

Real-time Camera Tracking for Gameplay Automation in Video Games

A navigational assistance module for video games that works in real-time using only a single camera's footage

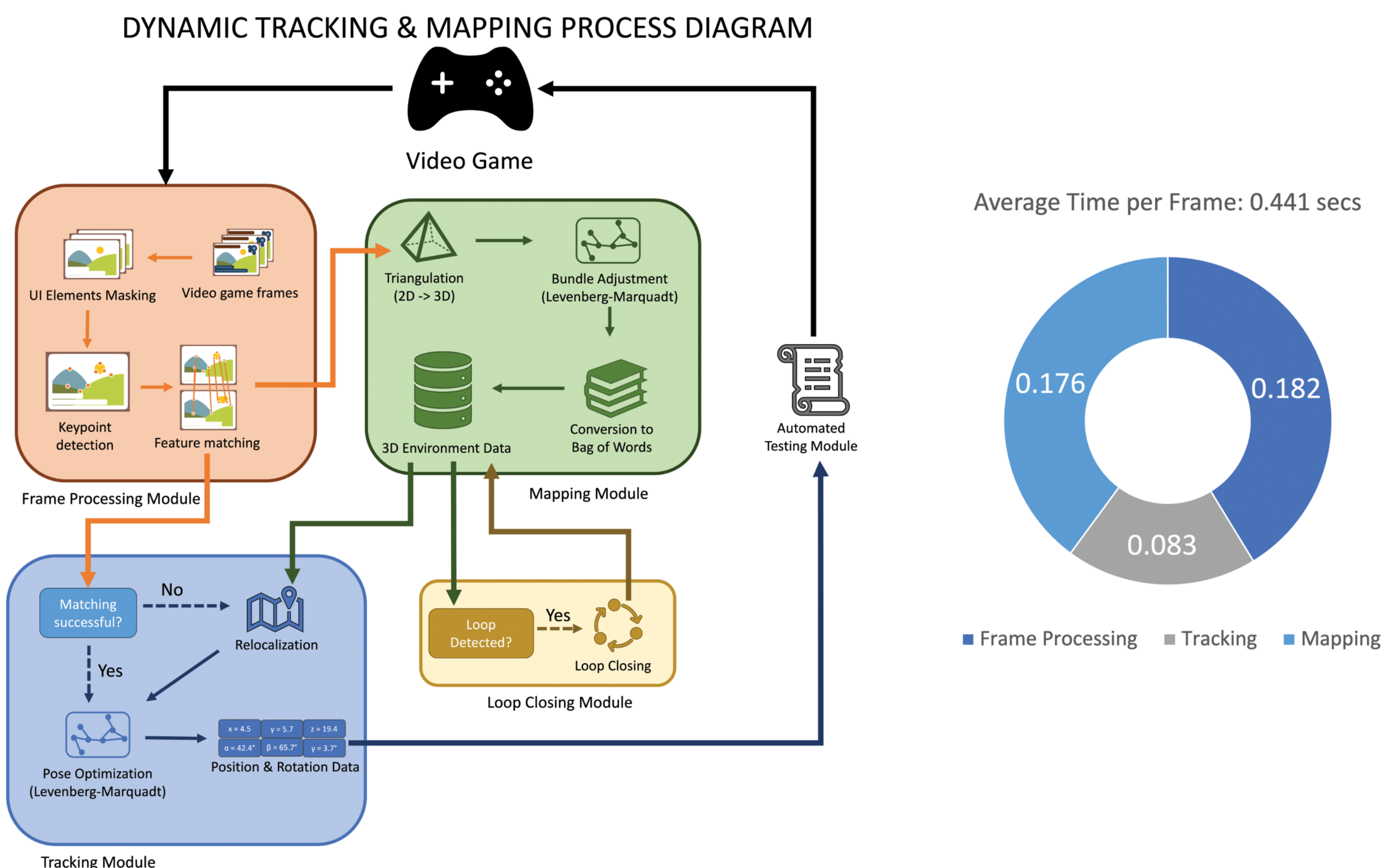
Mahesh Jayasankar

Igor Gilitschenski

ACADEMIC SUPERVISOR

Yonas Bedasso

INDUSTRY SUPERVISOR



PROJECT SUMMARY

The developed solution is based on monocular simultaneous localization and mapping (SLAM) techniques such as ORB-SLAM [1]. The video game footage is first extracted frame by frame and pre-processed to remove UI elements using a convolutional neural network. Key image features, such as object corners, are detected using ORB feature detection [2]. The image features are compared with past frames, and the relative movement between frames (called the pose) is estimated using the Levenberg-Marquadt optimization algorithm. During this process, the solution also keeps track of the key 3D environment information to improve long-term consistency of the tracked movements. When revisiting a location, the solution efficiently detects similarity with past frames using DBoW [3] and uses this information to eliminate drift, which is caused due to accumulation of error across frames. The solution is integrated within an automated gameplay testing tool, operating in real-time alongside the video game itself.

REFERENCES

1. Mur-Artal, R., Montiel, J.M.M. and Tardos, J.D., 2015. ORB-SLAM: a versatile and accurate monocular SLAM system. IEEE transactions on robotics, 31(5), pp.1147-1163.
2. Rublee, Ethan & Rabaud, Vincent & Konolige, Kurt & Bradski, Gary. (2011). ORB: an efficient alternative to SIFT or SURF. Proceedings of the IEEE International Conference on Computer Vision. 2564-2571. 10.1109/ICCV.2011.6126544.
3. D. Galvez-López and J. D. Tardos, "Bags of Binary Words for Fast Place Recognition in Image Sequences," in IEEE Transactions on Robotics, vol. 28, no. 5, pp. 1188-1197, Oct. 2012, doi: 10.1109/TRO.2012.2197158.

