M6 VISUAL ODOMETRY PARAMETERS OPTIMIZATION FOR AUTONOMOUS UNDERWATER VEHICLES

Pep Lluís Negre Carrasco[172], Gabriel Oliver-Codina[171]

Abstract

Visual odometry algorithms are used in a wide range of applications to provide reliable estimates of the movement of the camera frame. Combined with inertial measurement units and other low-cost sensors, they can be used as input of high-level algorithms like Kalman filters or SLAM to estimate the global vehicle pose. However, visual odometers are often complex algorithm pipelines with large parameter sets involved in every process stage. In this paper we focus on optimizing the parameters of a visual odometer that has proved to work in underwater environments. We present the results of a exhaustive optimization process to reduce the errors in rotation and translation comparing vision-based pose estimates with ground truth in real underwater environments.

Keywords – Visual Odometry, Underwater Vehicles, Stereo Images, Motion Estimation

INTRODUCTION

Although there are many visual odometers in the literature, we focus our work on visual odometry systems that use stereo cameras in order to obtain precise 6 DOF camera poses. The common algorithm pipeline [1] for stereo visual odometers is based in the following steps: first, keypoints (landmarks) are identified in each camera frame and feature descriptors for these points are extracted. Then, the depth for every landmark is estimated using stereo, structure from motion or a separate depth camera. Subsequently, features are matched across time frames and the rigid-body transformation that best aligns the features between frames is estimated. The result of this process is an estimation of camera motion between frames and therefore it is necessary to integrate this data over time to obtain the vehicle's absolute position and orientation. Every algorithm stage has its parameter set that must be configured for every application. The present work investigates the effect of these parameters on the final odometry performance and the possible correlations between them.

STEREO VISUAL ODOMETRY PARAMETERS

Libviso2 [2] is a publicly available visual odometer that has proved to work in real-time in Autonomous Underwater Vehicles (AUVs) [3]. This algorithm will be taken as the basis for the parameter optimization process. The total number of configurable parameters for this odometer is 15, so that an overall optimization with 10 possible values for each parameter would result in an intractable amount of combinations. Therefore, it is important to carefully study the influence of each parameter and take into account only those combinations which can result in a significantly improved odometry. In the following we give a brief overview of the main parameters of every stage considered for the study.

Keypoint and feature extraction. The features used by libviso2 are simple blob and corner detector masks which provide a large amount of interest points. To reduce the quantity of features, libviso2 applies a peak threshold technique based on finding local maximums. This technique is controlled by two main parameters: Non-maxima-suppression-window-size defines the quadratic region, centered around the pixel under consideration, used to compute the local maximum search. Non-maxima-suppression-threshold defines the interest point peakiness threshold for which the pixel will be considered a keypoint or not.

Feature matching across frames. The matching process across frames uses current and reference images from left and right cameras to estimate the motion. Libviso2 searches feature coincidences between these four images in order to classify a match as a valid matching. To discard possible outliers between current and reference frames, the algorithm uses the parameter Outlier-flow-tolerance which defines the maximum x/y-disparity, in pixels, between features over the images loop.

Motion estimation. The features extracted in the previous stage are used to compute the camera motion. 3D coordinates of the features in the reference image pair are calculated through triangulation. Gauss-Newton minimization of the reprojection error of these 3D points onto the current left and right images leads to the rigid transformation from the previous image's camera pose to the current camera pose . Since the images used by the algorithm are previously rectified it makes no sense to apply an inlier threshold greater than 1 pixel to discard erroneous 3D reprojection points. However, before the computation of the 3D poses, libviso2 filters the matchings obtained in the previous stage to equally distribute the features over the images. This is achieved using the bucketing technique which consists in dividing the image in a grid of subimages and force a fixed maximum number of matches from each cell. The parameters governing this algorithm are: Bucket-size which defines the grid size and the Maximum-features-per-bucked which defines the maximum number of features per bucked.

The set of possible combinations of these parameters are optimized for experiments carried out in laboratory and real sea conditions. The work exposes the statistical analysis developed to find correlations between parameters and odometry performance, as well as the procedure to easily configure the visual odometer in new underwater scenarios.

REFERENCES

[1] H. Moravec, "Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover," Ph.D. dissertation, Stanford University, 1980.

[2] A. Geiger and J. Ziegler, "Stereoscan: Dense 3d reconstruction in real- time" in IEEE Intelligent Vehicles Symposium, jun 2011.

[3] S. Wirth, P. Negre and G. Oliver, "Visual Odometry for Autonomous Underwater Vehicles" in IEEE Oceans Conference, Bergen (NO), June 2013.