Proposal for Long Term Visual Localization

Yusa Omer Altintop, Irem Arpag

March 15, 2021

1 Problem Definition

Machines need to know their exact position in the world to make intelligent decisions on a global scale and localization system tries to provide it by using visual information to infer the exact position and orientation of the device that is used to capture the image; this idea is the challenge of visual localization, i.e., estimating the position and orientation of the camera pose accurately. Visual localization problem has a vital importance in Computer Vision including autonomous vehicles and multiple approaches are presented to solve it, most of which based on deep learning networks in recent years. The set of training images using to test visual localization algorithms assume to cover all relevant viewing conditions, that is under all probable illumination conditions (dawn, day, sunset, night), seasonal conditions (summer, winter, spring etc.) or weather conditions (snowy, rainy) etc., but in practice it is not the case and additionally many scenes are dynamic over time. More robust algorithms are needed to cope with such long-term visual localization problems that make the challenge more strenuous and motivated.

2 Motivation

- We really like the problem of Long-Term Visual Localization under Changing Conditions as mentioned above and we want to explore the limit of SOTA methods in this problem.

3 Datasets

It was an outstanding problem not to have proper benchmark datasets to handle changes in scene appearance over time in visual localization, but it is over lately with the presentation of "Evaluating Visual Localization in Changing Conditions" in CVPR 2018 that provides benchmark datasets for evaluating 6 Degree-of-Freedom (DOF) pose estimation on-line at www.visuallocation.net . The datasets are available for academic use only not commercial, it is necessary to cite regularly related academic papers, no limitations on the frequency of submission. We are planning to work on RobotCar Seasons dataset, as an outdoor dataset, it is captured in the city of Oxford over a period of 12 months in different weather conditions (rain, night etc.). As SOTA of the dataset Kapture R2D2 APGeM "Robust Image Retrieval based Visual Localization Using Kapture" [1] took the first rank with RobotCar Seasons Dataset with its achievement for the scenario "Visual Localization for Autonomous Vehicles Challenge" performed during CVPR 2020.

4 Baseline

The method consists of two main components; the SFM-based mapping pipeline and components; the SFM-based mapping pipeline and image registration-based localization pipeline. The approach follows the workflow of image retrieval as well as structure based methods and combines functionalities provided by the COLMAP SFM library as well as their local features R2D2 [2] local descriptors which is a sparse key point extractor that jointly performs detection and description but separately estimates key point reliability, and

their global image representation AP-GeM [3], the model participated in ICCV 2019 that directly optimizes the Average Precision (AP) approximated by histogram binning to make it differentiable. Since the method is suitable to evaluate multiple datasets covering a large variety of application scenarios and challenging situations, it facilitates future large scale visual localization researchers to implement and improve their own models by taking it as a base.

REFERENCES

[1] Martin Humenberger, Yohann Cabon, Nicolas Guerin, Julien Morat, Jerome Revaud, Philippe Rerole, Neo Pion, Cesar de Souza, Vincent Leray, Gabriela Csurka. Robust Image Retrieval – based Visual Localization using Kapture. Arvix.2007.1386v2, 2020

[2] Jerome Reuaud, Philippe Weinzaepfel, Cesar De Souza, and Martin Humenberger, R2D2: Reliable and Repeatable Detectors and Descriptors. In NeurIPS 2019

[3] Jerome Reuaud, Jon Almazan, Rafael Sampaio de Rezende, and Cesar Roberto de Souza. Learning with Average Precision: Training Image Retrieval with a Listwise Loss. In ICCV, 2019