Visual SLAM for humanoids on embedded devices

Nikolaos Tavoularis[1#,], **Stylianos Piperakis**[2*] , **Emmanouil Hourdakis**[*] and , **Panos Trahanias**[2]

[1] Computer Science Department, University of Crete, Heraklion, Greece

[2] Institute of Computer Science, Foundation for Research and Technology, Heraklion , Greece

Presenting author: N.Tavoylaris, email: tavu@ics.forth.gr

ABSTRACT

Localization is an essential piece of information for autonomous mobile robots that is usually computed based on the kinematic information of the platform. For humanoids, however, kinematic computations are often inaccurate, because they do not account for the dynamic effects caused by slippage, discontinuous ground contacts and actuation errors. In this context, visual SLAM algorithms can provide an off-the-shelf and platform independent, method to compute the state of a humanoid accurately, using a light and low-cost sensor that is easily mountable on the robot.

When visual SLAM is applied in the humanoid domain however, it faces extra challenges. This is mainly because the motion of a camera mounted on a humanoid has distinctive differences to the motion models assumed in traditional visual SLAM systems. It has a much wider spread, compared to the one for wheeled robots, and follows the oscillating trajectory of the center of mass, as designated by the bipedal gait [1]. This results in blurriness during the image acquisition process, and reduced performance of the image registration methods, which in turn limit the performance of the visual SLAM system.

To confront the aforementioned problems, we suggest a new dense visual SLAM algorithm for humanoids that monitors the robot's stability, in order to improve the processes of pose estimation and map building. To accomplish this, we extend the Kinect Fusion visual SLAM pipeline, and include a loop closure mechanism, based on pose-graph optimization. For each frame, we evaluate the robot's stability, and create new nodes in the graph only when the robot is stable. To formulate the constraints in each node, we extract 3D match descriptors [2] by focusing on scene locations that contain rich geometric information. Once a loop closure is detected, the algorithm rectifies the pose estimate, based on the location of the tracked features, rollbacks the map to a state prior to optimization, and merges the corrected voxel grid with the existing one. As a result, our algorithm does not allow erroneous map estimates, caused by unstable gaits phases, to be integrated with the rest of the map. Consequently, it protects the map model from being corrupted, thus improving the overall pose estimation process.

We evaluate our algorithm by tracking the pose of an RGB-D camera that is mounted on a NAO humanoid, and executing gait patterns with different oscillating Center of Mass trajectories. In the current work, we present preliminary results for different blocks of the algorithm, including *(i)* a method for evaluating the consistency of a map, by projecting the voxel-grid on the view-port of the camera and comparing it to the actual acquired image, *(ii)* the detection of 3D match features on false normal locations and *(iii)* preliminary results from the loop closure block, using the ICP estimate and covariance as constraints.

REFERENCES

[1]S. Piperakis, S. Timotheatos and P. Trahanias, 2019. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), "Unsupervised Gait Phase Estimation for Humanoid Robot Walking"

[2] Zeng Andy et al. 2017. Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), "3DMatch: Learning Local Geometric Descriptors from RGB-D Reconstructions"