An Evaluation Of Features For Pose Estimation And Its Application to Free Viewpoint Video

Corentin Térence Eloi Chéron, Master of Science in Computer Science

University of Dublin, Trinity College, 2018

Supervisors: Dr Konstantinos Amplianitis and Prof Aljosa Smolic

As consumers start using augmented and virtual reality with personal devices, new ways of creating high-quality 3D textured videos (free viewpoint videos) become necessary. Smartphone video clips of an actor from different viewpoints can affordably produce FVV even in outdoor environments. However, the cameras need to be accurately localised, and current methods require up to twelve cameras with small baseline angles to produce high-quality models. In this dissertation, we evaluate the capacity of different hand-crafted and learned features to estimate relative 3D pose, a critical step in the Structure-from-Motion reconstruction process. We developed a unified workflow based on COLMAP to compare those features with the highest rigour and match the FVV production pipeline. We evaluate various configurations for each of the SIFT, ASIFT, Sift-Affine, LIFT and SuperPoint features against new wide-viewpoint datasets with varying geometric complexity. The results show that the traditional hand-crafted features SIFT and SIFT-Affine are the most efficient to estimate wide-baseline camera poses regarding the number of keypoints. SuperPoint overtakes LIFT and reaches state-of-the-art performances in some configurations and shows an impressive match ratio in all the situations but fails to register cameras with the widest baseline. When applied to an FVV dataset, SIFT provides the best speed due to its refined implementation. Using the latest work in semantic segmentation, we evaluate the effect of matching feature regrouped semantically and show an improvement in the pose accuracy. The improvement observer from LIFT to SuperPoint and recent work on auxiliary learning applied to camera relocalisation show good promises in designing a new deep learning feature for wide-baseline applications.