



(12) 发明专利

(10) 授权公告号 CN 108986162 B

(45) 授权公告日 2022.02.22

(21) 申请号 201810686554.X

(22) 申请日 2018.06.28

(65) 同一申请的已公布的文献号  
申请公布号 CN 108986162 A

(43) 申请公布日 2018.12.11

(73) 专利权人 杭州吉吉知识产权运营有限公司  
地址 310000 浙江省杭州市滨江区西兴街  
道江淑路260号10242室

(72) 发明人 王梓里

(74) 专利代理机构 浙江千克知识产权代理有限公司 33246

代理人 雷娴

(51) Int. Cl.

G06T 7/70 (2017.01)

G06T 7/194 (2017.01)

(56) 对比文件

US 2013293532 A1, 2013.11.07

CN 106056664 A, 2016.10.26

CN 105513127 A, 2016.04.20

CN 106371104 A, 2017.02.01

CN 108171796 A, 2018.06.15

CN 104658012 A, 2015.05.27

CN 105045263 A, 2015.11.11

CN 106971403 A, 2017.07.21

CN 104331699 A, 2015.02.04

CN 105184852 A, 2015.12.23

CN 103955920 A, 2014.07.30

CN 107687850 A, 2018.02.13

CN 104143194 A, 2014.11.12

CN 105761242 A, 2016.07.13

CN 107767425 A, 2018.03.06

审查员 向奎

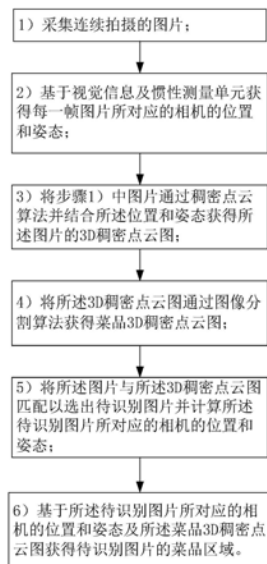
权利要求书1页 说明书5页 附图1页

(54) 发明名称

基于惯性测量单元和视觉信息的菜品和背景分割方法

(57) 摘要

本发明涉及图像处理技术领域,具体为一种基于惯性测量单元和视觉信息的菜品和背景分割方法,包括以下步骤:1)采集连续拍摄的图片;2)基于视觉信息及惯性测量单元获得每一帧图片所对应相机的位置和姿态;3)将步骤1)中图片通过稠密点云算法并结合所述位置和姿态获得所述图片的3D稠密点云图。本发明利用惯性测量单元和视觉信息的融合获得菜品以及周围物体的3D点云,使用菜品和背景的3D点云信息来分割图像中的菜品,相对现有通用算法的优点是不需要事先训练,精度更高,使用场景更广。



1. 基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:包括以下步骤:
  - 1) 采集连续拍摄的图片;
  - 2) 基于视觉信息及惯性测量单元获得每一帧图片所对应的相机的位置和姿态;所述惯性测量单元还获得了所述图片的稀疏点云数据,所述稀疏点云数据辅助计算所述图片中每一帧图片的图片位姿;
  - 3) 将步骤1) 中图片通过稠密点云算法并结合所述位置和姿态获得所述图片的3D稠密点云图;
  - 4) 将所述3D稠密点云图通过图像分割算法获得菜品3D稠密点云图;
  - 5) 将所述图片与所述3D稠密点云图匹配以选出待识别图片并计算所述待识别图片所对应的相机的位置和姿态;
  - 6) 基于所述待识别图片所对应的相机的位置和姿态及所述菜品3D稠密点云图获得待识别图片的菜品区域。
2. 根据权利要求1所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:图片采集频率大于10赫兹。
3. 根据权利要求1所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:图片采集设备在采集所述图片时围绕菜品在经度方向上旋转的角度不小于60度。
4. 根据权利要求1所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:所述步骤2) 具体为基于视觉惯性里程计获得每一帧图片所对应的相机的位置和姿态。
5. 根据权利要求4所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:所述步骤2) 通过视觉惯性里程计还获得了所述图片的稀疏点云数据,所述稀疏点云数据辅助计算所述图片所对应的相机的位置和姿态。
6. 根据权利要求1所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:所述步骤3) 具体为计算不同位置和姿态的相机所对应的图片的对应像素之间的视差,并使用全局优化获得图片的3D稠密点云图。
7. 根据权利要求1所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:所述步骤4) 中的图像分割算法采用高度作为区别特征,并对所述3D稠密点云图进行垂直扫描和聚类以将菜品对应的点云分割出来形成菜品3D稠密点云图。
8. 根据权利要求1所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:所述步骤5) 具体为将所述待识别图片通过步骤3) 中的3D稠密点云获取待识别图片所对应的相机的位置和姿态。
9. 根据权利要求1所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:所述步骤5) 具体为通过查询步骤2) 中的图片所对应的相机的位置和姿态来获取待识别图片所对应的相机的位置和姿态。
10. 根据权利要求1所述的基于惯性测量单元和视觉信息的菜品和背景分割方法,其特征在於:所述步骤6) 具体为将3D稠密点云图在与所述待识别图片所对应的相机的位置和姿态相同的方向上投影获得相应的2D稠密点云图,并将待识别图片与2D稠密点云图匹配以获得待识别图片的菜品区域。

## 基于惯性测量单元和视觉信息的菜品和背景分割方法

### 技术领域

[0001] 本发明涉及图像处理技术领域,具体为一种基于惯性测量单元和视觉信息的菜品和背景分割方法。

### 背景技术

[0002] 菜品的图像识别中,背景图案会对识别产生干扰作用,如果能够事先把菜品从背景中分割出来,可以大大提高菜品识别的鲁棒性。现有技术都是通过深度学习来分割菜品和背景,需要大量的人工标记,需要事先训练模型,计算效率低,使用复杂。

[0003] 专利申请号为CN201610694814.9的发明专利公开了一种图像前背景分割网络模型的训练方法,包括:获取待训练的样本图像的特征向量,其中,所述样本图像为包含有前景标注信息和背景标注信息的样本图像;对所述特征向量进行卷积处理,获取特征向量卷积结果;对所述特征向量卷积结果进行放大处理;判断放大后的所述特征向量卷积结果是否满足收敛条件;若满足,则完成对用于分割图像前背景的卷积神经网络模型的训练;若不满足,则根据放大后的所述特征向量卷积结果调整所述卷积神经网络模型的参数并根据调整后的所述卷积神经网络模型的参数对所述卷积神经网络模型进行迭代训练,直至迭代训练后的特征向量卷积结果满足所述收敛条件。该方法的图像前背景分割网络模型需要大量的人工标记,需要事先训练模型,计算效率低。

[0004] 专利申请号为CN201310687416.0的发明专利公开了一种背景分割方法,所述方法包括:为每个像素建立一个高斯混合模型GMMs;更新所述GMMs模型的充分统计量;基于更新后的充分统计量计算GMMs模型的参数,获得新的GMMs模型;根据所述新的GMMs模型确定当前像素为背景或前景。该背景分割方法需要大量的人工标记,需要事先训练模型,使用复杂。

[0005] 利用视觉信息和惯性测量单元对3D模型的重建算法已进入到应用阶段,使用类似的技术,可以解决很多问题,比如在物体识别方面,不仅可以利用图片的明暗信息,还可以利用图片的深度信息,以达到更好的识别效果。因此,本申请通过视觉信息和惯性测量单元来对图像进行菜品和背景分割。

### 发明内容

[0006] 本发明针对现有技术存在的问题,提出了一种基于惯性测量单元和视觉信息的菜品和背景分割方法,不需要事先训练,精度更高,使用场景更广。

[0007] 本发明解决其技术问题所采用的技术方案是:一种基于惯性测量单元和视觉信息的菜品和背景分割方法,包括以下步骤:

[0008] 1) 采集连续拍摄的图片;

[0009] 2) 基于视觉信息及惯性测量单元获得每一帧图片所对应的相机的位置和姿态;

[0010] 3) 将步骤1)中图片通过稠密点云算法并结合所述位置和姿态获得所述图片的3D稠密点云图;

- [0011] 4) 将所述3D稠密点云图通过图像分割算法获得菜品3D稠密点云图;
- [0012] 5) 将所述图片与所述3D稠密点云图匹配以选出待识别图片并计算所述待识别图片所对应的相机的位置和姿态;
- [0013] 6) 基于所述待识别图片所对应的相机的位置和姿态及所述菜品3D稠密点云图获得待识别图片的菜品区域。
- [0014] 作为优选,图片采集频率大于10赫兹。
- [0015] 作为优选,图片采集设备在采集所述图片时围绕菜品在经度方向上旋转的角度不小于60度。
- [0016] 作为优选,所述步骤2) 具体为基于视觉惯性里程计获得每一帧图片所对应的相机的位置和姿态。
- [0017] 作为优选,所述步骤2) 通过视觉惯性里程计还获得了所述图片的稀疏点云数据,所述稀疏点云数据辅助计算所述图片所对应的相机的位置和姿态。
- [0018] 作为优选,所述步骤3) 具体为计算不同位置和姿态的相机所对应的图片的对应像素之间的视差,并使用全局优化获得图片的3D稠密点云图。
- [0019] 作为优选,所述步骤4) 中的图像分割算法采用高度作为区别特征,并对所述3D稠密点云图进行垂直扫描和聚类以将菜品对应的点云分割出来形成菜品3D稠密点云图。
- [0020] 作为优选,所述步骤5) 具体为将所述待识别图片通过的步骤3) 中的3D稠密点云图获取待识别图片所对应的相机的位置和姿态。
- [0021] 作为优选,所述步骤5) 具体为通过查询步骤2) 中的图片所对应的相机的位置和姿态来获取待识别图片所对应的相机的位置和姿态。
- [0022] 作为优选,所述步骤6) 具体为将3D稠密点云图在与所述待识别图片所对应的相机的位置和姿态相同的方向上投影获得相应的2D稠密点云图,并将待识别图片与2D稠密点云图匹配以获得待识别图片的菜品区域。
- [0023] 本发明的有益效果是,本发明利用惯性测量单元和视觉信息的融合来获得菜品以及周围物体的3D点云,使用菜品和背景的3D点云信息来分割图像中的菜品,相对现有通用算法的优点是不需要事先训练,精度更高,使用场景更广。

## 附图说明

- [0024] 图1为本发明一种基于惯性测量单元和视觉信息的菜品和背景分割方法的流程图。

## 具体实施方式

- [0025] 以下是本发明的具体实施例并结合附图,对本发明的技术方案作进一步的描述,但本发明并不限于这些实施例。
- [0026] 如图1所示,一种基于惯性测量单元和视觉信息的菜品和背景分割方法,包括以下步骤:
- [0027] 1) 采集连续拍摄的图片。用户可以通过手机、相机等装有摄像头的设备拍摄图片。用户在拍摄图片时,需要从俯视的角度围绕菜品旋转,图片采集设备在采集所述图片时围绕菜品在经度方向上旋转的角度不小于60度,否则采集的图片不能够获取有效的3D稠密点云

图,在纬度方向上不做要求,可以是任意值。图片采集频率需要大于10赫兹,否则采集的图片同样不能够获取有效的3D稠密点云图。图片拍摄的时间可以根据图片采集的频率决定,当图片采集频率高时,拍摄的时间可以相对较短,如4至8秒钟。当图片采集频率低时,拍摄的时间可以相对较长,如9至12秒钟。

[0028] 2) 基于视觉信息及惯性测量单元获得每一帧图片所对应的相机的位置和姿态。具体为基于视觉惯性里程计获得每一帧图片所对应的相机的位置和姿态。视觉惯性里程计包括惯性测量单元,所述惯性测量单元所提供加速度值和角速度值的频率需要大于100赫兹,否则同样不能够获取有效的3D稠密点云图。目前视觉惯性里程计(VIO)是主流的基于惯性测量单元和视觉信息来计算摄像头位置和姿态的方法。VIO全称是:Visual-Inertial Odometry,已经有很多成熟的算法。视觉和惯性测量单元融合目前主要基于三类方法。

[0029] 基于滤波的方法:扩展卡尔曼滤波器(EKF)框架通常由预测步骤和更新步骤组成。对于基于滤波的VIO方法,惯性传感器测量得到的三轴加速度和转动速度,可用于计算3D刚体运动的数据驱动动力学模型或先验分布,并且可在预测步骤中用来进行运动预测。相机能够提供特征和移动平台之间的角度和距离测量,可用作测量模型或似然分布,并用于在更新步骤中更新预测。状态向量只包含当前状态,由于线性化误差和计算能力的限制,通常只能构建很少的landmark/mappoint,或者创建structureless的状态向量(将landmark/mappoint边缘化),典型代表为MSCKF。边缘化是将旧状态融合进当前状态的先验中,滤波方法主要的缺陷也就存在于边缘化过程中,首先当前测量的structure信息需要延迟处理,降低当前状态的更新精度,其次边缘化对线性化近似和外点(outlier)都比较敏感,容易造成滤波器状态失准。

[0030] 基于优化的方法:基于优化的方法主要依赖于特征提取的图像处理技术和图像配准的最优化,而惯性测量被视为先验,可当作正则项或完全忽略。在大多数情况下,基于优化的方法有两个阶段:构图和跟踪。在构图阶段,由各种特征检测器从图像中提取特征,如3D空间中的角点,边缘或其他地标。然后,对于检测到的所有特征,定义两帧图像之间的重投影误差。该误差用作要优化的损失函数,以便找到特征或地标的坐标。在跟踪阶段中,地图中特征和地标的坐标被用于定义两帧图像之间的重投影误差,并且再次应用优化算法以计算出移动平台的位置和方向的变化。将估计问题分为两个阶段的想法可以获得快速的跟踪结果,然而构图处理还是耗时的。引入关键帧的概念能够边缘化旧的状态而保持一个有界优化窗口,保证实时性。

[0031] Fixed-Lag Smoothing,也就是滑动窗口优化(Sliding Window Optimization),状态向量包含随时间滑动的窗口内多个状态,但是也需要将旧状态边缘化到高斯先验中,因此在滤波算法中存在的边缘化问题,这里都存在。但是由于采用了多个状态的窗口,状态估计更精确,同时可以建立鲁棒的优化代价函数,降低外点对状态估计的影响。状态向量中可以加入测量的structure,但是太多structure会造成高斯先验矩阵是稠密的,通常会想办法减少structure的数量。

[0032] 基于滤波方法和优化方法的结合:基于滤波和基于优化的方法都可以在贝叶斯推理下形成。经由近似线性化,两者的联系可以通过迭代EKF来明确表示。当近似线性化仅是单个步骤时,包含前通和后通的基于平滑的方法,等价于利用最小二乘问题中信息矩阵的Cholesky分解来求解的基于优化的方法。为了降低计算复杂度,通过保留仅有的关键帧或

滑动窗口来减少要估计的状态变量。特别地,滑动窗的方案或滚动时域估计将损失函数划分为两部分,流行的方式是在一个部分中优化结果,在另一部分中用EKF来忽视过去的状态。

[0033] Full Smoothing,也就是batch non-linear least-squares或者全局优化,状态向量包含所有相机位姿和structure信息,做大规模的非线性优化,显然是最精确的方法,虽说稀疏性降低了计算量,但是依然比滤波大,主要工作集中于降低计算量,可以采用只优化关键帧,以及将优化放在独立线程中,最新的研究采用因子图,增量式地根据测量更新部分状态,典型代表为ISAM。当引入惯性测量单元后,由于惯性测量单元频率通常在100Hz-1kHz量级,无法每次惯性测量单元测量都更新状态向量。通常的做法是在关键帧之间对惯性测量单元进行积分,初始状态由上一帧状态提供,根据惯性测量单元的微分运动方程,积分出当前状态,再将视觉测量引入更新。然而,由于优化过程中上一帧的状态是会变化的,这样积分的初始状态改变了,需要重新算积分,为了避免每一次优化调整后都需要重复计算惯性测量单元积分,利用预积分preintegration将相对运动约束再参数化。

[0034] 步骤2)中通过视觉惯性里程计还获得了所述图片的稀疏点云数据,所述稀疏点云数据辅助计算所述图片中每一帧图片的图片位姿。通过VIO能够得到图像的稀疏点云,但是这些点云因为太稀疏,不足以用来分辨桌面背景和菜品,所以我们只利用稀疏点云辅助得到图片采集设备的位姿。

[0035] 3)将步骤1)中图片通过稠密点云算法并结合所述位置和姿态获得所述图片的3D稠密点云图。步骤3)具体为计算不同位置和姿态的相机所对应的图片的对应像素之间的视差,并使用全局优化获得图片的3D稠密点云图。使用专门提取稠密点云的算法来计算菜品的3D模型,其基本原理是计算不同位置和姿态的相机所对应的图片的对应像素之间的视差,视差越大的点,对应深度越大。最后使用全局优化得到和所有菜品框架一致的稠密点云。

[0036] 4)将所述3D稠密点云图通过图像分割算法获得菜品3D稠密点云图。步骤4)中的图像分割算法采用高度作为区别特征,并对所述3D稠密点云图进行垂直扫描和聚类以将菜品对应的点云分割出来形成菜品3D稠密点云图。有了稠密点云后,再使用算法对点云进行分割,基本原理是将高度最小的一部分点云判定为桌面,之上部分聚拢的点云判定为菜品。

[0037] 5)将所述图片与所述3D稠密点云图匹配以选出待识别图片并计算所述待识别图片所对应的相机的位置和姿态。待识别图片所对应的相机的位置和姿态可以直接通过查询步骤2)中的图片所对应的相机的位置和姿态来获取。或者待识别图片可通过步骤3)中的3D稠密点云图来获取待识别图片所对应的相机的位置和姿态,将待识别图片与3D稠密点云图的投影匹配来获取待识别图片所对应的相机的位置和姿态,通过2D图像和3D点云的匹配,可以计算拍摄这张图片的摄像头在3D空间中所对应的相机的位置和姿态。

[0038] 6)基于所述待识别图片所对应的相机的位置和姿态及所述菜品3D稠密点云图获得待识别图片的菜品区域。步骤6)具体为将菜品3D稠密点云图在与待识别图片所对应的相机的位置和姿态相同的方向上投影获得相应的菜品2D稠密点云图,并将待识别图片与菜品2D稠密点云图匹配以获得待识别图片的菜品区域。

[0039] 实施例一,

[0040] 1)采集连续拍摄的图片,用户围绕菜品在经度方向上旋转90度来拍摄菜品,拍摄

图片的时长是10秒。

[0041] 2) 基于视觉惯性里程计获得每一帧图片所对应的相机的位置和姿态,视觉惯性里程计采用基于滤波方法和优化方法的结合。

[0042] 3) 将步骤1) 中图片通过稠密点云算法并结合所述位置和姿态获得所述图片的3D稠密点云图。

[0043] 4) 将所述3D稠密点云图通过图像分割算法获得菜品3D稠密点云图,图像分割算法采用高度作为区别技术特征。

[0044] 5) 将所述图片与所述3D稠密点云图匹配以选出待识别图片并获取待识别图片所对应的相机的位置和姿态,待识别图片所对应的相机的位置和姿态可以直接通过查询步骤2) 中的图片所对应的相机的位置和姿态来获取。

[0045] 6) 基于所述待识别图片所对应相机的位置和姿态及所述菜品3D稠密点云图获得待识别图片的菜品区域。

[0046] 实施例二

[0047] 1) 采集连续拍摄的图片,用户围绕菜品在经度方向上旋转90度来拍摄菜品,拍摄图片的时长是10秒。

[0048] 2) 基于视觉惯性里程计获得每一帧图片所对应的相机的位置和姿态,视觉惯性里程计采用基于滤波方法和优化方法的结合。

[0049] 3) 将步骤1) 中图片通过稠密点云算法并结合所述位置和姿态获得所述图片的3D稠密点云图。

[0050] 4) 将所述3D稠密点云图通过图像分割算法获得菜品3D稠密点云图,图像分割算法采用高度作为区别技术特征。

[0051] 5) 将所述图片与所述3D稠密点云图匹配以选出待识别图片并获取待识别图片所对应的相机的位置和姿态,待识别图片通过所述3D稠密点云图来获取待识别图片所对应的相机的位置和姿态,即将待识别图片与3D稠密点云图的投影匹配来获取待识别图片所对应的相机的位置和姿态。

[0052] 6) 基于所述待识别图片所对应相机的位置和姿态及步所述菜品3D稠密点云图获得待识别图片的菜品区域。

[0053] 本文中所描述的具体实施例仅仅是对本发明精神作举例说明。本发明所属技术领域的技术人员可以对所描述的具体实施例做各种各样的修改或补充或采用类似的方式替代,但并不会偏离本发明的精神或者超越所附权利要求书所定义的范围。

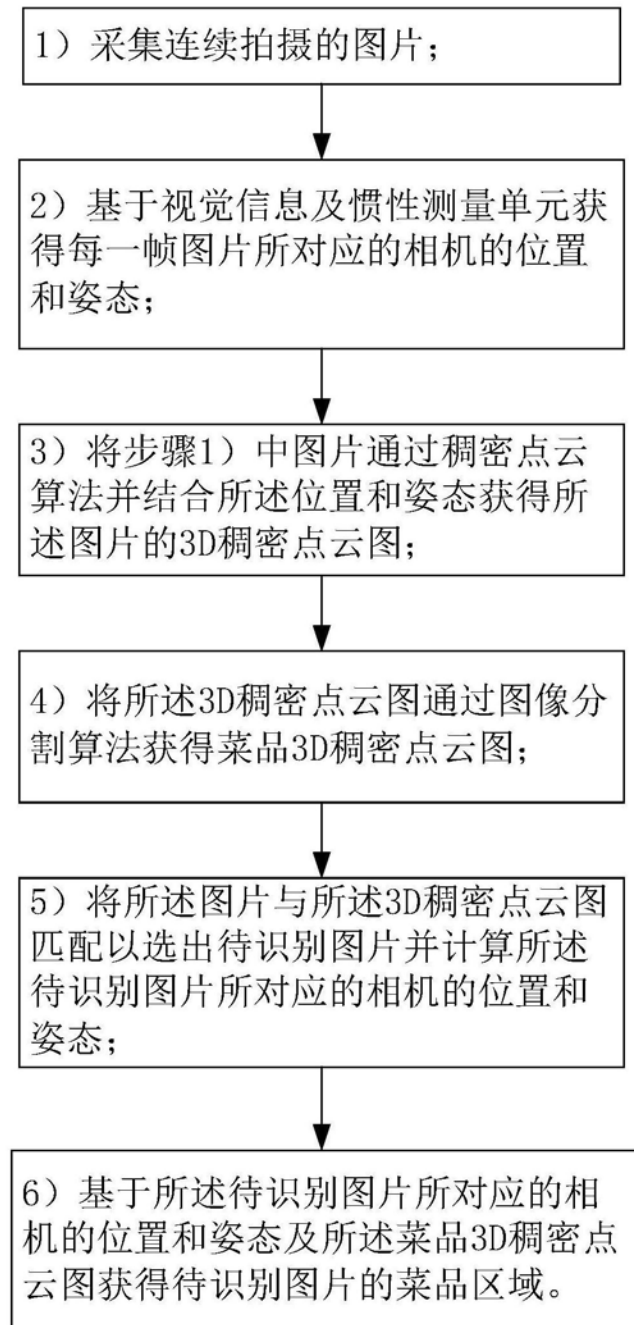


图1