2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Enhancing Visual Inertial SLAM with Magnetic Measurements

Bharat Joshi and Ioannis Rekleitis

***Abstract*— This paper presents an extension to visual inertial odometry (VIO) by introducing tightly-coupled fusion of magnetometer measurements. A sliding window of keyframes is optimized by minimizing re-projection errors, relative inertial errors, and relative magnetometer orientation errors. The results of IMU orientation propagation are used to efficiently transform magnetometer measurements between frames producing relative orientation constraints between consecutive frames. The soft and hard iron effects are calibrated using an ellipsoid fitting algorithm. The introduction of magnetometer data results in significant reductions in the orientation error and also in recovery of the true yaw orientation with respect to the magnetic north. The proposed framework operates in all environments with slow-varying magnetic fields, mainly outdoors and underwater. We have focused our work on the underwater domain, especially in underwater caves, as the narrow passage and turbulent flow make it difficult to perform loop closures and reset the localization drift. The underwater caves present challenges to VIO due to the absence of ambient light and the confined nature of the environment, while also being a crucial source of fresh water and providing valuable historical records. Experimental results from underwater caves demonstrate the improvements in accuracy and robustness introduced by the proposed VIO extension.**

I. Introduction

Magnetic measurements are an often neglected source of information mainly because of their sensitivity to ambient noise; however, there are several situations in which they can provide useful information with minimal cost and low computational overhead. In this work we are targeting the underwater domain with the use of a sensor suite [1]. The used sensor suite, comprised of a stereo camera, a 9-axis IMU, a water depth sensor, and a pencil beam mechanically scanning sonar, can be deployed by a human to collect data inside an underwater cave; see Fig. 1. We propose a tightly coupled optimization-based fusion of visual, inertial, and magnetometer information. Magnetometer measurements are added as new factors to the keyframe-based sliding window optimization graph proposed in [2]. We leverage the IMU preintegration algorithm from [3], [4] to efficiently compute magnetometer residuals for all measurements. Since the IMU preintegration terms are already defined in the relative inertial error, the additional computational cost is relatively small. The magnetometer data are sensitive to the local magnetic field and require an explicit calibration procedure. As such, the magnetometer measurements are calibrated to account for soft and hard iron effects using an ellipsoid fitting algorithm. Experiments have shown that after calibration, the local magnetic noise is relatively low throughout the trajectory. Furthermore, magnetic data present an absolute measurement based on the magnetic field of Earth and while noisy they are consistent over long trajectories. Our target application is the mapping of underwater caves.

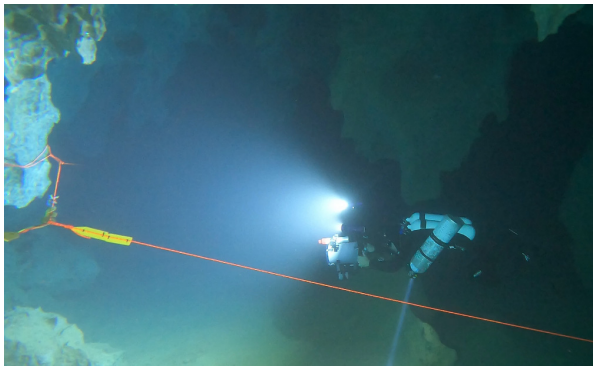

Fig. 1: Sensor suite deployed inside the Cueva del Agua, Spain.

Mapping underwater cave systems is extremely important for environmental protection, fresh water management [5], and resource utilization [6]. Moreover, caves provide valuable historical evidence as they present an undisturbed time capsule [7]–[10], and information about geological processes [11]. Diver centered mapping is a dangerous, labor-intensive, tedious, and slow task. In addition to limited visibility, color absorption, hazing, and lightning variations, there is no ambient light and the environment is often cluttered and fragile, making navigation difficult, with inadvertent motion maneuvers often reducing visibility to zero. Furthermore, the confined environment prevents resurfacing in case of a problem or emergency. During operations inside underwater caves, there are two constraining factors. First, the passages are often narrow, and the view on the way in is completely different from the view on the way out, making maneuvering difficult. Second, there is significant water flow that makes staying in one spot to collect data in 360 degrees very challenging. These conditions contribute to reducing the number of loop closures that are feasible.

Moreover, underwater cave systems provide an exciting opportunity to use magnetometer measurements as they are devoid of significant magnetic disturbances due to ferromagnetic materials. The introduction of magnetic field measurements in the state estimation process, proposed in

The authors are with the Department of Computer Science and Engineering, University of South Carolina, Columbia, SC 29208, USA. bjoshi@email.sc.edu, yiannisr@cse.sc.edu.

This research has been supported in part by the National Science Foundation under grants 1943205 and 2024741. The authors also acknowledge the help of the Woodville Karst Plain Project (WKPP), El Centro Investigador del Sistema Acuífero de Quintana Roo A.C. (CINDAQ), Ricardo Constantino, Project Baseline, and Evan Kornacki in collecting data, providing access to challenging underwater caves, and mentoring us in underwater cave exploration. The authors are also grateful for equipment support by Halcyon Dive Systems, Teledyne FLIR LLC, and KELDAN GmbH lights.

this paper, leads to two significant contributions. First of all, absolute orientation measurements are introduced to orient the produced trajectory with respect to the magnetic north. This makes the produced trajectories compatible with the existing man-made maps of the caves. Second, the introduction of magnetometer data constrains the produced trajectories along the yaw direction, producing much more consistent results, eliminating the orientation drift that plagued earlier deployments [12].

To the best of our knowledge, our work is the first to use IMU preintegration to introduce high frequency magnetometer measurements in the optimization framework. The proposed framework has been tested in a variety of underwater caves in Florida and Mexico. Qualitative results demonstrate the alignment of the produced trajectories to the existing maps of the caves, and also a significant reduction in orientation error (drift). The produced trajectories resulted in significantly reduced error, compared to the baseline trajectories obtained using COLMAP [13] (a bundle adjustment, global optimization framework).

II. Related Work

Vision-based state estimation has proven to be extremely challenging due to the varying lightning conditions, scattering and haziness from floating particulates, and color variations due to light absorption in the water. Several visual and visual-inertial state estimation packages have exhibited severe failures [14], [15]. Rahman *et al*. [16] proposed an extension of OKVIS [2] incorporating a pencil beam mechanically scanning sonar, a water depth sensor, and loop closure capabilities [17], [18]. The framework was then adapted for inexpensive action cameras, GoPro 9, [19] producing superior performance. More recent packages such as ORB-SLAM3 [20], OpenVINS [21], *etc*. have demonstrated better performance, but there is no in-depth analysis of their accuracy and robustness.

Magnetometers have traditionally been fused with acceleration and angular velocity measurements to estimate the vehicle's attitude; such systems are termed attitude and heading reference systems (AHRS) [22]–[24]. Some recent work on the fusion of magnetic field measurements with a visual sensor has gained traction. Wang *et al*. [25] performed visual, inertial, and magnetometer fusion first by initializing the VIO in the reference frame of Earth and then computing the error between magnetometer measurement and Earth's magnetic field. The authors included only magnetometer measurements near the keyframe, skipping most of the high-frequency magnetometer measurements. Siebler *et al*. proposed a particle filter method to use magnetic field distortions for magnetometer calibration and localization [26]; without focusing on continuous state estimation. The closest approach to ours used inter-frame preintegrated magnetometer measurements [27], [28]. However, it requires an array of magnetometers measuring magnetic field and it's gradient. Our work is focused on a single 9-axis IMU without requiring additional hardware enhancements.

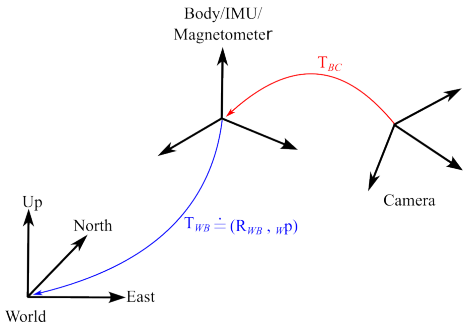

Fig. 2: The body pose $T_{WB} \doteq (R_{WB}, {}_W\mathbf{p})$ which coincides with 9-axis IMU is tracked w.r.t the world frame W. The camera pose in body frame T_{BC} is fixed and known from prior calibration. The world frame W coincides with East, North and Up direction in Earth coordinates.

The calibration of soft and hard iron offsets for magnetometer is very important as they can introduce huge errors in the estimation process. In [29] an iterative batch least squares estimation was introduced to calibrate a full three-axis magnetometer. Vasconcelos *et al*. used maximum likelihood estimation to find the optimal calibration parameters that best fit the reading of the onboard sensors to calibrate the 3-axis strapless magnetometer [30]. Kok *et al*. [31] proposed maximal likelihood calibration using orientation estimation from inertial sensors. Taking things to the next step, Solin *et al*. created a map of the magnetic field in an indoor environment by imposing a Gaussian process (GP) prior to the latent scalar potential of the magnetic field [32].

III. Visual-Inertial-Magnetometer Fusion

Visual-Inertial Odometry aims to estimate the pose of the moving camera-imu system by fusing inertial information with images. In this work, we extend visual-inertial odometry to fuse magnetometer measurements in a tightly coupled fashion using IMU preintegration.

A. Problem Formulation

We consider the visual inertial magnetometer odometry problem where we want to track the *state* of a sensing system (a handheld underwater sensor suite) equipped with an IMU, a stereo/monocular camera, and a magnetometer. The IMU frame coincides with the body frame "B" we want to track, and the transformation between the camera and the IMU (T_{BC}) is known by calibrating the extrinsic parameters and remains constant throughout the experiment; see Fig. 2 . In this work, we use a 9-axis IMU that includes an accelerometer, a gyroscope, and a magnetometer. As such, we assume that the magnetometer and IMU (gyroscope and accelerometer) axes are aligned and that there is no significant axis misalignment. When using stereo cameras, we assume that they are rigidly attached with known extrinsic parameters between the cameras and between the cameras and the IMU.

Nonlinear keyframe-based sliding window optimization is performed to estimate body poses and 3D landmark positions by minimizing the reprojection error of landmarks seen in the camera frame. Thus, we need to estimate the variables

$\mathcal{X} = \{\mathcal{X}_\mathrm{B}, \mathcal{L}\}$ where $\mathcal{L}$ represents 3D landmark positions visible in the sliding window and $\mathcal{X}_\mathrm{B} = [\mathbf{x}_1, ..., \mathbf{x}_K]$ holds the system states at camera times 1 to K with K being the total number of keyframes in the sliding window. The system state $\mathbf{x}_k$ at time t_k holds position ${}_\mathrm{W}\mathbf{p}^k$ and orientation $\mathbf{q}^k_\mathrm{WB}$ in the world frame, velocity in the inertial frame ${}_\mathrm{W}\mathbf{v}^k$, as well as biases of the gyroscope $\mathbf{b}^k_g$ and accelerometer $\mathbf{b}^k_a$. The system state can be written as

$$\mathbf{x}_k = [{}_\mathrm{W}\mathbf{p}^k, \mathbf{q}^k_\mathrm{WB}, {}_\mathrm{W}\mathbf{v}^k, \mathbf{b}^k_g, \mathbf{b}^k_a] \tag{1}$$

where $\mathbf{q}_\mathrm{WB}$ is the quaternion representation of the orientation R_WB.

B. Fusing visual, inertial and magnetometer measurements

The keyframe-based visual-inertial SLAM is formulated as joint nonlinear optimization that maximizes the posterior probability of the system state $\mathcal{X}$. Using the problem formulation proposed in [2], [18], the following cost function is minimized

$$J_{VI}(\mathcal{X}) = \sum_{k=1}^{K} \sum_{j \in \mathcal{J}_k} \|\mathbf{e_r}^{j,k}\|^2_{\mathbf{W_r}^{j,k}} + \sum_{k=1}^{K} \|\mathbf{e_i}^k\|^2_{\mathbf{W_i}^k} + \|\mathbf{e_p}\|^2 \tag{2}$$

where k denotes the camera frame index and j denotes the landmark index. The cost function contains the reprojection $\mathbf{e_r}$, inertial $\mathbf{e_i}$, and marginalization $\mathbf{e_p}$ residuals weighted by their respective information matrices.

The reprojection error for the landmark ${}_W\mathbf{l}_j \in \mathcal{J}_k$ is calculated as $\mathbf{e_r}^{i,j} = h({}_W\mathbf{l}_j) - \mathbf{z}^{j,k}$ where $\mathcal{J}_k$ is the set of all the landmarks visible in the keyframe k. Here, $h(.)$ denotes the camera projection model and $\mathbf{z}^{j,k}$ the measurements in the image coordinates. For more details, please refer to [2]. The inertial residuals $\mathbf{e_i}$ are obtained using the IMU preintegration theory proposed in [3], [33] which is detailed in Section III-C. We employ a marginalization strategy similar to [2] to obtain the marginalization prior error term, $\mathbf{e_p}$. Whenever a new frame is inserted in the optimization window, the marginalization operation is classified into two cases. If the oldest frame in the sliding window is not a keyframe, it is marginalized together with the oldest speed and bias states, and all its landmark measurements are dropped to maintain sparsity. In the case where the oldest frame is a keyframe, only the landmarks that are visible in that frame but not in the most recent keyframe are marginalized out.

We introduce the magnetometer residuals based on IMU preintegration into Eq. 2 to obtain the cost function used in this work as

$$J(\mathcal{X}) = J_{VI}(\mathcal{X}) + \sum_{k=1}^{K} \sum_{j \in \mathcal{M}_k} \|\mathbf{e_m}^{j,k}\|^2_{\mathbf{W_m}^k} \tag{3}$$

where $\mathcal{M}_k$ denotes all magnetometer measurements attached to the system state $\mathbf{x}_k$ by the error term. In the remainder of Section III, we use the output of the IMU preintegration algorithm to derive the magnetometer error term $\mathbf{e_m}^{j,k}$ and the residual weights $\mathbf{W_m}^k$.

C. IMU Preintegration

In this section, we review the IMU preintegration formulation, which in turn will be used to derive the magnetometer residuals in Section III-D. The IMU preintegration formulation is based on [3] inspired by continuous-time quaternion kinematics from [33] and uses IMU bias manipulation according to [4].

The accelerometer and gyroscope measurements in body frame B at time t are affected by the additive white noise $\boldsymbol{\eta}$ and slowing varying bias $\mathbf{b}$

$$\begin{aligned} {}_\mathrm{B}\tilde{\boldsymbol{w}}(t) &= {}_\mathrm{B}\boldsymbol{w}(t) + \mathbf{b}_g(t) + \boldsymbol{\eta}_g \\ {}_\mathrm{B}\tilde{\mathbf{a}}(t) &= \mathrm{R}^\mathrm{T}_\mathrm{WB}({}_\mathrm{W}\mathbf{a}(t) - {}_\mathrm{W}\mathbf{g}) + \mathbf{b}_a(t) + \boldsymbol{\eta}_a \end{aligned} \tag{4}$$

The accelerometer and gyroscope noise is modeled as additive Gaussian noise with $\boldsymbol{\eta}_a \sim \mathcal{N}(\mathbf{0}, \sigma_a^2 \cdot \mathbf{I})$ and $\boldsymbol{\eta}_w \sim \mathcal{N}(\mathbf{0}, \sigma_w^2 \cdot \mathbf{I})$, where $\mathbf{I}$ being the identity matrix. IMU biases are modeled as a slowly varying random walk with $\dot{\mathbf{b}}_{w_t} = \boldsymbol{\eta}_{b_w}$ and $\dot{\mathbf{b}}_{a_t} = \boldsymbol{\eta}_{b_a}$ where $\boldsymbol{\eta}_{b_w} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{b_w} \cdot \mathbf{I})$ and $\boldsymbol{\eta}_{b_a} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{b_a} \cdot \mathbf{I})$. Most modern 9-axis IMUs provide high-frequency synchronized accelerometer, gyroscope, and magnetometer measurements, as shown in Fig. 3. As such, ${}_\mathrm{W}\mathbf{p}^i, \mathbf{q}^i_\mathrm{WB}$ and ${}_\mathrm{W}\mathbf{v}^i$ can be propagated in the time interval $[t_k, t_{k+1}]$ using accelerometer and gyroscope measurements. Although in practice IMU measurements may not perfectly synchronize with image timestamps, the sensor setup is calibrated using Kalibr [34] beforehand. Propagation in the world frame requires knowledge of the initial state of the system at time t_k. Whenever the system state changes during optimization, repropagation is needed. Thus, the IMU preintegration is performed in the body frame to avoid repropagation whenever system state is updated during optimization for computational efficiency.

The propagation is performed in body frame B_k instead of the world frame as

$$\begin{aligned} \mathrm{R}^k_\mathrm{BW}\ {}_\mathrm{W}\mathbf{p}^{k+1} &= \mathrm{R}^k_\mathrm{BW}({}_\mathrm{W}\mathbf{p}^k + {}_\mathrm{W}\mathbf{v}^j\Delta t_k - \frac{1}{2}{}_\mathrm{W}\mathbf{g}{\Delta t_k}^2) + \boldsymbol{\alpha}^{k+1}_k \\ \mathrm{R}^k_\mathrm{BW}\ {}_\mathrm{W}\mathbf{v}^{k+1} &= \mathrm{R}^k_\mathrm{BW}({}_\mathrm{W}\mathbf{v}^k - {}_\mathrm{W}\mathbf{g}\Delta t_k) + \boldsymbol{\beta}^{k+1}_k \\ \mathbf{q}^k_\mathrm{BW} \otimes \mathbf{q}^{k+1}_\mathrm{WB} &= \boldsymbol{\gamma}^{k+1}_k \end{aligned} \tag{5}$$

where $\boldsymbol{\alpha}^{k+1}_k$, $\boldsymbol{\beta}^{k+1}_k$ and $\boldsymbol{\gamma}^{k+1}_k$ are the preintegration terms which only depend on the inertial measurements and biases. These preintegration terms along with their covariance ${}_\mathrm{B}\mathbf{P}^{k+1}_k$ can be updated iteratively using discrete accelerometer and gyroscope measurements; see [3], [4] for details. The preintegrated terms $\boldsymbol{\alpha}^{k+1}_k$, $\boldsymbol{\beta}^{k+1}_k$ and $\boldsymbol{\gamma}^{k+1}_k$ are updated using the first-order approximation with respect to their biases if the bias estimates are relatively unchanged as proposed in [4]. Otherwise, we redo the propagation using the updated bias estimates. The inertial error $\mathbf{e_i}^k$ is obtained from Eq. 5 as

$$\mathbf{e_i}^k = \begin{bmatrix} \mathrm{R}^k_\mathrm{BW}({}_\mathrm{W}\mathbf{p}^{k+1} - {}_\mathrm{W}\mathbf{p}^k - {}_\mathrm{W}\mathbf{v}^j\Delta t_k + \frac{1}{2}{}_\mathrm{W}\mathbf{g}{\Delta t_k}^2) - \boldsymbol{\alpha}^{k+1}_k \\ \mathrm{R}^k_\mathrm{BW}({}_\mathrm{W}\mathbf{v}^{k+1} - {}_\mathrm{W}\mathbf{v}^k + {}_\mathrm{W}\mathbf{g}\Delta t_k) - \boldsymbol{\beta}^{k+1}_k \\ 2[(q^k_\mathrm{WB})^{-1} \otimes q^{k+1}_\mathrm{WB} \otimes (\boldsymbol{\gamma}^{k+1}_k)^{-1}]_{xyz} \\ \mathbf{b}^{k+1}_g - \mathbf{b}^k_g \\ \mathbf{b}^{k+1}_a - \mathbf{b}^k_a \end{bmatrix}. \tag{6}$$

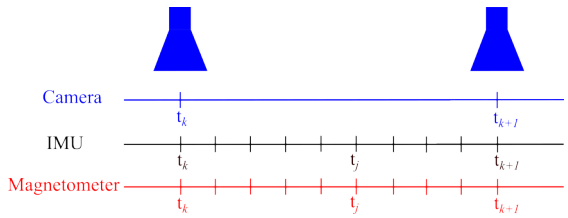

Fig. 3: Illustration of visual, inertial and magnetometer measurement timing

The preintegration term allows us to estimate the body orientation at time $j \in (k, k+1]$ during recursive updates which we will use to obtain the magnetometer error in the next section.

D. Magnetometer Residuals

Magnetometer and IMU measurements obtained from 9-axis IMU used in attitude and heading reference systems (AHRS) are synchronized as shown in Fig. 3. The uncertainty of magnetometer measurements is modeled as Gaussian noise such that the magnetometer measurement is

$$ {}_{\mathrm{B}}\tilde{\mathbf{m}} = \mathrm{R}^{\mathrm{T}}_{\mathrm{WB}}\ {}_{\mathrm{W}}\mathbf{m} + \boldsymbol{\eta}_m \tag{7} $$

where ${}_{\mathrm{W}}\mathbf{m}$ is the earth's magnetic field at the location in world coordinate frame which is aligned with earth's east, north and up (ENU) direction. The magnetometer additive white noise can be expressed as $\boldsymbol{\eta}_m \sim \mathcal{N}(\mathbf{0}, \sigma_m^2 \cdot \mathbf{I})$. We estimate the magnetometer white noise similar to IMU [35] using the allan-variance plot.

However, the magnetometer measures the superposition of Earth's magnetic field and local magnetic field because of the presence of magnetic materials in the sensor's vicinity. Thus, the magnetometer is calibrated before experiments to estimate the soft and hard iron effects. The calibration procedure is explained in the next section. During VIO initialization, we align the orientation of the body frame in the ENU direction. When properly calibrated, the Earth's magnetic field points in the north-down direction. First, the upward-pointing gravity direction is obtained from the accelerometer measurements. The cross product of the magnetic field and the gravity direction provides the east direction. The remaining axis can be obtained from the cross product of the estimated axes. This is a common procedure used in AHRS systems such as the Madgwick filter [22] and the complementary filter [23]. Formulating the magnetometer residual directly in terms of Eq. 4 has one significant drawback that any error during the initial alignment is also included in the estimation process. Thus, we formulate the magnetometer residuals as relative orientation constraint between two consecutive frames.

For the magnetometer residual formulation, we assume that the magnetometer samples are calibrated following the procedure in Section III-E. Using the preintegrated orientation term allows us to use all magnetometer measurements between consecutive frames. Given the current consecutive frame state estimates $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$, we define the magnetometer residual for measurement at time $t_j \in (t_k, t_{k+1}]$ as

$$ {\mathbf{e}_{\mathbf{m}}}^{j,k} = \mathrm{R}^k_{\mathrm{BW}}\ \mathrm{R}^{k+1}_{\mathrm{WB}}\ {}_{\mathrm{B}}\tilde{\mathbf{m}}^{k+1} - \Delta\mathrm{R}^j_k\ {}_{\mathrm{B}}\tilde{\mathbf{m}}^j \tag{8} $$

where $\Delta\mathrm{R}^j_k$ is the propagated orientation between the time interval $[t_k, t_j]$ and transforms the magnetometer measurement from time j to k. This propagated orientation is the rotation matrix representation of the intermediate preintegration result $\boldsymbol{\gamma}^j_k$. Rotation $\mathrm{R}^k_{\mathrm{BW}}\ \mathrm{R}^{k+1}_{\mathrm{WB}}$ represents relative orientation during the time interval $[t_k, t_{k+1}]$ in the body frame B_k and transforms the magnetometer measurement from time t_{k+1} to t_k. By expressing the error term in the body frame, we can recursively calculate $\boldsymbol{\gamma}^j_k$ efficiently as $t_j < t_{k+1}$. Thus, the magnetometer error using preintegrated orientation allows us to use all the magnetometer measurements in the time interval $(k, k+1]$ to optimize the relative orientation between consecutive frames. The magnetometer residual does not depend on the position estimate; as such, it does not directly affect the state's position estimate. However, a better orientation estimate will certainly yield a better position accuracy in the long run.

The magnetometer measurement estimated at t_k using quaternion preintegration depends on the gyroscope noise. However, the gyroscope noise is already considered while computing $\boldsymbol{\gamma}^j_k$ and is significantly smaller than the magnetometer noise. Also, we found that the propagated orientation variance was very small compared to the magnetometer variance. Thus, for simplicity and computational efficiency, we assume that the magnetometer residual weights depend only on the magnetometer sensor noise. Moreover, when the state connected to the magnetometer residuals is marginalized, the magnetometer residuals are added as the marginalization prior along with the inertial and visual residuals.

E. Soft and Hard Iron Effect Calibration

The full magnetometer measurement model taking into account the magnetic disturbances and sensor imperfections can be modeled [30] as

$$ {}_{\mathrm{B}}\hat{\mathbf{m}} = \mathrm{S}\mathrm{R}^{\mathrm{T}}_{\mathrm{WB}}\ {}_{\mathrm{W}}\mathbf{m} + \mathbf{h} + \boldsymbol{\eta}_m \tag{9} $$

where S represents the soft iron matrix and $\mathbf{h}$ represents the hard iron effect. Hard iron effects arise because of the permanent magnetization of the material and also depend on the fixed sensor recording mechanism. In particular, we can think of the hard iron effect as a constant bias $\mathbf{h}$. Soft iron effects are due to the magnetization of ferromagnetic materials due to the local magnetic field and depend on the orientation with respect to the local magnetic field [31]. The soft iron effect can be represented as a 3×3 symmetric matrix S. We do not consider other sources of magnetometer errors that could arise from non-orthogonality of the magnetometer axes or differences in sensitivity along the three magnetometer axes [29].

Without the soft and hard iron offset, if we rotate a magnetometer sensor then the magnetic field vector falls on the surface of sphere with a radius equal to the magnitude of the earth's magnetic field. Eq. 9 can be seen as a linear transformation that maps points on a sphere to an ellipsoid [36]. Thus, the magnetometer calibration problem can be interpreted as ellipsoid fitting of the points to the sphere surface. We solve the ellipsoid fit problem detailed

in Kok *et al.* [31] to obtain the estimate of $A = S^{-1}$ and the hard iron offset $\mathbf{h}$. For correct calibration, the magnetometer needs to be rotated along all three axes several times. Magnetometer measurements are corrected in the body frame as follows:

$$ {}_B\tilde{\mathbf{m}} = A({}_B\hat{\mathbf{m}} - \mathbf{h}) \tag{10} $$

These corrected measurements are used to calculate magnetometer residuals in the sliding-window optimization as explained in Section III-D. In an indoor environment, these hard and soft iron effects continue to change as we move through the different areas. The underwater cave environment is ideally suited for magnetometer fusion as the likelihood of magnetic disturbance is minimal. As such, we performed the calibration at the beginning and assumed that the local magnetic field did not change significantly. During deployment, we performed calibration at the beginning and at the end of the dive, both calibrations resulted in similar values, validating the above assumption. On older datasets, we did not record an explicit magnetometer calibration sequence. In such a case, these are not enough magnetometer measurements in all directions; thus full mapping from ellipse to sphere is degenerate. Thus, we only calibrate the hard iron effect. Figure 4 presents the ellipsoids fitted during complete calibration for the new sensor and for partial calibration of the old sensor.

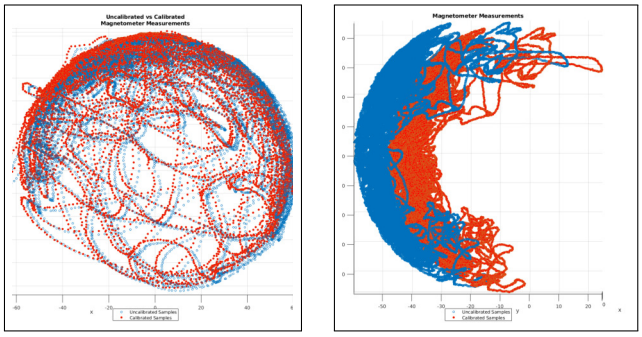

Fig. 4: Magnetometer calibration ellipses fitted from explicit calibration sequence (left) and partial calibration (right).

IV. Experimental Results

We present evaluation results from two experiments in the Devil's Eye cave system, FL, USA and one experiment in the Dos Ojos Cenote, QR, Mexico. In all experiments the diver enters and exits the cave from the same location.

A. Datasets

All the datasets are collected utilizing the sensor suite described in Rahman *et al.* [1]. The first dataset was collected using IDS UI-3251LE cameras in a stereo configuration and Microstrain 3DM-GX5-25 AHRS in the Devil's cave system, Florida. We call this dataset cave1_short and images are captured at 15Hz. The cave1_short dataset spans approximately 220m. The second dataset was collected with the same sensor suite in Dos Ojos Cenote, QR, Mexico. This trajectory is very long (∼680m) with a duration of more than 50 minutes. COLMAP was unable to finish the complete trajectory due to computational and memory constraints. Thus, this data set is used only for qualitative evaluation.

The third dataset was collected using a similar sensor suite in monocular setup with Flir BFS-U3-16S7C-C camera and microstrain 3DM-GV7 AHRS. We call this dataset cave1_long with total trajectory length of approximately 506m. Both datasets are very challenging in terms of low brightness, illumination changes due to artificial lighting, and low visibility.

B. Results

Due to the absence of GPS or motion capture systems underwater, COLMAP [13] is used to produce baseline trajectories. COLMAP is a Structure-from-Motion (SfM) framework that performs joint optimization of camera poses and map structure using global bundle adjustment. Since one of the datasets only contains monocular images, we consider the COLMAP trajectory up to scale for uniformity. Thus, the trajectories are aligned with the COLMAP baseline trajectories using *sim3* alignment [37].

We compare the performance of our magnetometer formulation with the VIO formulation from OKVIS [2] in terms of absolute trajectory error (ATE) [38]. We include the results of the VIO-only case without fusing magnetometer measurements and compare them with the proposed method after *sim(3)* alignment. Each method is run three times and the absolute trajectory error in terms of degree/meters is reported in Table I. As seen in the table, there is a significant reduction in both translation and rotation RMSE. In particular in the cave1_long dataset, both the rotation and translation error decreased significantly from ∼15°to ∼6°and from ∼14m to ∼3m.

TABLE I: Three times run mean absolute trajectory error (ATE) for VIO with and without magnetometer compared with COLMAP baseline after *sim(3)* alignment in terms of degree/meters.

dataset	length	VIO	VIO+MAG
cave1_short	220m	6.31°/6.45m	4.48°/4.98m
cave1_long	506m	14.65°/14.57m	5.83°/3.37m

We also computed the relative trajectory error (RPE) that computes the relative error between states at different times with the focus on the yaw error [39]. RPE shows the local drift over different sections of the trajectory and is unaffected by previously accumulated error. The relative yaw error calculated for cave1_long trajectory after aligning the initial pose is shown in Fig. 6. We can see that when using the magnetometer the yaw error remains constant over large distances, whereas the VIO accumulates error when distances keep increasing.

Two different deployments from the Devil's Eye cave system are presented in Fig. 5. The first deployment, a short foray of 100 meters penetration inside the cave, can be seen in the left image, and the drift in the trajectory (blue) without magnetometer data fusion is clear. The visual inertial magnetometer fusion (green) follows much more closely the ground truth trajectory from COLMAP. More

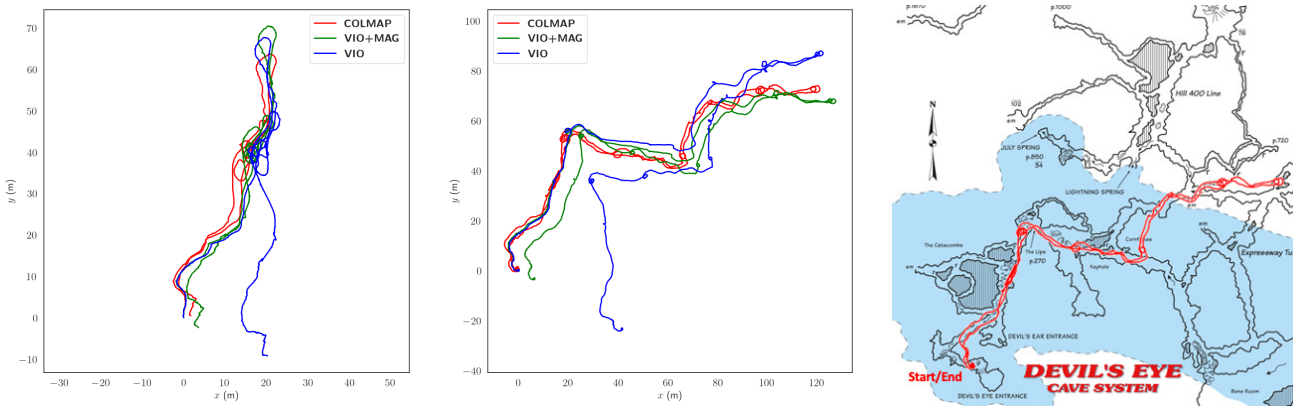

Fig. 5: Trajectories from two deployments at the Devil's Eye cave system, FL, USA. The first two images show the trajectories, COLMAP (red) treated as ground truth, OKVIS VIO [2] (blue), and OKVIS with magnetometer (green) the proposed method. The first trajectory was carried out with the old sensor suite [1] and was a short (∼220m) trajectory, the second was done with the new system and went further (∼500m) trajectory. The final figure presents a section of the man-made map with the ground truth trajectory superimposed (the blue background indicates the Santa Fe River over the cave). These trajectories are aligned at the origin (0,0) only to show endpoint error.

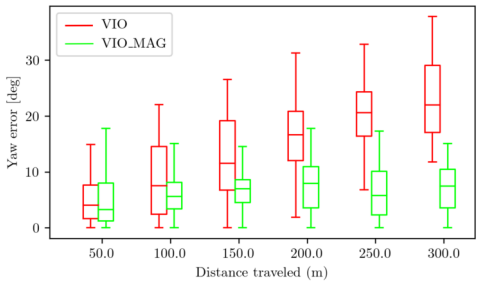

Fig. 6: Relative trajectory error (yaw) comparison between VIO and VIO+magnetometer.

dramatic improvements are presented in the middle figure with trajectory of more than 500 meters. Compared to the globally optimized (COLMAP) trajectory there was reduction of positional error from 2.8% to 0.6%. The right image presents a partial map of the cave system with the COLMAP trajectory superimposed. The map is not up to scale, thus some passages are not scaled uniformly. Nonetheless, the trajectory follows the main passages of the cave system.

Figure 7 presents the trajectory resulting from the deployment of the old sensor suite at the Cenote Dos Ojos. During the 50-minute deployment, the sensor covered approximately 680 meters inside a highly decorated cave. The large volume of data made the production of a COLMAP-based ground truth trajectory infeasible due to computation and memory constraints. As such, this trajectory is used only as a qualitative comparison. As can be seen, the VIO diverged by almost 100 meters. The trajectory obtained from magnetometer fusion follows the path out much more closely to the path taken on the way in. It is worth noting that the sensor's trajectory does not follow exactly the same path; however, the confined structure of the cave passage constrains the inbound and the outbound trajectories fairly close.

V. Conclusion

In this paper we presented a novel approach for augmenting VIO with magnetic field data. The experimental results

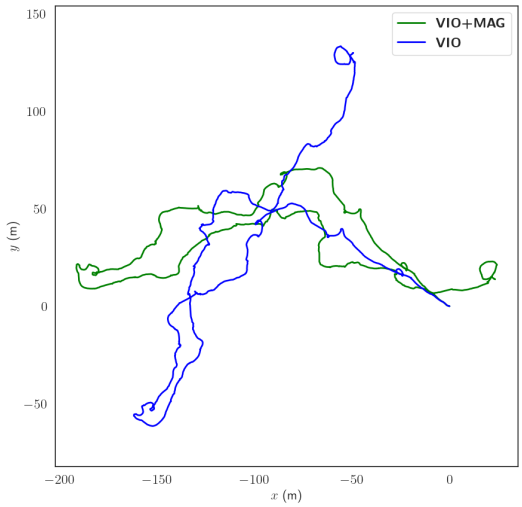

Fig. 7: Trajectories (∼680m) from a deployment at Cenote Dos Ojos cave system, QR, Mexico. OKVIS VIO [2] (blue), and OKVIS with magnetometer (green) the proposed method. These trajectories are aligned at the origin (0,0) only to show endpoint error.

demonstrated significant improvements in the accuracy of the resulting trajectories, significant reduction in the orientation error, and also the ability to orient the resulting trajectory with the magnetic north. During the estimation process, the magnetometer and the accelerometers provide absolute attitude information (with respect the magnetic north and the gravity vector).

Future work will explore the performance of the proposed system during shipwreck mapping, where the sensor traverses near metal structures. During initial calibration we expect to recover the true yaw of the sensor with respect to the magnetic north, enabling better positioning of the underwater structure in space. Furthermore, deployments of the proposed system on an autonomous underwater vehicle will require placing the magnetometer at some distance from the motors. A future experimental study will evaluate the magnetic noise levels for popular vehicles, such as the BlueRoV2 [40] or the Aqua2 AUV [41].

References

[1] S. Rahman, N. Karapetyan, A. Quattrini Li, and I. Rekleitis, "A modular sensor suite for underwater reconstruction," in *MTS/IEEE OCEANS - Charleston*, IEEE, 2018, pp. 1–6.

[2] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual–inertial odometry using nonlinear optimization," *The International Journal of Robotics Research*, vol. 34, no. 3, pp. 314–334, 2015.

[3] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.

[4] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 1–21, 2016.

[5] "Climate Change and Sea-Level Rise in Florida," Florida Ocean and Coastal Council, Tallahasee, FL, Tech. Rep., 2010.

[6] Z. Xu, S. W. Bassett, B. Hu, and S. B. Dyer, "Long distance seawater intrusion through a karst conduit network in the Woodville Karst Plain, Florida," *Sc. Rep.*, vol. 6, 2016.

[7] A. Abbott, "Mexican skeleton gives clue to American ancestry," *Nature News*, May 2014, Springer Nature.

[8] B. L. MacDonald, J. C. Chatters, E. G. Reinhardt, F. Devos, S. Meacham, D. Rissolo, B. Rock, C. Le Maillot, D. Stalla, M. D. Marino, *et al.*, "Paleoindian ochre mines in the submerged caves of the yucatán peninsula, quintana roo, mexico," *Science advances*, vol. 6, no. 27, eaba1219, 2020.

[9] D. Rissolo, A. N. Blank, V. Petrovic, R. C. Arce, C. Jaskolski, P. L. Erreguerena, and J. C. Chatters, "Novel application of 3d documentation techniques at a submerged late pleistocene cave site in quintana roo, mexico," in *2015 Digital Heritage*, IEEE, vol. 1, 2015, pp. 181–182.

[10] S. de Azevedo, M. C. Bortolini, S. L. Bonatto, T. Hünemeier, F. R. Santos, and R. González-José, "Ancient remains and the first peopling of the americas: Reassessing the hoyo negro skull," *American Journal of Physical Anthropology*, vol. 158, no. 3, pp. 514–521, 2015.

[11] N. Kresic and A. Mikszewski, *Hydrogeological Conceptual Site Models: Data Analysis and Visualization*. Boca Raton, LA: CRC Press, 2013.

[12] B. Joshi, M. Xanthidis, M. Roznere, N. J. Burgdorfer, P. Mordohai, A. Q. Li, and I. Rekleitis, "Underwater exploration and mapping," in *IEEE OES AUV Symposium*, Singapore, Sep. 2022, pp. 1–7.

[13] J. L. Schönberger and J.-M. Frahm, "Structure-from-motion revisited," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[14] A. Quattrini Li, A. Coskun, S. M. Doherty, *et al.*, "Experimental comparison of open source vision based state estimation algorithms," in *International Symposium of Experimental Robotics (ISER)*, Tokyo, Japan, Mar. 2016.

[15] B. Joshi, S. Rahman, M. Kalaitzakis, *et al.*, "Experimental Comparison of Open Source Visual-Inertial-Based State Estimation Algorithms in the Underwater Domain," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, Nov. 2019, pp. 7221–7227.

[16] S. Rahman, A. Quattrini Li, and I. Rekleitis, "Sonar Visual Inertial SLAM of Underwater Structures," in *IEEE International Conference on Robotics and Automation*, Brisbane, Australia, May 2018, pp. 5190–5196.

[17] S. Rahman, A. Quattrini Li, and I. Rekleitis, "An Underwater SLAM System using Sonar, Visual, Inertial, and Depth Sensor," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, (IROS ICROS Best Application Paper Award. Finalist), Nov. 2019, pp. 1861–1868.

[18] S. Rahman, A. Quattrini Li, and I. Rekleitis, "SVIn2: A Multi-sensor Fusion-based Underwater SLAM System," *International Journal of Robotics Research*, vol. 41, no. 11-12, pp. 1022–1042, Jul. 2022.

[19] B. Joshi, M. Xanthidis, S. Rahman, and I. Rekleitis, "High definition, inexpensive, underwater mapping," in *IEEE International Conference on Robotics and Automation (ICRA)*, Philadelphia, PA, USA, 2022, pp. 1113–1121.

[20] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM3: An accurate open-source library for visual, visual–inertial, and multimap slam," *IEEE Transactions on Robotics*, 2021.

[21] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, and G. Huang, "Openvins: A research platform for visual-inertial estimation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2020, pp. 4666–4672.

[22] S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan, "Estimation of imu and marg orientation using a gradient descent algorithm," in *2011 IEEE International Conference on Rehabilitation Robotics*, 2011, pp. 1–7.

[23] R. G. Valenti, I. Dryanovski, and J. Xiao, "Keeping a good attitude: A quaternion-based orientation filter for imus and margs," *Sensors*, vol. 15, no. 8, pp. 19 302–19 330, 2015.

[24] H. Zhao and Z. Wang, "Motion measurement using inertial sensors, ultrasonic sensors, and magnetometers with extended kalman filter for data fusion," *IEEE Sensors Journal*, vol. 12, no. 5, pp. 943–953, 2012.

[25] J. Wang, L. Li, H. Yu, X. Gui, and Z. Li, "Vimo: A visual-inertial-magnetic navigation system based on

non-linear optimization," *Sensors*, vol. 20, no. 16, 2020.

[26] B. Siebler, S. Sand, and U. D. Hanebeck, "Localization with magnetic field distortions and simultaneous magnetometer calibration," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3388–3397, 2021.

[27] D. Caruso, A. Eudes, M. Sanfourche, D. Vissière, and G. le Besnerais, "Robust indoor/outdoor navigation through magneto-visual-inertial optimization-based estimation," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 4402–4409.

[28] D. Caruso, A. Eudes, M. Sanfourche, D. Vissiere, and G. le Besnerais, "An inverse square root filter for robust indoor/outdoor magneto-visual-inertial odometry," in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2017, pp. 1–8.

[29] D. Gebre-Egziabher, G. H. Elkaim, J. D. Powell, and B. W. Parkinson, "Calibration of strapdown magnetometers in magnetic field domain," *Journal of Aerospace Engineering*, vol. 19, no. 2, pp. 87–102, 2006.

[30] J. F. Vasconcelos, G. Elkaim, C. Silvestre, P. Oliveira, and B. Cardeira, "Geometric approach to strapdown magnetometer calibration in sensor frame," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1293–1306, 2011.

[31] M. Kok and T. B. Schön, "Magnetometer calibration using inertial sensors," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5679–5689, 2016.

[32] A. Solin, M. Kok, N. Wahlström, T. B. Schön, and S. Särkkä, "Modeling and interpolation of the ambient magnetic field by gaussian processes," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1112–1127, 2018.

[33] J. Solà, *Quaternion kinematics for the error-state kalman filter*, 2017.

[34] P. Furgale, J. Rehder, and R. Siegwart, "Unified temporal and spatial calibration for multi-sensor systems," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 1280–1286.

[35] "IEEE standard for specifying and testing single-axis interferometric fiber optic gyros," *IEEE Std 952-2020 (Revision of IEEE Std 952-1997)*, pp. 1–93, 2021.

[36] Y. Wu and W. Shi, "On calibration of three-axis magnetometer," *IEEE Sensors Journal*, vol. 15, no. 11, pp. 6424–6431, 2015.

[37] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376–380, 1991.

[38] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of rgb-d slam systems," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 573–580.

[39] Z. Zhang and D. Scaramuzza, "A tutorial on quantitative trajectory evaluation for visual(-inertial) odometry," in *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2018.

[40] B. Robotics, *BlueROV2*, `https://bluerobotics.com/store/rov/bluerov2/`, 2023.

[41] G. Dudek, M. Jenkin, C. Prahacs, *et al.*, "A visually guided swimming robot," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Edmonton AB, Canada, Aug. 2005, pp. 1749–1754.