ENABLING ROBUST, REAL-TIME VERIFICATION OF VISION-BASED NAVIGATION THROUGH VIEW SYNTHESIS

Marius NEUHALFEN[⋆†‡] *and Jonathan GRZYMISCH*[⋆]
[⋆]*GNC, AOCS and Pointing Division, European Space Agency, ESTEC*
Noordwijk, The Netherlands
[†]*RWTH Aachen University, Aachen, Germany*
neuhalfen.marius@gmail.com · jonathan.grzymisch@esa.int
[‡]Corresponding author

Abstract

In this work, we propose a novel approach for the verification of Image Processing (IP) algorithms used in Vision-Based Navigation (VBN) that works by extending existing datasets with accurate novel viewpoints using View Synthesis.

Today, verification of Image Processing algorithms for VBN mostly relies on surrogate data, either rendered synthetically or acquired in a physical robotic testbed. Both types of data have their own advantages and should be used complementarily to ensure robustness. For closed-loop testing, the real-time processing of these images is essential. However, this is difficult to achieve with both types of imagery mentioned.

Our work addresses this issue. We propose augmenting existing image datasets, acquired offline, in real time by using image transformations. These transformations are used to synthesize novel viewpoints by using the relative pose data of the source images from their associated ground truth labels. This technology opens up multiple use-cases. Firstly, it enables running closed-loop tests for the Image Processing in software without having to rely on performance models. It can by extension also be useful for running large-scale Monte Carlo campaigns. Secondly, it can be used for *densifying* existing datasets, meaning filling holes (regions with lower image density) with synthesized views. Third, it can be used as a Data Augmentation technique for increasing robustness of Neural Networks.

To verify our approach, we seek to obtain a model which relates *distance of View Synthesis* to *quality of synthesized views*. To measure the latter propose a hierarchy of metrics to assess pairwise correspondence between images that are used to compare existing source images to their synthesized counterparts. These metrics are designed to be of particular use for applications relying on Pose Estimation, Feature Tracking and Landmark Matching. The most advanced metric relates the distance of keypoints between the synthesized and existing images. Their detection is performed by a Convolutional Neural Network based on the state-of-the-art High Resolution Network (HRNet) which we adapted for satellite-type imagery. We release this version open-source alongside this work.

For the *distance of View Synthesis* we introduce a new norm between poses which expresses the ability to synthesize a target view from a source one. We call it the *Non-Planar Distance* and it possesses multiple useful properties. Firstly, it fulfills the necessary mathematical properties of a norm on the space of unit quaternions. Secondly, it is normalized between 0 and 1 and can thus be used to compare between datasets even at different pose-scales. Thirdly, it can be inverted to create a *dataset-independent* measure of density from nearest-neighbour distance.

We motivate the development of this measure by a proof of the anisotropy of View Synthesis. We show that the Non-Planar Distance better predicts the accuracy of the View Synthesis (in the sense of keypoint distance) than other commonly used distance norms. The verification is then performed exclusively with publicly available datasets for VBN for which we developed a taxonomy, which we include alongside a compilation of these datasets.

We develop multiple different types of image transformations for View Synthesis depending on available metadata. We compare them using the above mentioned metrics and show that using the most sophisticated technique, IP algorithms can be verified with the augmented data up to performance requirements typical of the mentioned use-cases.