目录

[1 Introduction 2](#_Toc104329387)

[2 Related work 4](#_Toc104329388)

[3 Closed-form solution: Known biases and IMU-Camera calibration 6](#_Toc104329389)

[3.1 Solution of the linear system in the absence of noise 10](#_Toc104329390)

[3.2 Analysis of the rank of A and minimal cases 12](#_Toc104329391)

[3.2.1 Two features seen in three images 14](#_Toc104329392)

[3.2.2 One feature seen in four images 14](#_Toc104329393)

[4 Unknown biases and/or IMU-Camera calibration 16](#_Toc104329394)

[4.1 Unknown extrinsic calibration, known IMU biases 16](#_Toc104329395)

[4.2 Unknown extrinsic calibration, unknown IMU biases 17](#_Toc104329396)

[5. Solution in the presence of noise 20](#_Toc104329397)

**Closed-form Solutions for Vision-aided Inertial Navigation**

视觉辅助惯性导航的封闭式解决方案

This report focuses on motion estimation using inertial measurements and observations of naturally occurring point features. To date, this task has primarily been addressed using filtering methods, which track the system state starting from known initial conditions. However, when no prior knowledge of the initial system state is available, (e.g., at the onset of the system’s operation), the existing approaches are not applicable. To address this problem, in this work we present algorithms for computing all the observable quantities (platform attitude and velocity, feature positions, IMU biases, and IMU-camera calibration) in closed form directly from the sensor measurements, without any prior knowledge. As a key contribution of this work, we identify and analyze the properties of minimal problems that have a finite number of solutions, as well as singular trajectories, in which solutions cannot be computed. Additionally, to address the presence of noise in the measurements, we present a quadratically constrained least-squares solution and an iterative maximumlikelihood estimator.

本报告侧重于使用惯性测量和对自然发生的点特征的观察进行运动估计。迄今为止，这项任务主要是使用滤波方法来解决的，该方法从已知的初始条件开始跟踪系统状态。然而，当没有初始系统状态的先验知识可用时，（例如，在系统开始运行时），现有方法不适用。为了解决这个问题，在这项工作中，我们提出了直接从传感器测量以封闭形式计算所有可观察量（平台姿态和速度、特征位置、IMU 偏差和 IMU 相机校准）的算法，没有任何先验知识。作为这项工作的一个关键贡献，我们识别和分析了具有有限数量解决方案的最小问题的属性，以及无法计算解决方案的奇异轨迹。此外，为了解决测量中存在的噪声问题，我们提出了一个二次约束最小二乘解和一个迭代最大似然估计器。

# Introduction

In recent years, the topic of motion estimation using visual and inertial measurements (often termed vision-aided inertial navigation) has attracted significant research interest (e.g., see [1–6] and references therein). Both cameras and MEMS inertial sensors are compact, inexpensive, and have low power requirements. Moreover, these sensors can operate in virtually any environment, and allow for full-3D pose estimation, thus providing a very versatile solution for navigation. The vast majority of existing techniques for navigation using camera and IMU measurements employ either a recursive Bayesian estimation approach [1–4], or a smoothing formulation [5]. In both cases, an accurate initial guess (prior estimate) for the state is necessary for reliable estimation. This is due to the fact that both types of methods rely on linearization of the measurement models, and thus in the absence of an accurate initial estimate, large linearization errors can lead to divergence.

近年来，使用视觉和惯性测量（通常称为视觉辅助惯性导航）的运动估计主题引起了极大的研究兴趣（例如，参见 [1-6] 和其中的参考文献）。相机和 MEMS 惯性传感器都很紧凑、价格低廉，并且具有低功耗要求。而且，

这些传感器几乎可以在任何环境中运行，并允许进行全 3D 姿态估计，从而为导航提供了一个非常通用的解决方案。绝大多数现有的使用相机和 IMU 测量进行导航的技术采用递归贝叶斯估计方法 [1-4] 或平滑公式 [5]。在这两种情况下，对状态的准确初始猜测（先前估计）对于可靠估计是必要的。这是因为这两种方法都依赖于测量模型的线性化，因此在没有准确的初始估计的情况下，大的线性化误差会导致发散。

In current practice, to initialize any of the state estimation methods discussed above, one typically uses domainspecific knowledge on a case-by-case basis. For instance, in certain applications, additional sensors (e.g., inclinometer and/or GPS) may be available, or it may be known that the platform is initially at rest. However, no generally applicable methods exist for determining all the observable system states directly from the sensor data, without use of a prior or of domain-specific knowledge. To address this problem, this report presents the following key contributions:

在当前实践中，为了初始化上面讨论的任何状态估计方法，通常会根据具体情况使用特定领域的知识。例如，在某些应用中，额外的传感器（例如，倾角仪和/或 GPS）可能是可用的，或者可能知道平台最初是静止的。然而，在不使用先验知识或特定领域知识的情况下，不存在直接从传感器数据确定所有可观察系统状态的普遍适用的方法。为了解决这个问题，本报告提出了以下主要贡献：

* We propose algorithms for computing, in closed form, all the observable quantities of the system directly from the visual and inertial measurements. In this work, we do not require that the 3D positions of the features are known in advance (i.e., we do not utilize fiducial points). Instead, the feature coordinates with respect to the camera are also computed as part of the solution.
* 我们提出算法，以封闭形式直接从视觉和惯性测量中计算系统的所有可观察量。在这项工作中，我们不需要预先知道特征的 3D 位置（即，我们不使用基准点）。反而，相对于相机的特征坐标也被计算为解决方案的一部分。
* We carry out a detailed analysis of the properties of the problem, and identify minimal problems for which a solution exists. Moreover, we determine singular cases for the platform trajectory and the features’ spatial configuration, which result in either multiple or infinite solutions. For instance, we prove that when the camera is moving with a constant acceleration, two discrete solutions for the trajectory exist in general.
* 我们对问题的性质进行详细分析，并确定存在解决方案的最小问题。此外，我们确定了平台轨迹和特征空间配置的奇异情况，这导致了多个或无限的解决方案。例如,我们证明，当相机以恒定加速度移动时，通常存在两个离散的轨迹解。
* In the presence of noise, the formulation proposed here results in a quadratically-constrained least-squares problem, which can be solved analytically. This has the advantage of not requiring any prior information about the state, but it does not carry any optimality properties. To properly treat the noise in the sensor data, we propose a nonlinear iterative maximum-likelihood estimator (MLE) for estimating the observable quantities of the system
* 在存在噪声的情况下，此处提出的公式会导致二次约束的最小二乘问题，该问题可以解析求解。这具有不需要任何关于状态的先验信息的优点，但它不具有任何最优性。为了正确处理传感器数据中的噪声，我们提出了一种非线性迭代最大似然估计器（MLE）来估计系统的可观察量
* Our main motivation for the closed-form solutions presented in this work is the initialization of recursive (e.g., [1–3]) and iterative estimators (e.g., the MLE described in Section 5.2). However, the methods presented here can also be applied for improving the robustness of vision-aided inertial navigation methods. For instance, the solutions of the minimal problems can be utilized within a RANSAC framework for outlier rejection. Moreover, the closed-form solutions can be run in parallel to a filtering algorithm (e.g., the EKF) as a sanity check, in order to detect possible divergence. In what follows, we present the details of our work.
* 我们在这项工作中提出的封闭形式解决方案的主要动机是递归（例如，[1-3]）和迭代估计器（例如，第 5.2 节中描述的 MLE）的初始化。然而，这里介绍的方法也可以用于提高视觉辅助惯性导航方法的鲁棒性。例如，
* 最小问题的解决方案可以在 RANSAC 框架内用于异常值拒绝。此外，封闭形式的解决方案可以与过滤算法（例如，EKF）并行运行，作为健全性检查，以检测可能的分歧。在下文中，我们将介绍我们的工作细节。

# Related work

In recent publications, the observability properties of the vision-aided inertial navigation system have been examined [1,2]. These works show that, in the absence of reference points with known global coordinates, the global position of the IMU, as well as the rotation about the axis of gravity (i.e., the yaw) are not observable. On the other hand, the following quantities are in general observable:

在最近的出版物中，已经检查了视觉辅助惯性导航系统的可观察性属性 [1,2]。这些工作表明，在没有已知全局坐标的参考点的情况下，IMU 的全局位置以及围绕重力轴的旋转（即偏航）是不可观测的。另一方面，通常可以观察到以下量：

(O1) The IMU attitude with respect to the horizontal plane (i.e., the roll and pitch), IMU 相对于水平面的姿态（即横滚和俯仰），

(O2) The IMU trajectory (position, velocity, and orientation) with respect to the initial IMU frame, 相对于初始 IMU 帧的 IMU 轨迹（位置、速度和方向），

(O3) The feature positions with respect to the initial IMU frame. 相对于初始 IMU 帧的特征位置。

(O4) The IMU biases,

(O5) The transformation between the IMU and camera frames (ie., the camera-to-IMU calibration). IMU 和相机帧之间的转换（即相机到 IMU 的校准）

These results can be justified intuitively: if no known landmarks are available, the only source of absolute pose information is the gravity vector, which is sensed by the IMU accelerometers. Therefore, we can estimate the global orientation with respect to the horizontal plane (i.e., the plane normal to gravity), while all other quantities can be determined only in the local frame of the first IMU pose.

这些结果可以直观地证明：如果没有已知的地标可用，则绝对姿势信息的唯一来源是重力矢量，它由 IMU 加速度计感应。因此，我们可以估计相对于水平面（即垂直于重力的平面）的全局方向，而所有其他量只能在第一个 IMU 位姿的局部框架中确定。

The results of [1, 2] provide valuable intuition into the properties of the problem at hand, and serve as a motivation for further study. Since we know which quantities are observable in general, the next step is to identify (i) the minimal conditions (number of features and images) required for a solution, and (ii) the number of possible solutions. The aforementioned works do not address these questions. The same holds for [5], in which an iterative method for determining the IMU orientation and attitude is proposed, but the conditions under which a solution exists are not studied. In Section 3.2, these limitations will be addressed

[1, 2] 的结果为手头问题的性质提供了有价值的直觉，并作为进一步研究的动力。由于我们知道通常可以观察到哪些量，下一步是确定 (i) 解决方案所需的最小条件（特征和图像的数量），(ii) 可能解决方案的数量。上述作品没有解决这些问题。 [5] 也是如此，其中提出了一种确定 IMU 方向和姿态的迭代方法，但没有研究解决方案存在的条件。在第 3.2 节中，将解决这些限制

In a paper that will soon appear [7], the visual and inertial measurements are used to analytically determine the initial velocity and attitude, as well as the feature positions. Compared to that work, in this report we (i) offer a detailed characterization of the special cases, under which a solution exists, (ii) propose solutions that can address the presence of noise in the sensor data, and (iii) we describe solutions for the observable quantities O4-O5 described above.

在一篇即将发表的论文 [7] 中，视觉和惯性测量用于分析确定初始速度和姿态，以及特征位置。与这项工作相比，在本报告中，我们 (i) 提供了特殊情况的详细描述，在这些情况下存在解决方案，

(ii) 提出可以解决传感器数据中存在噪声的解决方案，以及 (iii) 我们描述上述可观察量 O4-O5 的解决方案。

We point out that methods for motion estimation from features have been an important area of computer vision research (e.g., determination of the epipolar geometry from five point correspondences [8,9]). However, it is important to note that, in the absence of fiducial points, camera-only approaches can only estimate the camera motion up to an unknown scale. In contrast, the availability of the inertial measurements makes it possible to estimate (i) the absolute scale of the platform’s motion, as well as (ii) the absolute orientation of the platform with respect to the horizontal plane. In recent publications, the second of these properties has been exploited. Specifically, [10, 11] examine the problem of motion estimation with a known vertical direction, and derive algorithms for its solution. However, these approaches do not use the inertial measurements to estimate the scale of the camera motion, and require that the camera is at rest when images are recorded (to allow for estimating its orientation using the accelerometer measurements). In our work, both of these limitations are lifted

我们指出，根据特征进行运动估计的方法一直是计算机视觉研究的一个重要领域（例如，从五点对应确定核几何形状 [8,9]）。然而，重要的是要注意，在没有基准点的情况下，相机方法只能估计到未知尺度的相机运动。相比之下，惯性测量的可用性使得可以估计 (i) 平台运动的绝对比例，以及 (ii) 平台相对于水平面的绝对方向。在最近的出版物中，这些属性中的第二个已被利用。具体来说，[10, 11] 研究了具有已知垂直方向的运动估计问题，并推导出其解决方案的算法。然而，这些方法不使用惯性测量来估计相机运动的规模，并要求在记录图像时相机处于静止状态（以允许使用加速度计测量来估计其方向）。在我们的工作中，这两个限制都被解除了

# Closed-form solution: Known biases and IMU-Camera calibration

封闭式解决方案：已知偏差和 IMU-Camera 校准

We first study the case in which the IMU biases and the transformation between the IMU and camera are known (e.g., from prior calibration), and our goal is to estimate the observable quantities O1-O3 defined in Section 2. The case where all quantities are unknown and need to be estimated from the data is treated in Section 4.

我们首先研究 IMU 偏差和 IMU 与相机之间的转换已知的情况（例如，通过先前的校准），我们的目标是估计第 2 节中定义的可观察量 O1-O3。所有量的情况是未知的，需要从数据中估计，在第 4 节中处理。

Consider the case where N images are recorded at the time instants t0, t1, . . . , tN−1. By employing a suitable image processing algorithm (e.g., KLT tracking [12]), we track M feature points in the images. For the solution presented here, we do not require all features to be tracked in all images. In addition to the feature observations, the IMU (gyroscope and accelerometer) measurements for the time interval [t0, tN−1] are available. In the remainder of the section, we show how the feature observations and IMU measurements can be used to estimate the observable quantities O1-O3.

考虑在时刻 t0、t1、... . . , tN-1记录 N 个图像的情况。。通过采用合适的图像处理算法（例如，KLT 跟踪 [12]），我们跟踪图像中的 M 个特征点。对于此处介绍的解决方案，我们不需要在所有图像中跟踪所有特征。除了特征观察，

时间间隔 [t0, tN−1] 的 IMU（陀螺仪和加速度计）测量可用。在本节的其余部分，我们将展示如何使用特征观察和 IMU 测量来估计可观察量 O1-O3。

The measurements of the IMU gyroscopes and accelerometers are given by the following equations [13]:

IMU 陀螺仪和加速度计的测量值由以下等式 [13] 给出：

where denotes the 3D rotational velocity vector expressed in the IMU frame, is the IMU acceleration in the global frame, is the gravitational acceleration vector expressed in the global frame, , and ba are the gyroscope and accelerometer biases, while , and represent the noise in the gyroscope and accelerometer measurements, respectively. In the remainder of this section, we assume that the biases are known and removed from the measurements. Moreover, in order to obtain a closed-form solution, we will ignore the measurement noise, which is addressed properly in Section 5.

其中， 表示 IMU 坐标系中表示的 3D 旋转速度矢量， 是全局坐标系中的 IMU 加速度， 是全局坐标系中表示的重力加速度矢量，bg 和 ba 是陀螺仪,加速度计偏差，而 和 表示陀螺仪和加速度计测量中的噪声，在本节的其余部分，我们假设偏差是已知的并从测量中移除。此外，为了获得封闭形式的解决方案，我们将忽略测量噪声，这在第 5 节中得到了适当的解决。

To estimate the IMU orientation change in the interval , we integrate the following differential equation [14]:

为了估计区间 内的 IMU 方向变化，我们整合了以下微分方程 [14]：

in [t0, ti ]. This yields the rotation matrix B0 Bi C = B0 B C(ti), which describes the IMU rotation between times t0 and ti . Note that in practice, we may not directly integrate the above equation, but use a quaternion-based representation of the orientation, or roll-pitch-yaw angles. These representations are equivalent for our purposes, and we here use (3) to simplify the presentation. The global position of the IMU at time ti is computed as:

在 。这产生了旋转矩阵 ，它描述了时间 t0 和 ti 之间的 IMU 旋转。请注意，在实践中，我们可能不会直接对上述方程进行积分，而是使用基于四元数的方向表示或滚动-俯仰-偏航角。这些表示对于我们的目的是等效的，我们在这里使用 (3) 来简化表示。 IMU 在时间 ti 的全局位置计算如下：



where is the initial velocity, and ∆ti = ti −t0. Using Eq. (2) and the above equation we obtain (see Appendix A):

其中 是初始速度，Δti = ti -t0。使用方程式。 （2）和上面我们得到的方程（见附录A）：

图片包含 图示

描述已自动生成

Note that Eq. (5) involves the global position, velocity, and orientation of the IMU, which are impossible to determine using only visual and inertial data. Therefore, our next step is to remove these quantities. To this end, we note that the IMU position at time with respect to is given by . Using this expression, we can re-arrange Eq. (5) to obtain

请注意，方程式。 (5) 涉及 IMU 的全局位置、速度和方向，仅使用视觉和惯性数据是无法确定的。因此，我们的下一步是删除这些数量。为此，我们注意到 IMU 在时间 相对于 的位置由 给出。使用这个表达式，我们可以重新排列方程。 (5) 获得

where is the IMU velocity at time t0 expressed with respect to frame , and is the gravity vector expressed with respect to the same frame.

其中 是相对于帧 表示的时间 t0 的 IMU 速度，并且 是相对于同一帧表示的重力矢量。

Eq. (7) demonstrates that in order to estimate the platform trajectory using the inertial measurements, the initial velocity, B0 v0, as well as the gravity vector in the local frame, B0 g are required (note that s(ti) depends only on the IMU measurements [Eq. (6) and (3)]). If these two vectors are known, we can employ Eq. (7) to determine the platform position in the local frame for any time instant of interest. By differentiation of the position we can also determine the platform velocity. Moreover, the IMU orientation in the local frame can be determined by integration of Eq. (3), which only requires the gyroscope measurements. The above discussion shows that to determine the platform trajectory in the local frame (O2 in Section 2), it suffices to know the vectors B0 v0 and B0 g. Moreover, note that, since the gravity vector is normal to the horizontal plane, B0 g defines the orientation of the horizontal plane with respect to {B0}. Thus, estimating B0 g is equivalent to estimating the IMU orientation with respect to the horizontal plane (O1 in Section 2).

方程。 (7) 证明，为了使用惯性测量来估计平台轨迹，需要初始速度 以及局部坐标系中的重力矢量 （请注意， 仅取决于IMU 测量 [Eq. (6) 和 (3)])。如果这两个向量是已知的，我们可以使用方程。(7) 确定任何感兴趣时刻在本地框架中的平台位置。通过位置的微分，我们还可以确定平台速度。此外，局部坐标系中的 IMU 方向可以通过对等式的积分来确定。 (3)，它只需要陀螺仪测量。上面的讨论表明，要确定局部坐标系中的平台轨迹（第 2 节中的 O2），知道向量 和 就足够了。此外，请注意，由于重力矢量垂直于水平面，因此 定义了水平面相对于 的方向。因此，估计 等效于估计 IMU 相对于水平面的方向（第 2 节中的 O1）。

In summary, we see that if we are able to determine B0 v0 and B0 g, we can immediately obtain the observable quantities O1 and O2 described in Section 2. We next show how we can employ the feature observations to achieve this goal, as well as to compute the feature positions (O3). Assuming an intrinsically calibrated camera, the observation of the j-th feature at time ti is described by the perspective camera model:

总之，我们看到，如果我们能够确定 和 ，我们可以立即获得第 2 节中描述的可观察量 O1 和 O2。接下来我们将展示如何利用特征观察来实现这一目标，以及来计算特征位置（O3）。假设一个内在校准的相机，

在时间 对第 j 个特征的观察由透视相机模型描述：

文本

中度可信度描述已自动生成

where Cipj is the position of the j-th feature with respect to the camera frame at time ti , and nij is the measurement noise. We use the set Sm to describe all pairs of indices {i, j} that describe the available measurements.

其中 是第 j 个特征在时间 相对于相机帧的位置， 是测量噪声。我们使用集合 来描述描述可用测量的所有索引对 {i, j}。

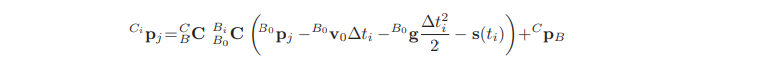
Using basic properties of frame transformations, we can express the vector Cipj as follows:

利用帧变换的基本性质，我们可以将向量 表示如下：



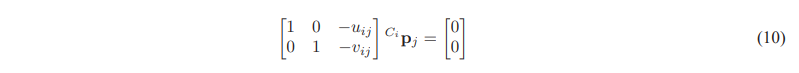
where B0 pj is the position of the feature with respect to {B0}, while { C BC, C pB} denotes the constant transformation (rotation and translation) between the IMU and camera frames. Using Eq. (7), we can rewrite (9) as

其中 是特征相对于 {B0} 的位置，而 表示 IMU 和相机帧之间的恒定变换（旋转和平移）。使用方程式(7), 我们可以将 (9) 改写为



On the right-hand side of above equation the unknown quantities are the vectors B0 pj , B0 v0, and B0 g, while all other terms are known. Proceeding further, we employ Eq. (8), to obtain (ignoring the measurement noise):

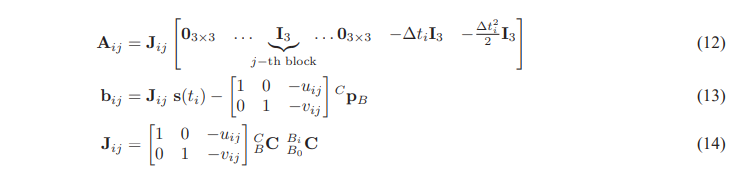
在上述等式的右侧，未知量是向量，而所有其他项都是已知的。进一步进行，我们使用方程式。 (8)、获得（忽略测量噪声）：



By re-arranging terms, the above equation can be written as Aijx = bij , where x is the following (3M + 6) × 1 vector:

通过重新排列项，上述等式可以写成 ，其中 x 是以下 (3M + 6) × 1 向量：





We therefore see that from the observation of the j-th feature in the i-th image we obtain one equation of the form Aijx = bij . By collecting the equations resulting from all feature measurements, we obtain the linear system

因此，我们看到，通过观察第 i 个图像中的第 j 个特征，我们得到了一个形式为 的方程。通过收集所有特征测量产生的方程，我们得到线性系统

where A is a matrix with block rows Aij , and b is a block vector with block elements bij , for all {i, j} ∈ Sm. It is important to observe that both A and b can be computed using (i) the feature measurements, (ii) the IMU measurements, and (iii) the known IMU-camera transformation, while all the unknown quantities are included in the vector x, which appears linearly in the above equation

其中 A 是具有块行 的矩阵，b 是具有块元素 的块向量，对于所有 。重要的是要观察到 A 和 b 都可以使用

1. 特征测量值、
2. IMU 测量值
3. 已知的 IMU 相机变换来计算，而所有未知量都包含在向量 x 中,在上述等式中呈线性

## Solution of the linear system in the absence of noise

无噪声线性系统的求解

If the matrix A has full column rank, ie., if rank(A) = 3M + 6, then the solution to Eq. (15) is unique, given by

如果矩阵 A 具有满列秩，即如果 rank(A) = 3M + 6，则方程的解。 (15) 是唯一的，由下式给出

An interesting observation is that this solution allows us to determine not only the direction of the gravity vector in {B0}, but also the magnitude of the gravitational acceleration, g = ||B0 g||2. While this may be useful in certain settings, in the vast majority of applications the norm of the gravity vector is known in advance with high precision. Knowing the magnitude of gravity provides us with an additional constraint, which we can use to expand the situations under which a solution is possible.

一个有趣的观察是，这个解不仅可以让我们确定 中重力矢量的方向，还可以确定重力加速度的大小，。虽然这在某些设置中可能很有用，在绝大多数应用中，重力矢量的范数是预先以高精度已知的。知道重力的大小为我们提供了一个额外的约束，我们可以用它来扩展可能解决方案的情况。

Specifically, if the rank of A equals 3M+5 (i.e., one less than the number of its columns), the matrix has a nullspace of dimension one: dim(N (A)) = 1. In this case Eq. (15) has an infinite number of solutions, described by x = A+b + αn, where A+ is a pseudoinverse of A, n is a basis vector for N (A), and α is an arbitrary scalar [15]. However, if we know the magnitude of the gravitational acceleration, g, then we can enforce the constraint that the norm of the vector formed by the last three elements of x is equal to g [Eq. (11)]:

具体来说，如果 A 的秩等于 3M+5（即，比它的列数少 1），则矩阵有一个维度为 1 的零空间：dim(N (A)) = 1。 (15) 有无数个解，由 描述，其中 A+ 是 A 的伪逆，n 是 N (A) 的基向量，α 是任意标量 [15]。然而，

如果我们知道重力加速度 g 的大小，那么我们可以强制执行由 x 的最后三个元素形成的向量的范数等于 g [Eq. (11)]:



This is a quadratic equation in α and has two real roots, α1 and α2. Thus, we obtain two discrete solutions for x:

这是 α 中的二次方程，有两个实根，α1 和 α2。因此，我们获得了 x 的两个离散解：



It should be noted that if the nullspace of A is of dimension higher than one, then the solutions of (15) contain more than one arbitrary scalars. In that case, even if the known-gravity constraint is used, it is not possible to obtain a finite number of solutions. We now summarize the above results as follows:

需要注意的是，如果 A 的零空间的维数大于 1，则 (15) 的解包含多个任意标量。在这种情况下，即使使用已知重力约束，也不可能获得有限数量的解。我们现在将上述结果总结如下：

**Result 1**. If A has full column rank (i.e., dim(N (A)) = 0), there exists a unique solution for x, given by Eq. (16). On the other hand, if dim(N (A)) = 1, there exist two solutions in general. These can be found by numerically computing n and A+, solving Eq. (17) for α, and substituting in (18).

结果 1. 如果 A 具有完整的列秩（即，dim(N (A)) = 0），则存在 x 的唯一解，由方程式给出。 (16)。另一方面，如果 dim(N (A)) = 1，则一般存在两种解。这些可以通过数值计算 n 和 A+ 来找到，求解方程。 (17) 代入 (18)。

In either case, these solutions provide, in closed form, the position of all the features, the initial IMU velocity, and the gravity vector, expressed with respect to the first IMU frame. As previously explained, this allows us to compute the observable quantities O1-O3 in the system, which is our key objective.

在任何一种情况下，这些解决方案都以封闭形式提供所有特征的位置、初始 IMU 速度和重力矢量，相对于第一个 IMU 帧的表示。如前所述，这使我们能够计算系统中的可观察量 o1-o3，这是我们的主要目标。

## Analysis of the rank of A and minimal cases

A和最小案例的等级分析

We next examine the rank of A and the dimension of N (A) in important cases of interest, and identify the minimal requirements (in terms of the number of features and images) that allow for computing a finite number of solutions.

接下来，我们在感兴趣的重要案例中检查 A 的等级和 N (A) 的维数，并确定允许计算有限数量的解决方案的最低要求（就特征和图像的数量而言）。

We first derive a condition for the number of measurements needed. Since A contains two rows for each feature observation [Eq. (12)], the total number of rows in the matrix is 2K, where K is the total number of measurements available: K = PM j=1 nj , where nj is the number of times feature j has been observed. Therefore, rank(A) ≤ 2K, which, by application of the rank-nullity theorem [15], yields the inequality dim(N (A)) ≥ 3M + 6 − 2K. Using this inequality and Result 1, we obtain the following result:

我们首先推导出所需测量次数的条件。由于 A 对于每个特征观察都包含两行 [Eq. (12)]，矩阵中的总行数为 2K，其中 K 是可用测量的总数： ，其中 是已观察到特征 j 的次数。因此，rank(A) ≤ 2K，其中，通过应用秩零定理 [15]，得出不等式 dim(N (A)) ≥ 3M + 6 − 2K。使用这个不等式和结果 1，我们得到以下结果：

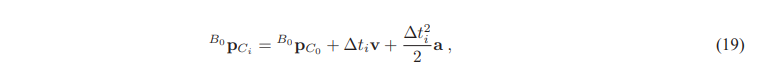
**Result 2.** For a unique solution to exist, a necessary, but not sufficient, condition is that the number of measurements satisfies K ≥ 3 2M + 3. On the other hand, for a finite number of solutions (one or two) to exist, a necessary condition is K ≥ 3 2M + 5 2 . If neither of the above necessary conditions are met, then it is guaranteed that infinite solutions exist.

结果 2. 对于唯一解的存在，一个必要但不充分的条件是测量次数满足 K ≥ 3 2M + 3。另一方面，对于有限数量的解（一个或两个）存在, 一个必要条件是 K ≥ 3 2M + 5 2 。如果以上两个必要条件都不满足，

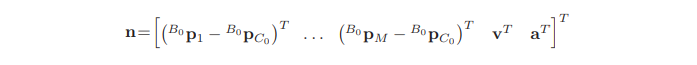
则保证存在无限解。

If the feature measurements meet the above necessary conditions, then we can proceed to compute dim(N (A)), to verify its dimension (since the above conditions are not sufficient, when they are met it may still be possible that infinite solutions exist). For arbitrary trajectories and feature placements, the dimension of N (A) can be numerically computed, on a case-by-case basis. However, by employing an analytical approach to determining the rank of A, we can identify an important special case. Specifically, in Appendix B.1, we show that if the camera moves with a constant acceleration, that is, if

如果特征测量满足上述必要条件，那么我们可以继续计算dim(N(A))，以验证其维度（由于上述条件不充分，当满足时，仍有可能存在无限解）。对于任意轨迹和特征放置，N (A) 的维数可以根据具体情况进行数值计算。但是，通过采用分析方法来确定 A 的等级，我们可以确定一个重要的特殊情况。具体来说，在附录 B.1 中，我们展示了如果相机以恒定加速度移动，也就是说，如果



then A has a nullspace of dimension at least one, and the following vector is one basis vector for this nullspace:

那么 A 有一个维数至少为 1 的零空间，下面的向量是这个零空间的一个基向量：

If enough feature measurements are available, this will be the only basis vector for the nullspace of A, and thus dim(N (A)) = 1. Thus, by application of Result 1, we conclude that

如果有足够的特征测量可用，这将是 A 的零空间的唯一基向量，因此 dim(N (A)) = 1。因此，通过应用结果 1，我们得出结论

**Result 3**. When the camera moves with a constant acceleration, two discrete solutions for the camera trajectory exist in general, given by (18).

结果 3. 当相机以恒定加速度移动时，相机轨迹的两个离散解通常存在，由 (18) 给出。

Further analyzing this result, it is interesting to explore what happens if a = 0, i.e., if the camera moves with a constant velocity. In that case, the last 3 × 1 block of n is zero, and α vanishes from Eq. (17). In turn, this means that α cannot be determined, and infinite solutions exist, given by x = A+b + αn. It is important to observe, however, that since the last three elements of n are zero, these solutions all have the same value for B0 g. In other words:

进一步分析这个结果，有趣的是探索如果 a = 0 会发生什么，即如果相机以恒定速度移动。在这种情况下，n 的最后一个 3 × 1 块为零，并且 α 从等式中消失。 (17)。反过来，这意味着无法确定 α，存在无限解，由 给出。观察很重要，但是，由于 n 的最后三个元素为零，因此这些解对于 都具有相同的值。换句话说：

**Result 4.** When the camera is moving with a constant velocity, the orientation of the platform with respect to the horizontal plane can be uniquely determined, while an infinite number of solutions exist for the platform velocity and the feature positions. The physical interpretation of the infinite set of solutions is that the scale of the trajectory cannot be determined [1].

结果4.当相机匀速运动时，平台相对于水平面的方位可以唯一确定，而平台速度和特征位置的解存在无穷多个。无限组解的物理解释是无法确定轨迹的尺度 [1]。

We now turn our attention to identifying the minimal conditions, in terms of the number of features and images, that make it possible to obtain discrete solutions for the observable states. We start by noting that if only one image is available, the number of equations in Eq. (15) is always less than the number of unknowns. Moreover, in Appendix B.2.2, we show that if only N = 2 images are available, the dimension of the nullspace of A is at least 3, regardless of the number of features. The minimum number of images needed for a finite number of solutions is N = 3. For this case, by comparing the number of unknowns to the number of measurements, we conclude that the number of features must be at least M = 2. On the other hand, if only one feature is available (M = 1), we obtain the condition N ≥ 4, as the lower bound on the number of images (Appendix B.2.2).

我们现在将注意力转向根据特征和图像的数量来识别最小条件，这使得获得可观察状态的离散解成为可能。我们首先注意到，如果只有一张图像可用，则方程中的方程数。 (15) 总是小于未知数的个数。而且，

在附录 B.2.2 中，我们展示了如果只有 N = 2 张图像可用，则 A 的零空间的维度至少为 3，而与特征数量无关。有限数量的解决方案所需的最小图像数量为 N = 3。对于这种情况，通过将未知数与测量数进行比较，

我们得出结论，特征的数量必须至少为 M = 2。另一方面，如果只有一个特征可用（M = 1），我们得到条件 N ≥ 4，作为图像数量的下限（附录 B.2.2)。

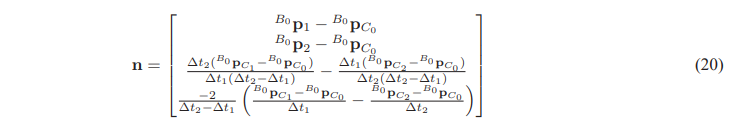
From the above we see that two minimal problems can be identified: (i) two features seen in three images, and (ii) one feature seen in four images. Both of these problems are minimal, since in the first case we use the minimum number of images possible, while in the second case we use the minimum number of features. In what follows, we describe the properties of these problems

从上面我们看到可以确定两个最小的问题：（i）在三个图像中看到两个特征，以及（ii）在四个图像中看到一个特征。这两个问题都很小，因为在第一种情况下，我们使用了尽可能少的图像数量，而在第二种情况下，我们使用了最少数量的特征。在接下来的内容中，我们描述了这些问题的性质

### Two features seen in three images

In Appendix B.2.3, we prove that when two features are observed in three images, then in general it is dim(N (A)) = 1. The only exception is when the three camera positions and the two feature points lie in the same plane, in which case 7 dim(N (A)) = 2. Barring this singular configuration, the nullspace of the matrix A is spanned by

在附录 B.2.3 中，我们证明了当在三个图像中观察到两个特征时，一般情况下 dim(N (A)) = 1。唯一的例外是当三个相机位置和两个特征点位于同一平面，在这种情况下 7 dim(N (A)) = 2。除非有这种奇异配置，否则矩阵 A 的零空间由



From the above expression, we see that if the camera is moving with a constant velocity, then the last three elements of n are zero, and the scale of the trajectory is undefined, while the attitude is uniquely determined. In all other cases, two distinct solutions exist.

由上式可知，如果相机匀速运动，则n的后三个元素为零，轨迹的尺度是不确定的，而姿态是唯一确定的。在所有其他情况下，存在两种不同的解决方案

### One feature seen in four images

By following a similar analysis in Appendix B.2.4, we show that the results for the minimal problem where one feature is seen in four different images are analogous to the previous case. The results for both cases can be summarized as follows:

通过遵循附录 B.2.4 中的类似分析，我们表明，在四个不同图像中看到一个特征的最小问题的结果与前一种情况类似。两种情况的结果可以总结如下：

**Result 5**. When two features are observed in three images or one feature is observed in four images, then we can compute two distinct solutions for all observable quantities, using Eq. (18). There exist two singular cases when an infinite number of solutions exist: (i) when all camera positions and the feature(s) lie in the same plane, and (ii) when the camera is moving with a constant velocity. In the latter case, the orientation is uniquely defined, but the scale is unobservable

结果 5. 当在三幅图像中观察到两个特征或在四幅图像中观察到一个特征时，我们可以使用方程式计算所有可观察量的两个不同解。 (18)。当存在无限数量的解决方案时，存在两种奇异情况：（i）当所有相机位置和特征位于同一平面时，(ii) 当相机以恒定速度移动时。在后一种情况下，方向是唯一定义的，但规模是不可观察的。

Fig. 1 depicts the two solutions arising in an example case where one feature is observed in four images. It is interesting to observe that, even though this is difficult to see in the plot, the two trajectories are not coplanar. As a final remark, we note that if either more images, or more features become available in either one of the two minimal scenarios discussed, then the problems become over-determined, and thus a unique solution can be obtained in general (barring singular cases, such as the one described in Result 3). This makes it possible to determine the conditions for obtaining a unique solution.

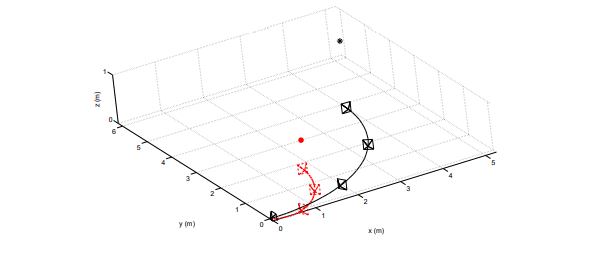
图 1 描述了在四个图像中观察到一个特征的示例情况下出现的两种解决方案。有趣的是，尽管这在情节中很难看到，但两条轨迹并不共面。最后，我们注意到，如果有更多图像，在所讨论的两个最小场景中的任何一个中，一个或多个特性变得可用，然后问题变得过度确定，因此通常可以获得唯一的解决方案（除非是特殊情况，例如结果 3 中描述的情况）。这使得可以确定获得唯一解的条件。

Figure 1: One feature observed in four camera images: This plot shows the two possible solutions for the trajectory and the feature position, both of which satisfy the measurements exactly. The two trajectories are not coplanar.

图 1：在四个摄像机图像中观察到的一个特征：该图显示了轨迹和特征位置的两种可能解决方案，两者都完全满足测量结果。两条轨迹不共面。

# Unknown biases and/or IMU-Camera calibration

Up to this point, we have examined the case where the extrinsic calibration between the IMU and the camera, as well as the IMU biases, are known in advance. We here study the scenario where one or both of these quantities are not kwown, and need to be estimated. As shown in [1], both the extrinsic calibration as well as the IMU biases are observable, for general camera trajectories

到目前为止，我们已经检查了 IMU 和相机之间的外部校准以及 IMU 偏差是预先知道的情况。我们在这里研究这些数量中的一个或两个都不知道并且需要估计的情况。如[1]所示，对于一般相机轨迹，外部校准和 IMU 偏差都是可观察到的

## Unknown extrinsic calibration, known IMU biases

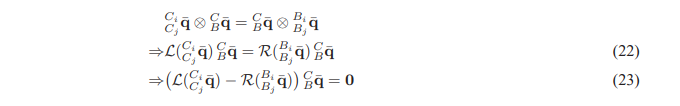
In this scenario, we consider the case where the transformation (rotation and translation) between the IMU and camera is not available while the IMU biases are known. By using the corrected IMU measurements, we can estimate the IMU orientation change Bi BjC in the time interval [ti , tj ], as in Eq. (3). In addition, by using only feature observations, we can 8 independently estimate the relative camera orientation during that interval, Ci CjC, by using image-based motion estimation algorithms such as [8]. We can then employ the following equation for the unknown C BC:

在这种情况下，我们考虑 IMU 和相机之间的转换（旋转和平移）不可用而 IMU 偏差已知的情况。通过使用校正后的 IMU 测量值，我们可以估计 IMU 方向变化 在时间间隔 中，如等式(3)。此外，通过仅使用特征观察，我们可以 8 通过使用基于图像的运动估计算法（例如 [8]）独立估计该时间间隔内的相对相机方向 。然后，我们可以对未知的 使用以下等式：



The above equation can be solved by transforming it into its equivalent unit-quaternion representation [16, 17]. Specifically, using this representation, we ha

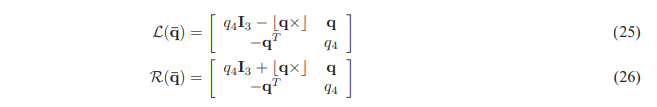
上述方程可以通过将其转换为等效的单位四元数表示来求解 [16, 17]。具体来说，使用这种表示，我们有



where for a 4 × 1 unit quaternion ¯q:



we define：



Therefore, we have a linear system with C Bq¯, the quaternion representing the IMU-camera rotation, as the unknown. By using all the available pairs of images, we construct an over-constrained linear system B C Bq¯ = 0, whose least-squares solution C Bq¯ is the right singular vector corresponding to the smallest singular value of B. For the solution to be unique, at least two pairs of images where the system rotates about different axes are required [17]. Therefore, at least three images are needed, in which at least six features (to ensure a unique solution for Ci CjC) are tracked. After C Bq¯ is estimated, we can directly obtain C BC [18]. Subsequently, by treating C pB in Eq. (9) as an extra unknown, we can construct a similar linear system as in Section 3, except that the vector of unknowns becomes (see Appendix A.2):

因此，我们有一个线性系统，其中 表示 IMU 相机旋转的四元数作为未知数。通过使用所有可用的图像对，我们构造了一个过约束的线性系统 ，其最小二乘解 是对应于 B 的最小奇异值的右奇异向量。

为了使解决方案独一无二，至少需要两对系统围绕不同轴旋转的图像[17]。因此，至少需要三个图像，其中至少跟踪六个特征（以确保 的唯一解决方案）。估计 后，我们可以直接得到 [18]。随后，

通过处理方程式中的 。 (9) 作为一个额外的未知数，我们可以构造一个与第 3 节类似的线性系统，只是未知数的向量变为（见附录 A.2）：

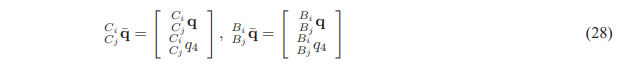


By solving this modified linear system, we can solve for the observable quantities O1-O3 discussed in Section 2 as well as the IMU-camera relative translation

通过求解这个修改后的线性系统，我们可以求解第 2 节中讨论的可观察量 O1-O3 以及 IMU 相机的相对平移

## Unknown extrinsic calibration, unknown IMU biases

We next consider the case where we have no prior information about either the IMU-camera transformation or the IMU biases. We here assume that the gyroscope and accelerometer biases are constant over the short time interval used for initialization. Similar to the previous subsection, by using only feature observations, we can estimate the relative camera pose change Ci CjC in the time interval [ti , tj ]. Let Ci Cj q¯ and Bi Bj q¯ be the quaternions that represent the relative pose changes of the camera and IMU, respectively:

我们接下来考虑没有关于 IMU 相机转换或 IMU 偏差的先验信息的情况。我们在这里假设陀螺仪和加速度计偏差在用于初始化的短时间间隔内是恒定的。与上一小节类似，仅使用特征观察，我们可以估计时间间隔 内的相对相机位姿变化 。令 和 分别为表示相机和 IMU 相对位姿变化的四元数：

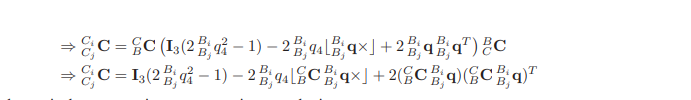
To derive an equation that will enable us to obtain a closed-form solution, we will employ the relationship between a quaternion and its corresponding rotation matrix [18]:

为了推导出使我们能够获得封闭形式解的方程，我们将使用四元数与其对应的旋转矩阵 [18] 之间的关系：

Using basic properties of rotation matrices, we have:

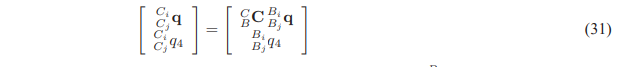
使用旋转矩阵的基本属性，我们有：





Converting to the equivalent quaternion representation, we obtain:

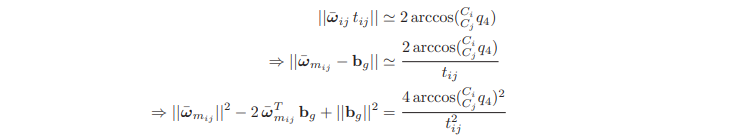
转换为等效的四元数表示，我们得到：



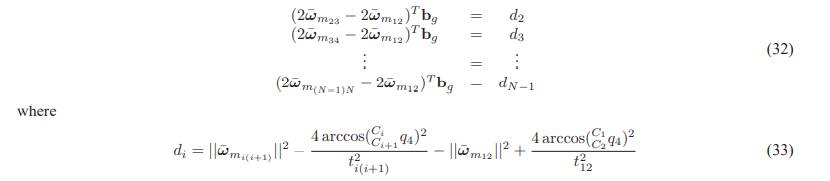
The above equation shows that the last element of the quaternion of the true IMU rotation (Bi Bj q4) can be estimated from feature observations (Ci Cj q4). If we let ω¯ ij denote the average true rotational velocity in the time interval [ti , tj ], and ω¯ mij represent the average measured rotational velocity in the same time, then, using the results of [18], we can write:

上式表明，真实 IMU 旋转的四元数的最后一个元素 可以从特征观测值 中估计出来。如果我们让表示时间间隔 内的平均真实旋转速度，并且 表示同时测量的平均旋转速度，那么，

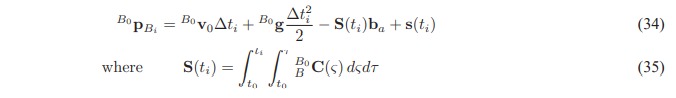
使用[18]的结果，我们可以写：



where . Choosing , we obtain (N − 1) equations of the above form. By subtracting the first one from all other equations, we get the linear system:

其中 。选择 j = i + 1, ∀i = 1。 . .(N - 1)，我们得到上述形式的 (N - 1) 个方程。通过从所有其他方程中减去第一个方程，我们得到线性系统：

Note that all quantities except for bg are known from the IMU and camera measurements. Therefore, the bias in the gyroscope measurements, bg, can be estimated by solving for the least-squares solution of the linear system constructed in Eq. (32). For a unique solution, at least five images are needed. After we have computed the gyro bias, we can obtain the corrected gyroscope measurements, and proceed to solve for the IMU-camera rotation matrix as presented in Section 4.1. Subsequently, we can solve for the remaining quantities linearly. Specifically, by taking into account the accelerometer bias, we rewrite Eq. (7) as follows (see Appendix A):

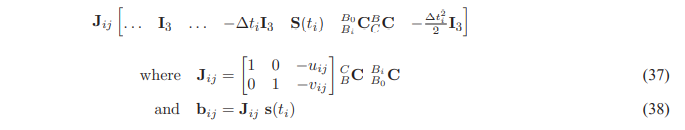
请注意，除了 之外的所有量都可以从 IMU 和相机测量中得知。因此，陀螺仪测量中的偏差 可以通过求解方程式中构造的线性系统的最小二乘解来估计。 (32)。对于一个独特的解决方案，至少需要五个图像。在我们计算了陀螺偏差之后，我们可以获得校正后的陀螺仪测量值，并继续求解 IMU 相机旋转矩阵，如第 4.1 节所示。随后，我们可以线性求解剩余量。具体来说，通过考虑加速度计偏差，我们重写方程。 (7) 如下（见附录A）：

By treating C pB and ba as unknowns, we can construct a similar linear system as in Section 3, except that the vector x in Eq. (11) is extended as (see Appendix A.3):

通过将 和 视为未知数，我们可以构造与第 3 节中类似的线性系统，除了等式中的向量 x。 (11) 扩展为（见附录 A.3）：



and becomes



Solving this modified linear system, we can estimate the accelerometer bias and IMU-camera relative translation in addition to the observable quantities O1-O3 in Section 2

求解这个修改后的线性系统，除了第 2 节中的可观察量 O1-O3 之外，我们还可以估计加速度计偏差和 IMU 相机相对平移

# Solution in the presence of noise