

COLMAP论文阅读笔记——sfm算法、不定期更新

原创

hiccupfrost

2021-03-12 09:18:46

1015

★ 收藏 14

版权

分类专栏：

slam初探

sfm

文章标签：

slam

计算机视觉

算法



slam初探 同时被 2 个专栏收录 ▾

1 订阅 2 篇文章

订阅专栏

文章目录

structure from motion revisited

- 第一步 对应点搜索
 - 特征提取feature extraction
 - 匹配matching
 - 几何验证geometric verification
- 增量式重建
 - 初始化
 - 图像配准
 - 三角剖分triangulation
 - BA
- 面临的挑战
- 贡献
- 场景图增强

structure from motion revisited

Sfm是对 **三维重建** 算法的统称，incremental sfm则表示他是增量式的，在鲁棒性准确性方面有巨大进步；但是灵活性、完整性以及可伸缩性仍然是打破技术壁垒的关键问题。论文后半部分讲述完整的重建流程。sfm 是一系列不同角度拍摄的图像的投影来重建3D结构的过程。

本文探讨的增量式sfm是具有迭代重构组件的顺序处理管道（一系列工序）。通常是先进行feature提取和匹配然后进行几何验证。生成的场景图是重建的基础，之后模型将植入仔细选择的双视角重建（重建作为一个名词，我还不能给出具体的形式），然后再递增注册新图像之前，将对场景中的点进行三角剖分（啥是三角剖分triangulating），过滤异常值，并且使用ba对重建进行优化。

详细介绍：

第一步 对应点搜索

首先会先在输入的图片中找到重叠（overlap）的场景。

输入的场景被定义为

$$I = \{I_i \mid i = 1 \dots N_I\}$$

在上述输入场景中找到重叠图像，标记相同的点的投影，输出是一组经过了几何验证的图像对以及每个点的图像投影图。

这里每个点是指对应点搜索中重叠部分的相同点

图像对被记作C

什么是几何验证——我猜测，是肉眼观察这些点确实是一个点

加粗部分为该步骤的整体概括。之后是细节步骤。

特征提取 feature extraction

对于每个 I_i , SFM都会检测到一个用以表示位置在 X_i 的特征集合

$$F_i = \{(x_j, f_j) \mid j = 1 \dots N_{F_i}\}$$

对这些特征有约束：对辐射和几何变化下应该是不变的，使得SFM可以在不同的图像中唯一识别他们。SIFT和其衍生特征点以及最新了解的那些特征点在鲁棒性方面是事实标准，另外，二进制特征binary feature损失部分鲁棒性提高效率。

这里提到了辐射radiometric，这也解释了之前标定的时候为什么不可以使用发光的屏幕

匹配matching

接下来，SFM会通过利用特征集合，上一步骤中的 F_i 来作为图像的外观描述来发现看到的相同场景部分的图像，朴素的方法会对每个图像对测试场景重叠。

他会基于 F_i 建立一个相似度的度量，为图像 I_b 中的每个特征找到在图像 I_a 中的最相似的特征来搜索对应的特征点，获得这个图像对的特征对应关系。并且这个方法具有计算复杂度

$$O(N_I^2 N_{F_i}^2)$$

且不可以用于大型图像的集合。

于是有许多方法致力于解决规模和效率的问题。（引用了七篇论文，09年-“一天之内造一个罗马城”之类的论文。

输出是一组可能重叠的图像对

$$C = \{\{I_a, I_b\} \mid I_a, I_b \in I, \quad a < b\}$$

和他们关联的特征对应关系矩阵

$$M_{ab} \in F_a \times F_b$$

几何验证geometric verification

第三部分将验证可能重叠的图像对 C 关联的特征对应关系（correspondences）

由于matching仅仅是基于外观的，那么不能保证相应的特征实际映射到同一场景点。因此，SFM通过尝试估计一个变换矩阵，该矩阵通过使用投影几何图形在图像对之间映射他们各自的特征点，来验证匹配。

根据图像对的空间配置，不同的映射将描述他们的几何关系。

单应性矩阵H 用以描述捕获到的平面场景的纯旋转或者纯平移的摄像机的转换

对极几何中通过本征矩阵E（essential）（已经标定）或者基础矩阵（fundamental）F（未标定）描述运动中相机的关系，并且可以通过三焦点张量扩展到三个视图。

三焦点张量——tritenor，在计算机视觉中，是一个 $3 \times 3 \times 3$ 的数组阵列，并入所有投影中的三个视图的几何关系，他与三个视图中的对应点或线的坐标相关联，这与场景结构无关，并且仅仅取决于三个视图之间的相对运动，（即姿态），以及其固有的校准参数（标定的内参）。因此，三焦点张量可以看作是三视图中基础矩阵F的推广，尽管27个元素，只有18个元素是独立的。

如果有效的变换在图像对之间映射了足够多的特征，那么将这个变换视为通过了几何验证。由于来自matching的信息（correspondences）经常被异常值污染，所以会采用RANSAC这类鲁棒的估计方法。该阶段的输出是一组通过几何验证的图像对 \bar{C} ，以及他们的关联的对应关系（correspondences） \bar{M}_{ab} ，同时可选的会有一个关于他们几何关系的描述 G_{ab} ，为了确定适当的关系，可以使用诸如GRIC的决策标准或者QDEGSAC的方法，这个阶段的输出是所谓的场景图，scene graph，其中图像作为节点，已经验证的图像对作为边缘。

增量式重建

重建阶段的输入是场景图，scene graph，输出则是被配准的图像的姿态估计，记作P

$$\mathcal{P} = \{P_c \in SE(3) \mid c = 1 \dots N_p\}$$

输出还包括重建的场景结构，以一组点的形式，这组点记作X

$$\mathcal{X} = \{X_k \in R^3 \mid k = 1 \dots N_X\}$$

初始化

SFM通过精心选择的两视图重建来初始化模型，选择适合的初始图像对十分重要，否则可能根本无法成功重建。此外，重建的鲁棒性，准确性、性能取决于增量过程的种子位置。从图像图中的稠密位置，意味着这里有很多视角有重叠的相机，开始初始化通常会导致冗余度提高，从而使鲁棒性和准确性更高，相反，由于BA处理了整个重建过程中累计的稀疏问题，因此从稀疏位置中进行初始化会降低运行时间。（我并没有理解这里的点，虽然作者好像已经说明了原因）

overall sparser problems这个问题，指的是什么呢？

图像配准

这一步叫做, Image registration, (原谅我一开始一直理解成了图像注册), 从度量重构开始, 使用特征对应关系于已经配准的图像(2d-3d的对应)中的三角点, 来解决一个PnP问题,使得新的图像可以配准至当前的模型。

PnP问题涉及姿态估计, P_c 和相机的内参(没有标定的话)。

因此, 通过新配准的图像的姿势 P_c 来扩展集合 \mathcal{P} ,这里由于2d-3d的对应关系经常被异常值污染, 因此通常使用RANSAC和最小姿态求解器处理已经标定的相机, 对于没标定的相机, 会有各种最小求解器, 作者提出了一种新颖的鲁棒的下一个最优图像选择方法, 用于准确姿态估计和可靠的三角剖分。(4.2节)

强调一下, 配准, register, 指的是将二维图片上的点匹配到三维世界中去。

三角剖分triangulation

新配准的图像必须观察现有已经存在的场景点, 此外, 他还可以通过三角剖分来扩展点集, 增加场景覆盖范围。在至少一幅图像或者更多副图像, 不仅覆盖了新的场景部分, 并且从一个新的视角覆盖了这个场景部分, 当这(些)图像被配准之后, 一个新的场景点 X_k 立刻就可以被三角化和增加到点集合 \mathcal{X} 中。

三角剖分是sfm的关键一步, 他通过冗余性增加了场景的稳定性, 并且通过提供附加的2d-3d对应关系来启用新图像的配准。多视角三角剖分的方法很多, 但是这些方法在sfm应用当中大多鲁棒性具有缺陷或者需要巨大的计算成本代价, 作者提出了一个鲁棒且高效率的三角剖分方法(4.3节)

BA

图像配准和三角测量是独立的过程, 即使他们的输出是高度相关的, 相机姿态的不确定性会传播到三角测量的点, 反之亦然, 附加的三角测量可能会通过增加冗余度来改善相机姿态, 如果没有进一步的改进, SFM通常会迅速漂移到不可恢复的状态。

BA是相机参数 P_c 以及场景点集合 X_k 的联合非线性优化, 可以最大限度的减少重投影的误差 E 。

$$E = \sum_j \rho_j (\|\pi(P_c, X_k) - x_j\|_2^2)$$

这里的 π 函数将场景点投影到图像空间, 损失函数 ρ_j 降低潜在异常值的权重。

Levenberg-Marquardt是解决BA问题的首选方法。BA问题中参数的特殊结构, 激发了Schur补充技巧。高亮部分的方法我都不了解。其中, 前者首先解决简化的相机系统, 然后通过反替换来更新点。由于相机的数量通常小于点的数量, 因此该方案通常更有效, 解决系统有两种选择, 精确和不精确的步长算法。

精确的算法将整个系统存储为空间复杂度为 $O(N_P^2)$ 的稀疏或者稠密矩阵，并且方法的时间复杂度是 $O(N_P^3)$ 。不精确的方法近似求解该系统，使用迭代求解器（计算方法那类求正规方程），例如preconditioned conjugate gradients (PCG)，时间、空间复杂度都是 $O(N_P)$ 。直接的算法是多达几百台相机的首选算法，但是在大规模的条件下代价还是巨大。尽管稀疏的直接方法在很大程度上减少了整个稀疏问题的复杂程度，但是由于通常具有更密集的连接图，因此他们对于大型非结构化照片集是禁止的。在这种情况下，可以选择间接方法。

特别是针对Internet土派年，BA将花费大量的时间来优化许多几乎重复的图片。

在4.5节，作者将讨论识别和参数化高度重叠的图像的方法，以实现提高密集情况下的BA步骤效率。

面临的挑战

一句话概括，完整性、鲁棒性难以保障。

导致原因归根于两点：

1. 这可能是由于对应搜索产生了不完整的场景图，例如由于匹配近似，因此既没有为完整模型提供必要的连通性，也没有为可靠的估计提供足够的冗余度。
2. 这可能是由于重建阶段由于缺少或不正确的场景结构而导致的配准图像而导致的；图像配准和三角剖分具有共生关系，即图像只能配准到现有的场景结构，而场景结构只能从配准的图像进行三角剖分[64]。

贡献

介绍了一种新算法，他改进了sfm中的主要挑战。

1. 几何验证策略，增加信息来增强场景显示，随后提高了初始化和三角剖分组件的鲁棒性
2. 次最优视图选择将最大化增量重建过程的鲁棒性和准确性
3. 健壮的三角剖分方法，它以降低的计算成本产生了比现有技术明显更完整的场景结构
4. 采用迭代BA，重新三角剖分和离群滤波策略，通过减轻漂移效应来显着提高完整性和准确性
5. 通过冗余视图挖掘对密集照片集进行更有效的BA参数化，这将导致系统在健壮性和完整性方面明显优于现有技术：同时保持了效率

场景图增强

scene graph augmentation

RANSAC算法的基本假设是样本中包含正确数据(inliers, 可以被模型描述的数据), 也包含异常数据(outliers, 偏离正常范围很远、无法适应数学模型的数据), 即数据集中含有噪声。这些异常数据可能是由于错误的测量、错误的假设、错误的计算等产生的。同时RANSAC也假设, 给定一组正确的数据, 存在可以计算出符合这些数据的模型参数的方法。

也就是说下文中计算的inliers数量是基于ransac方法的过程中求到的? 待验证

正如上文中提到的, 几何验证策略, 以通过几何关系来增加场景图。

首先我们要估计一个基础矩阵F, (基础矩阵F: 除了包含E的信息外, 还包含了两个摄像机的内参数。由于F包含了这些内参数, 因此它可以在像素坐标系将两个摄像机关联起来。), 一般, 这里估计F是由于没有标定内参, 因此估计F。

如果找到了至少 N_F 个像素(非异常值, 原文中用inliers表示, 但是这个inliers是如何产生的, 看到这里我还不是很理解), 那么将这对图像视为通过了几何验证, 接下来, 我们将对统一对图像通过确定单应性的inliers数量 N_H , 来对之前得到的变换结果进行分类

若近似模型选择方法GRIC。当 $N_H/N_F < \epsilon_{HF}$, 我们将这类变换情况视为这张图片到那张图片的一般场景中相机是移动的(常规)。

如果是已经标定的图像, 我们估计一个本征矩阵E, 以及获得他的inliers的数量 N_E 。如果 $N_E/N_F > \epsilon_{EF}$, 我们将认为标定结果正确。在标定结果正确的情况下, 若同时 $N_H/N_F < \epsilon_{HF}$, 则分解矩阵E, 从inliers的对应关系的进行三角化, 确定三角剖分角 α_m , 使用这个角度值, 可以将视角的转换中纯旋转(全景)和纯平移(平面)的情况。

此外, 互联网照片中的水印时间戳或帧(通称WTF)会错误地连接不同地标的图像。为了检测出这种图像, 我们在图像边缘处做一个相似度转换的估计, 若具有 N_s 个内点(理解成有这么多个点可能是WTF), 如

KaTeX parse error: Undefined control sequence: \or at position 22: ...F>\epsilon_{SF}\or N_S/N_E>\epsilon_{SE}...

那么认为这对图像是一个WTF, 就不会被插入到scene graph中去。

colmap 三维重建

04-20

三维重建效果很好的开源软件, 超过visualSFM的效果, 不限制图片数量

SFM方向开源软件COLMAP代码分析

Z5122的博客 5443

作为目前最优的SFM开源软件COLMAP(离线处理模式), 其详细介绍可参考: <http://colmap.github.io/>, 编译源码后运行程序界面如下: 其稀疏重建数据处理流程为: 1) "File->Proje...



请发表有价值的评论, 博客评论不欢迎灌水, 良好的社区氛围需大家一起维护。

抢沙发



评论