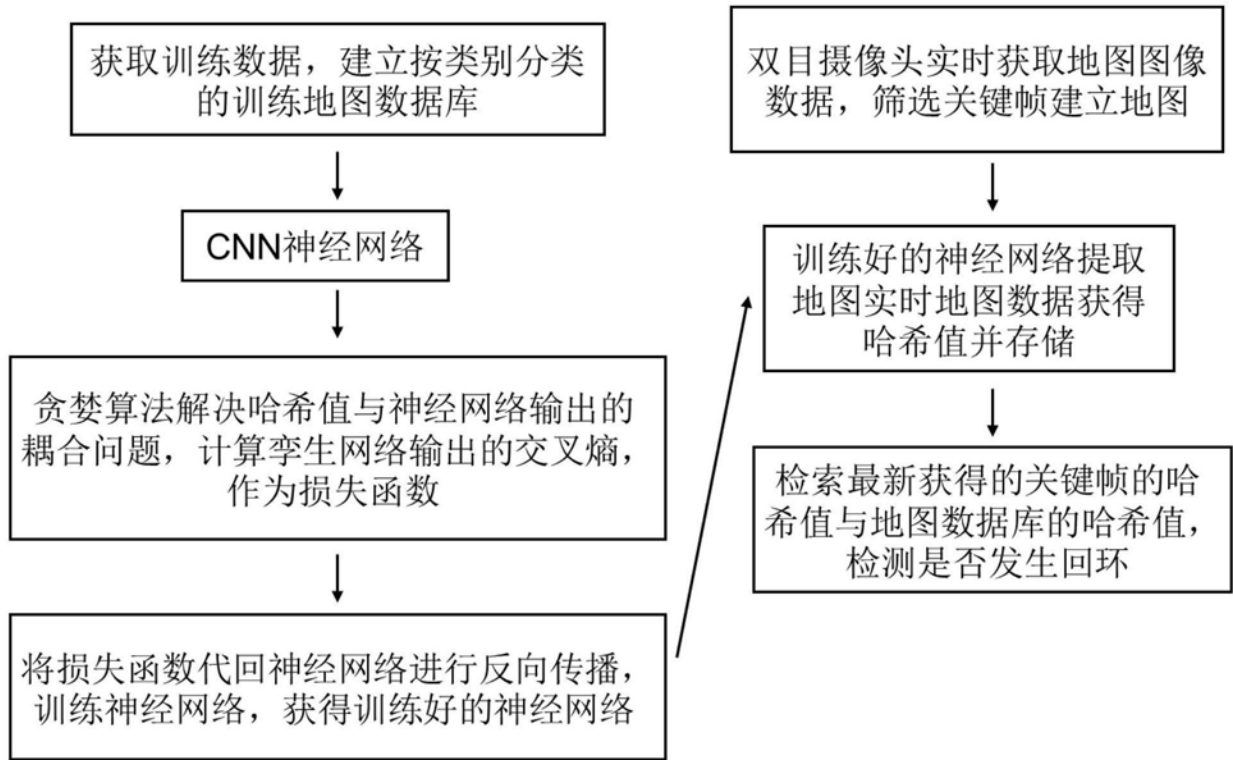


图1

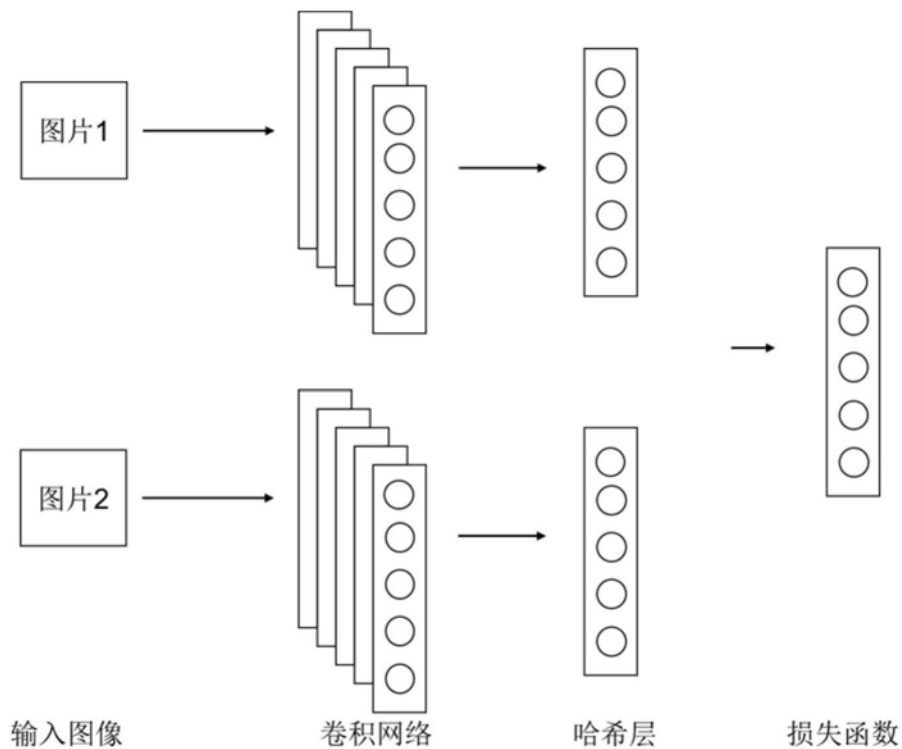


图2

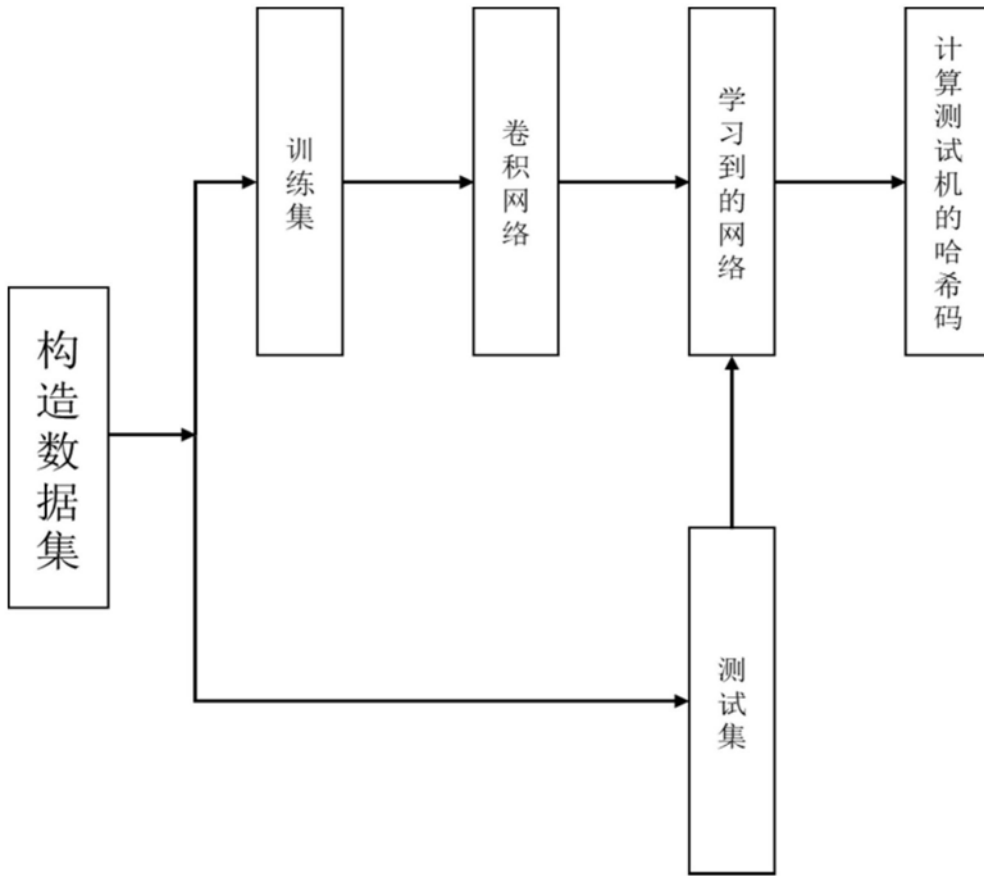


图3

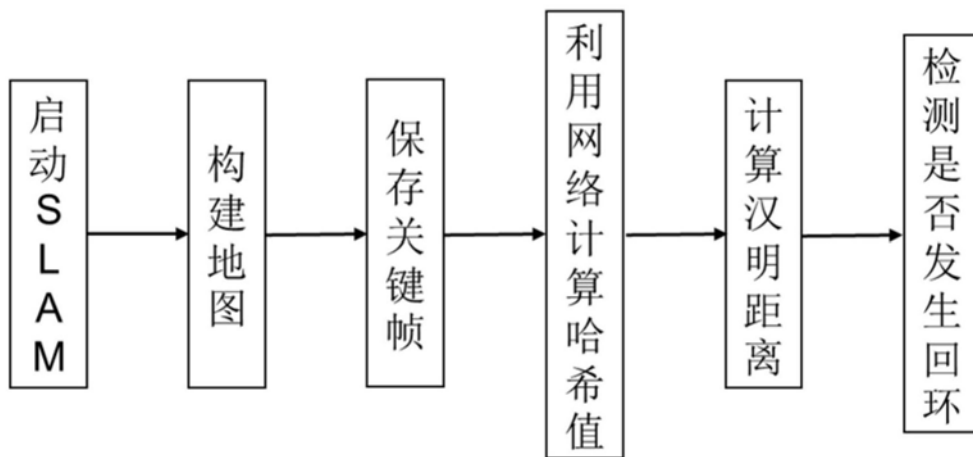


图4