



# **Development of a Novel Visual-Inertial SLAM System**

**A THESIS**

**Submitted to the Graduate School of  
Innovative Design Engineering,  
Egypt-Japan University of Science and Technology (E-JUST)**

**In Partial Fulfillment of the Requirements for the Degree  
of  
Doctor of Philosophy  
in  
Mechatronics and Robotics Engineering**

**by**

**Mohamed Hamdy Mohamed Abdallah Merzban**

**February 2015**

## Summary

Simultaneous Localization And Mapping (SLAM) is a system that comprises the estimation of both robot pose and map structure when moving in an unknown and unstructured environment. SLAM system may utilize sensors such as inertial sensors, laser range finders, GPS, and cameras. Laser range finders had long been used to sense the surrounding environment. Recently, there had been much interest in visual SLAM using camera only to reduce system cost. Estimation of robot pose using visual SLAM is not easy due to the absence of depth information in camera images and the strong nonlinearity of the perspective projection function. This also imposes computational challenges on visual SLAM techniques, especially when the robot moves in a large map. As the map size increases, the computational cost for updating the map and consulting it for localizing a robot becomes expensive.

In this work, a hybrid visual/inertial SLAM system is developed where the measurements from a single camera and an Inertial Measurement Unit (IMU) are fused together.

The main objective of the new approach is to reduce the computational load in large environments which can be described as the system scalability. The new system comprises two stages: at the first stage, incremental position changes, orientation, velocity, and 3-D relative positions of landmarks are estimated from the sensor measurements. A novel visual-inertial fusion technique for inertial and visual data is devised utilizing the inherent sparsity in the optimization matrix. While, in the second stage, a map of the environment is maintained using the results of the first stage. A novel map representation as a graph of interconnected landmark points is proposed. The landmark positions are extracted through an exactly linear optimization process of the map. Since map update is the most expensive process, this linearity is a significant improvement compared to the commonly used nonlinear

optimization techniques. The computational complexity of the second stage is of the order  $O(N^2)$ .

The relative motion estimation module is implemented and tested against standard benchmark data sets, a custom dataset is also used to investigate the performance and accuracy of the proposed system. The results indicate that the proposed system is stable and accurate compared to the state of the art methods. Simulations of the whole system are performed to verify the feasibility and accuracy of it.