

Orientation Tracking

Yao-Ting, Huang

Dept. Electrical and Computer Engineering
University of California, San Diego
La Jolla, United States
yah032@ucsd.edu

I. INTRODUCTION

In this project, we aim to solve the problem of tracking the 3D orientation of a rotating body using data from an Inertial Measurement Unit (IMU) and reconstructing a panoramic image from camera data. The orientation of a rotating body refers to the precise angular position and orientation of the body in 3D space, which is crucial in numerous robotic and autonomous systems. By leveraging IMU data, which includes measurements of angular velocity and linear acceleration, we can estimate the body's orientation over time. This orientation can then be used to align and stitch images captured by a camera mounted on the rotating body to create a panoramic view of the surrounding environment.

Orientation tracking and panoramic image stitching play a vital role in many applications, particularly in robotics and autonomous systems. For example, in robotics, accurate orientation estimation is necessary for tasks such as navigation, manipulation, and stabilization. In autonomous vehicles and drones, orientation tracking helps in precise motion planning and sensor fusion, while panoramic image stitching provides an extended view of the environment, improving situational awareness and obstacle detection. These tasks are essential for safe and efficient operation in dynamic, real-world environments.

To tackle this problem, we employ a two-step approach. First, we use a projected gradient descent algorithm to track the orientation of the body by estimating its quaternion trajectory based on IMU measurements. The quaternion representation allows for efficient and stable orientation estimation over time. Second, we utilize the estimated orientation trajectory to align and stitch camera images, ultimately constructing a panoramic image that reflects the body's motion and the surrounding environment. By combining these techniques, we can achieve robust orientation tracking and generate accurate panoramic visualizations, even in the presence of noise and sensor inaccuracies.

II. PROBLEM FORMULATION

The goal of this project is to estimate the 3D orientation of a rotating body using measurements from an Inertial Measurement Unit (IMU), and subsequently use these orientation estimates to reconstruct a panoramic image by stitching camera images captured during the rotation. This task involves solving two primary subproblems: orientation tracking and image stitching.

A. Orientation Tracking

We aim to track the orientation of the body over time using the IMU's angular velocity ($\omega_t \in \mathbb{R}^3$) and linear acceleration ($a_t \in \mathbb{R}^3$) measurements. The orientation at each time step t is represented by a unit quaternion $q_t \in \mathbb{H}_*$, where \mathbb{H}_* denotes the space of unit quaternions. We use the quaternion kinematics motion model to predict the body's orientation at the next time step:

$$q_{t+1} = f(q_t, \tau_t \omega_t) := q_t \circ \exp\left(\left[0, \frac{\tau_t \omega_t}{2}\right]\right) \quad (1)$$

Here, $f(q_t, \tau_t \omega_t)$ describes the quaternion update rule, where τ_t is the time step between consecutive measurements. The exponential map $\exp(\cdot)$ is used to update the quaternion based on the angular velocity.

The IMU measurements also provide the acceleration of the body, which in the world frame is expected to be $[0, 0, -g]^T$, where g is the acceleration due to gravity. The measured acceleration vector a_t in the IMU frame should correspond to the transformed gravity vector in the IMU frame, leading to the following observation model:

$$[0, a_t] = h(q_t) := q_t^{-1} \circ [0, 0, 0, -g] \circ q_t \quad (2)$$

To estimate the orientation trajectory $q_{1:T} = \{q_1, q_2, \dots, q_T\}$ over time, we formulate an optimization problem with the objective of minimizing a cost function that accounts for both the motion model prediction and the observation model alignment. The cost function $c(q_{1:T})$ consists of two terms: the first term penalizes the difference between the predicted orientation from the motion model and the estimated orientation, and the second term penalizes the discrepancy between the observed and predicted accelerations:

$$c(q_{1:T}) := \frac{1}{2} \sum_{t=0}^{T-1} \|2 \log (q_{t+1}^{-1} \circ f(q_t, \tau_t \omega_t))\|_2^2 \quad (3)$$

$$+ \frac{1}{2} \sum_{t=1}^T \| [0, a_t] - h(q_t) \|_2^2 \quad (4)$$

Also we need to ensure that the quaternion trajectory $q_{1:T}$ remains consistent with both the angular velocity measurements and the gravity measurements, subject to the constraint that all quaternions q_t remain unit-norm:

$$\|q_t\|_2 = 1, \quad \forall t \in \{1, 2, \dots, T\} \quad (5)$$

Hence, after gradient-descent step, we need to project the trajectories onto the space of unit-quaternions \mathbb{H}_* at all times, ensuring valid quaternion representations. The projection step in the gradient descent ensures that the quaternions are always normalized.

$$\Pi_{\mathbb{H}_*}(q) = \frac{q}{\|q\|_2} \quad (6)$$

B. Panorama Reconstruction

Once the orientation trajectory is estimated, we can proceed with the second subproblem of constructing a panoramic image from the camera images captured during the rotation. The camera images, which are paired with timestamps, need to be aligned with the corresponding orientation estimates. The challenge is to align these images with their respective quaternions by associating each camera image with the closest quaternion from the orientation trajectory.

The resulting panorama is obtained by stitching the aligned images together. While perfect stitching is not required, the orientation estimates play a crucial role in ensuring that the images are correctly positioned in relation to each other, thereby producing a coherent panoramic view of the surrounding environment.

This problem formulation combines the tasks of orientation tracking and image stitching to create a robust system for generating panoramic visualizations from rotational motion.

III. TECHNICAL APPROACH

To solve the problem of orientation tracking and panoramic image reconstruction, we employed a two-step approach: first, estimating the orientation trajectory using IMU data, and second, stitching the camera images based on the orientation estimates.

A. IMU Calibration

The first step in the orientation tracking process is the calibration of the IMU. IMUs typically suffer from biases and scale factors that need to be estimated and corrected for accurate orientation tracking. In this project, we utilize the static portions of the IMU data, where the body is not rotating, to calibrate the accelerometer and gyroscope. During these static periods, the accelerometer should measure approximately $[0, 0, -g]^T$ in the world frame of reference, where g is the gravitational acceleration. By comparing the raw accelerometer readings to the expected gravity vector, we can estimate and correct for biases and scale factors.

The gyroscope calibration is performed by using the VICON ground-truth data. The gyroscope's biases can be identified by comparing the IMU's angular velocity readings with the ground-truth angular velocities, ensuring accurate measurements during the integration process.

B. Orientation Tracking

Once the IMU is calibrated, we proceed with the orientation tracking step. The orientation of the body is represented using unit quaternions, which provide a compact and stable representation of rotations. The quaternion motion model, as described in Equation (1), is used to predict the orientation at each time step, given the current orientation q_t and the angular velocity ω_t .

This model is based on quaternion kinematics, where τ_t represents the time step between two consecutive measurements, and \exp is the exponential map for quaternions. This update rule integrates the angular velocity over time to estimate the new orientation.

To estimate the orientation trajectory, we formulate an optimization problem that minimizes the cost function given by Equation (3). The optimization is performed using a projected gradient descent algorithm to iteratively refine the quaternion estimates. The cost function consists of two terms: the first term minimizes the difference between the predicted orientation from the motion model and the actual orientation, while the second term minimizes the discrepancy between the predicted and observed accelerations.

The gradient descent algorithm updates the quaternion trajectory by minimizing the cost function while adhering to the unit-norm constraint, allowing us to obtain a smooth and accurate estimation of the body's orientation.

C. Cylindrical Projection for Panorama Reconstruction

In this step, we perform a more accurate cylindrical projection by utilizing not normalized pinhole rays, which include the focal length, and computing the intersection with a vertical cylinder. The cylinder has a radius equal to the average focal length f_{avg} , and the procedure is as follows:

- 1) **Compute the ray in camera coordinates:** For a pixel at coordinates (u, v) , the ray direction d in camera coordinates is computed as:

$$d = \begin{bmatrix} u - c_x \\ v - c_y \end{bmatrix} \cdot \frac{1}{f} \quad (7)$$

where c_x, c_y are the principal point coordinates, and f is the focal length. Here, we use the average focal length:

$$f_{\text{avg}} = \frac{f_x + f_y}{2} \quad (8)$$

with f_x and f_y being the horizontal and vertical focal lengths, respectively.

- 2) **Rotate the ray into world coordinates:** Using the rotation matrix R obtain from converting the optimized estimates of orientation $q_{1:T}$, which maps camera coordinates to world coordinates, we rotate the ray accordingly.
- 3) **Cylindrical mapping:** On the cylinder, which has a radius $r = f_{\text{avg}}$, the horizontal coordinate θ is determined by the angle:

$$\theta = \text{atan2}(X_{\text{world}}, Z_{\text{world}}) \quad (9)$$

where $(X_{\text{world}}, Z_{\text{world}})$ are the world coordinates of the ray. The vertical coordinate is taken from the world Y -coordinate (the ray's height).

- 4) **Map cylindrical coordinates to pixel coordinates in the panorama:** The cylindrical coordinates are mapped to pixel coordinates in the panorama:

- The horizontal coordinate u_{pano} is computed by linearly mapping θ from the range $[-\pi, \pi]$ to the range $[0, \text{pano_width}]$.
- The vertical coordinate v_{pano} is computed by linearly mapping the world Y -coordinate (after rotation) from the expected range $[0, \text{pano_height}]$.

D. Image Stitching

Once the cylindrical projections of all images have been calculated, we stitch them together by aligning the images based on their corresponding orientation estimates. For each image, we use the closest quaternion estimate q_t based on the image's timestamp. The cylindrical projections are then blended into a panoramic image.

To minimize artifacts in the stitched image, we apply a simple blending technique where overlapping regions between images are averaged.

Although the panorama might not be perfectly aligned due to sensor noise and inaccuracies in the orientation estimates, this approach provides a visually coherent result that captures the surrounding environment from the perspective of the rotating body.

IV. RESULTS

In this section, we present the results of our orientation tracking and panoramic image reconstruction. We evaluate the performance using both training and test datasets, analyzing the estimated orientation and the resulting panoramic images. Several plots are included to demonstrate the convergence of the optimization process, compare the estimated orientation to the ground truth, and show the quality of the constructed panorama.

A. Convergence of the Cost Function

We begin by evaluating the convergence of the cost function during the gradient descent optimization. The following figure shows the cost function value over time as the algorithm progresses, demonstrating how the cost decreases and the optimization converges to a minimum.

B. Comparison of Estimated Orientation with Ground Truth

Next, we compare the estimated roll, pitch, and yaw angles with the ground truth from the VICON system. The following plots show the roll, pitch, and yaw angles obtained from our orientation tracking algorithm, as well as the corresponding ground truth values. This comparison helps us assess the accuracy of the estimated orientation.

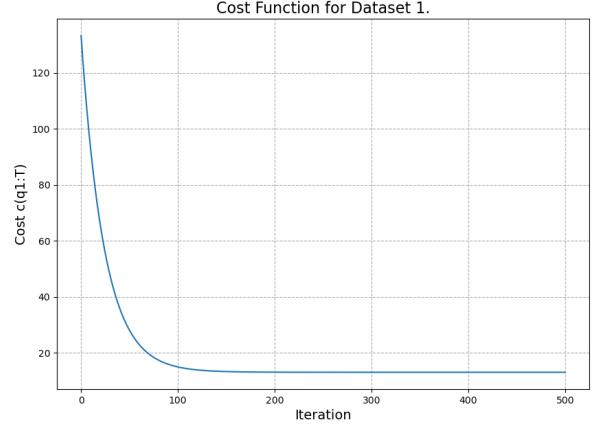


Fig. 1. Cost function convergence during gradient descent optimization. The curve shows the reduction in cost as the algorithm converges to the minimum value.

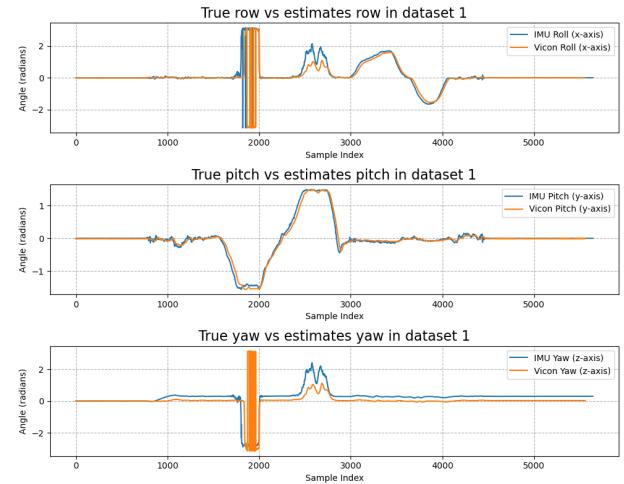


Fig. 2. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and yaw) angle from our method.

C. Panorama Reconstruction

Finally, we evaluate the quality of the panoramic image constructed using the estimated orientations. The following figures show the panoramic image generated by stitching camera images based on the estimated orientations, along with the corresponding ground truth orientation for comparison. The quality of the panorama is visually assessed by comparing the alignment and consistency of the stitched images. As for the image stitching method, we implement both overwriting and overlap-averaging methods and apply to the test cases to evaluate the differences.

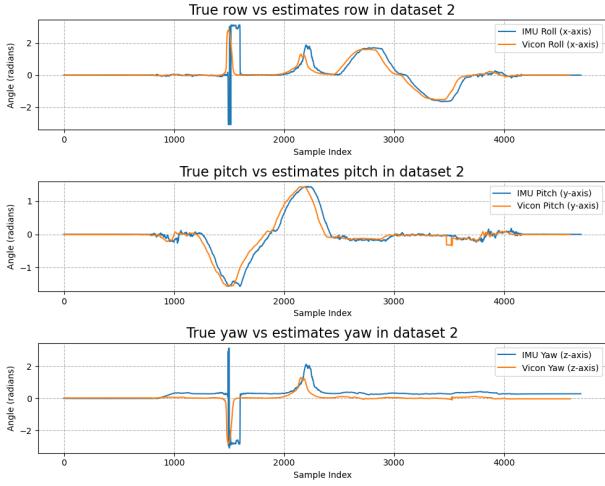


Fig. 3. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and ya) angle from our method.

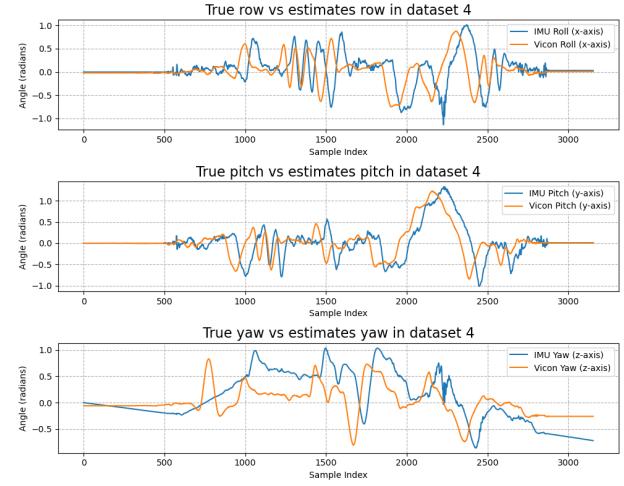


Fig. 5. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and ya) angle from our method.

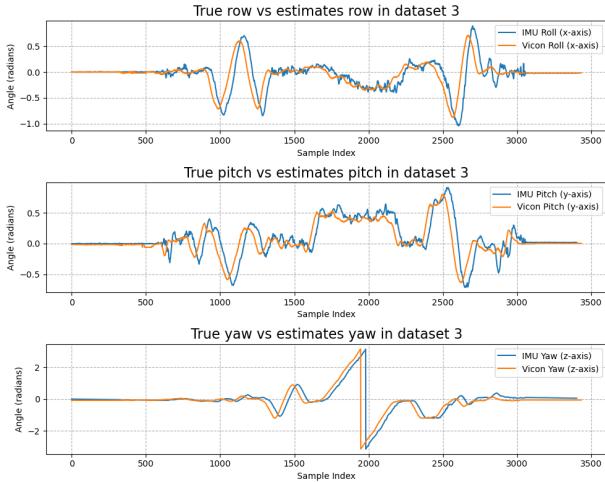


Fig. 4. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and ya) angle from our method.

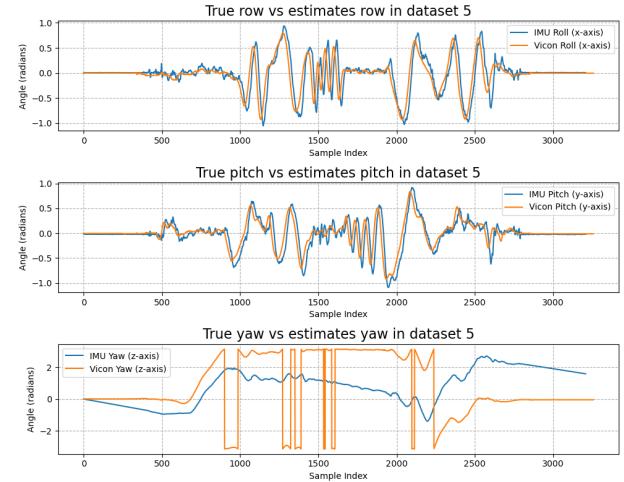


Fig. 6. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and ya) angle from our method.

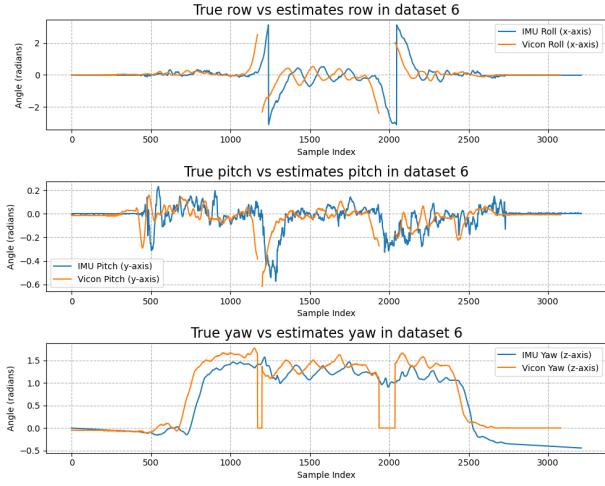


Fig. 7. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and ya) angle from our method.

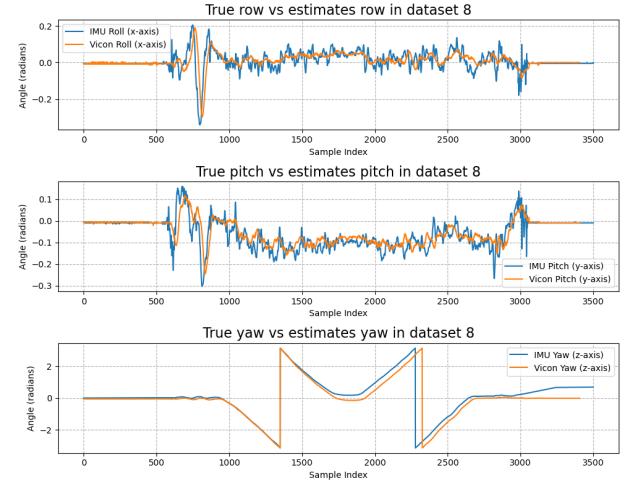


Fig. 9. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and ya) angle from our method.

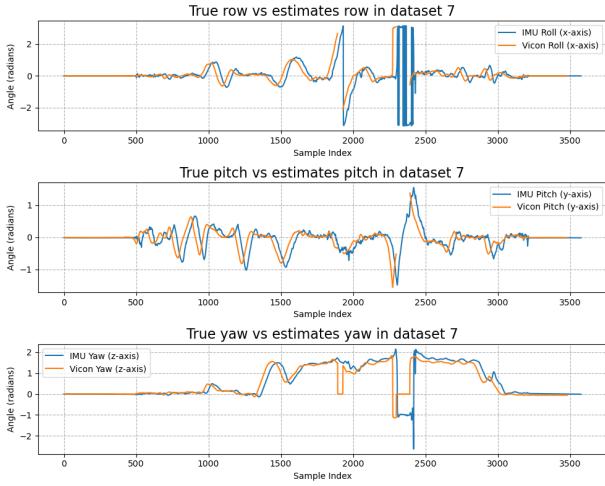


Fig. 8. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and ya) angle from our method.

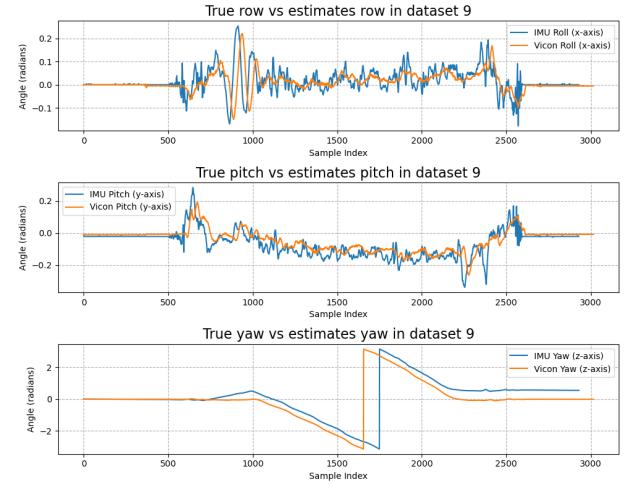


Fig. 10. Comparison of estimated angles with ground truth. The orange curve represents the ground truth, and the blue curve shows the estimated (roll, pitch, and ya) angle from our method.



Fig. 11. Panoramic image reconstructed using the estimated orientation. The image shows the stitched camera images, aligned using the estimated orientations over time.

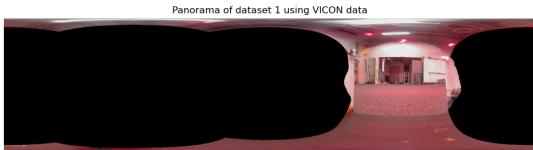


Fig. 12. Panoramic image reconstructed using the ground truth orientation. The image shows the stitched camera images, aligned using the ground truth orientations for comparison.



Fig. 13. Panoramic image reconstructed using the estimated orientation. The image shows the stitched camera images, aligned using the estimated orientations over time.

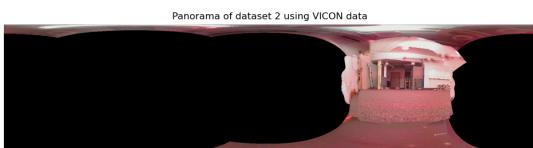


Fig. 14. Panoramic image reconstructed using the ground truth orientation. The image shows the stitched camera images, aligned using the ground truth orientations for comparison.

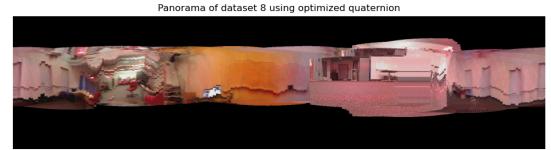


Fig. 15. Panoramic image reconstructed using the estimated orientation. The image shows the stitched camera images, aligned using the estimated orientations over time.



Fig. 16. Panoramic image reconstructed using the ground truth orientation. The image shows the stitched camera images, aligned using the ground truth orientations for comparison.

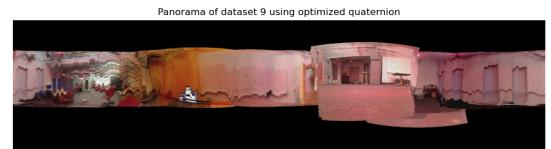


Fig. 17. Panoramic image reconstructed using the estimated orientation. The image shows the stitched camera images, aligned using the estimated orientations over time.



Fig. 18. Panoramic image reconstructed using the ground truth orientation. The image shows the stitched camera images, aligned using the ground truth orientations for comparison.



Fig. 19. Panoramic image reconstructed using the estimated orientation. The image shows the stitched camera images using overwriting methods, aligned using the estimated orientations over time.

Panorama of dataset 10 using optimized quaternion.



Fig. 20. Panoramic image reconstructed using the estimated orientation. The image shows the stitched camera images using overlap-averaging stitching method, aligned using the estimated orientations over time.

Panorama of dataset 11 using optimized quaternion.



Fig. 21. Panoramic image reconstructed using the estimated orientation. The image shows the stitched camera images using overwriting methods, aligned using the estimated orientations over time.

Panorama of dataset 11 using optimized quaternion.



Fig. 22. Panoramic image reconstructed using the estimated orientation. The image shows the stitched camera images using overlap-averaging stitching method, aligned using the estimated orientations over time.

D. Discussion

The results demonstrate that our method produces reasonably accurate estimates of the orientation, as seen in the comparison between the estimated roll, pitch, and yaw angles and the ground truth. The panoramic images generated from the estimated orientations align well with the ground truth, although minor misalignments can still be observed due to the noise in the IMU data and imperfections in the optimization process. Comparing both stitching methods, we found that the overlap-averaging method provides more smooth panoramic images and is less affected by the error of orientation estimates. The convergence curve of the cost function shows that the gradient descent algorithm successfully minimizes the error and converges to an optimal solution.

Overall, our approach provides a solid foundation for orientation tracking and panoramic image reconstruction, and the results suggest that further refinements can improve accuracy and robustness in real-world applications.