HARDWARE-DEPENDENCY-REDUCED VISUAL SLAM

Lian YU, Dongsheng LI, GuoYan WANG*, Fei ZHAO, Hongqi FAN

National Key Laboratory of Automatic Target Recognition, College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

Annotation: To address the poor real-time performance and hardware dependency of traditional visual Simultaneous Localization and Mapping on embedded platforms, this paper proposes a cloud-native collaborative framework enabling low-latency data transmission and cloud-based pose estimation. Experimental results demonstrate stable frame rates without map fragmentation or tracking failures in scenario testing. The architecture enhances environmental adaptability and computational robustness for visual Simultaneous Localization and Mapping on compact unmanned platforms, providing a high-precision solution for resource-constrained scenarios.

Keywords: Cloud-Native Architecture; Visual Simultaneous Localization and Mapping; Unmanned Systems; Real-Time Streaming Protocol;

1. Introduction

In modern unmanned combat systems, autonomous environmental perception and real-time localization capabilities constitute the critical operational advantages of unmanned platforms. As the core enabling technology for environmental awareness and navigation positioning, Simultaneous Localization and Mapping (SLAM) establishes spatial cognition for unmanned systems by simultaneously computing platform motion trajectories and constructing environmental topological models in real-time. Notably, visual SLAM has emerged as the preferred solution for autonomous localization in structured environments at low velocities, owing to its low sensor cost and strong environmental adaptability.

However, visual SLAM faces dual performance constraints: (1) algorithmic complexity arising from feature extraction (e.g., ORB feature detection) and nonlinear optimization processes (e.g., pose graph optimization), and (2) real-time performance degradation caused by resource contention between pose estimation and mapping threads. For small-scale unmanned systems with stringent size-, weight-, and power-constrained (SWaP-constrained) configurations, these limitations result in positioning accuracy and computational efficiency that fail to meet real-time requirements for battlefield reconnaissance tasks.

To address these technical bottlenecks, this paper proposes a cloud-native architecture-based collaborative computing framework for visual SLAM, effectively mitigating onboard hardware constraints.

2 Background and Related Work

2.1 Visual SLAM

Simultaneous Localization and Mapping (SLAM) is a foundational technology that enables autonomous device positioning and concurrent environmental mapping in unknown environments through multi-sensor data fusion. Its technical evolution can be delineated through the following developmental stages:

Early research predominantly utilized filtering frameworks, where state estimation and map