Low-Cost Webcam-Based Motion Capture System

Oskar Hessler & Ebbe Persson
Lund University & Robert Bosch AB

1 INTRODUCTION

IMU sensors (Inertial Measurement Units) from Bosch Sensortec combine accelerometers and gyroscopes to measure position and orientation. They are widely used for navigation and motion tracking in consumer electronics.

However, small measurement errors accumulate over time, causing sensor drift. Evaluating and correcting this drift usually requires expensive motion capture systems, creating a barrier for efficient development and testing.

This project aimed to create an affordable motion capture system using standard webcams for IMU development at Robert Bosch AB.

2 PROBLEM

"Can we build an accurate, low-cost motion capture system using common webcams to track three-dimensional movement and rotation?"

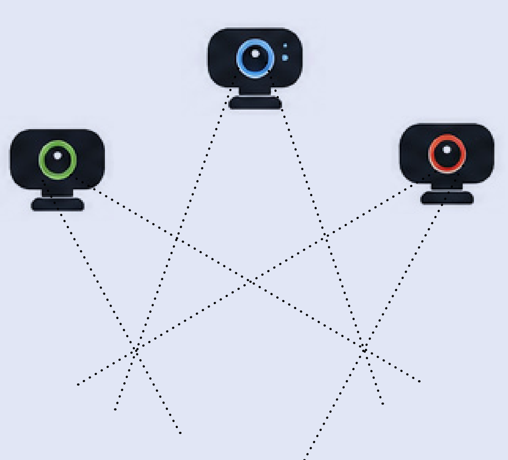
3 METHODOLOGY

The system was developed as a five-step pipeline where each stage builds on the previous one. The cameras are first physically arranged and calibrated to establish a shared coordinate system, before their video feeds are synchronized in time. Markers are then detected in each frame and reconstructed into three-dimensional motion data.

Setup → **Calibration** → **Synchronisation** → **Detection** → **3D-Reconstruction**

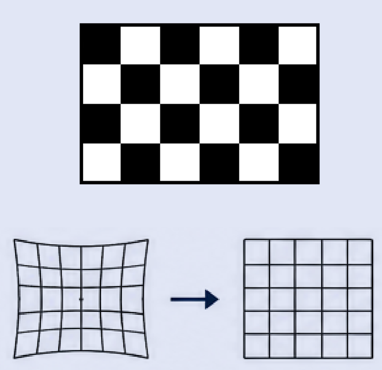
Setup

The system uses three webcams connected to a computer in a triangular arrangement. Capturing motion from multiple viewpoints improves depth perception and tracking reliability by keeping markers visible from several cameras during movement.



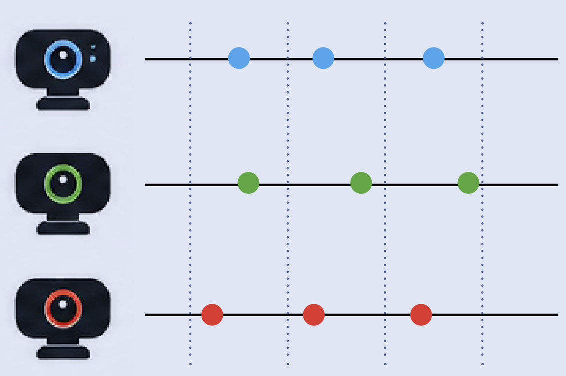
Calibration

A two-stage calibration process first corrects lens distortion using a chequerboard pattern. It then establishes a shared 3D coordinate system so all camera positions and orientations are known relative to one another.



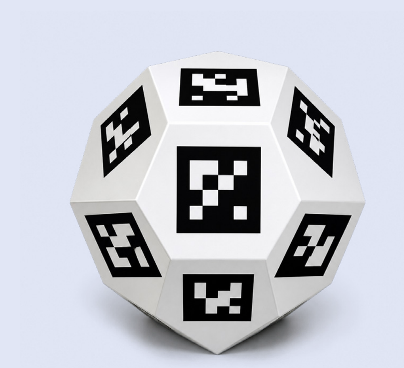
Synchronisation

To align all camera feeds in time, a software-based synchronization method matches frames using timestamps. This ensures all reconstructed motion represents the same moment in time.



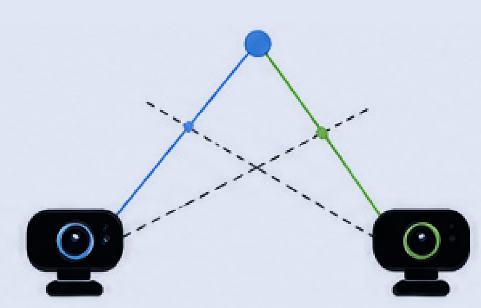
Detection

The system uses OpenCV to detect ArUco markers (less information-dense QR-codes) in each camera feed. By identifying each marker's unique pattern, the software estimates its position and orientation in the image. These markers are attached to the tracked object.



3D-Reconstruction

Data from at least two cameras is combined through triangulation to recover the markers' 3D positions, from which the object's trajectory and rotation are reconstructed.



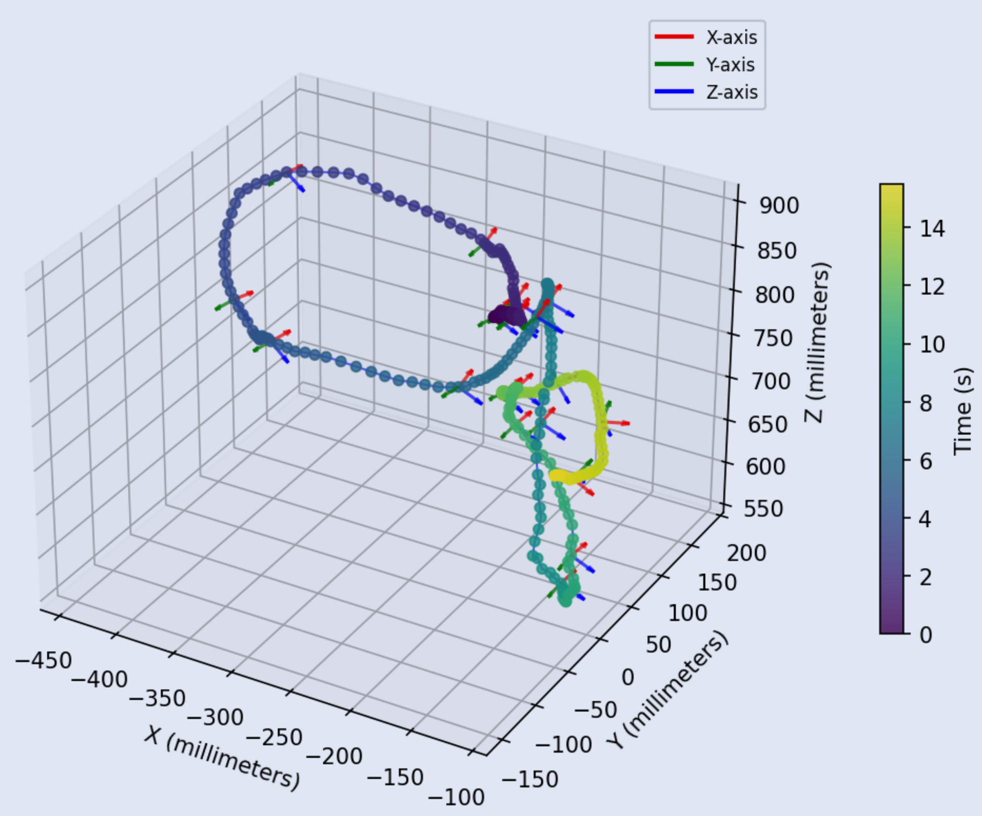
4 SOLUTION

The final system combines calibration, synchronization, marker detection, and triangulation into a single motion capture pipeline. A two-stage calibration process creates accurate camera models, while timestamp-based synchronization ensures that frames from all cameras represent the same instant in time. OpenCV-based ArUco detection identifies the markers in each image, and triangulation reconstructs their 3D positions. By using a rigid body template, the system can also estimate the object's pose even when some markers are temporarily hidden.

5 RESULTS

The system successfully tracked motion over more than 1600 frames, achieving a mean trajectory coverage of 92.9% across all markers. Most coverage losses were caused by temporary marker occlusion rather than software failure.

Synchronization accuracy averaged approximately 13.3 ms between cameras, remaining within the 20 ms target without observable drift. Reconstruction accuracy reached a re-projection error of 2-4 pixels and an extrinsic baseline accuracy of ±30 mm. Consistent marker step distances confirmed the stability of the rigid body measurements.



6 DISCUSSION

The results showed that system precision depended mainly on camera calibration quality rather than the reconstruction algorithms themselves. Markers near image edges produced more positional jitter due to residual lens distortion, which is common in camera-based tracking systems.

Synchronization performance was mainly limited by USB hardware and operating system scheduling, since real-time and post-processing approaches produced similar results. Coverage losses were primarily caused by the requirement that markers remain visible from at least two cameras. Improved calibration methods such as ChArUco boards could therefore significantly increase accuracy.

7 CONCLUSION

This project demonstrates that an accurate motion capture system can be built using affordable consumer webcams. The developed pipeline met the main performance targets and provides a low-cost alternative to commercial systems for evaluating IMU sensor drift.

The rigid body template also enables tracking of arbitrary three-dimensional objects. Future improvements, including ChArUco calibration, Kalman filtering, and higher-resolution cameras, could further improve precision and robustness.

8 ACKNOWLEDGEMENTS

We would like to thank Robert Bosch AB in Lund for supporting this bachelor's thesis, with special thank to Stefan Noll and Sebastian Farås for their guidance and feedback. We also thank our academic supervisor Christian Nyberg at Lund University for his supervision throughout the project.