3D Affine: An Embedding of Local Image Features for Viewpoint Invariance Using RGB-D Sensor Data

Local image features are invariant to in-plane rotations and robust to minor viewpoint changes. However, the current detectors and descriptors for local image features fail to accommodate out-of-plane rotations larger than 25°–30°. Invariance to such viewpoint changes is essential for numerous applications, including wide baseline matching, 6D pose estimation, and object reconstruction.

For that, we proposed a general embedding [1] that wraps a detector/descriptor pair in order to increase viewpoint invariance by exploiting input depth maps. The proposed embedding locates smooth surfaces within the input RGB-D images and projects them into a viewpoint invariant representation, enabling the detection and description of more viewpoint invariant features (See Figs. 1-3). Our embedding can be utilized with different combinations of descriptor/detector pairs, according to the desired application.

While standalone local image features fail to accommodate average viewpoint changes beyond 33.3°, our proposed embedding boosted the viewpoint invariance to different levels, depending on the scene geometry. Objects with distinct surface discontinuities were on average invariant up to 52.8°, and the overall average for all evaluated datasets was 45.4°. Similarly, out of a total of 140 combinations involving 20 local image features and various objects with distinct surface discontinuities, only a single standalone local image feature exceeded the goal of 60° viewpoint difference in just two combinations, as compared with 19 different local image features succeeding in 73 combinations when wrapped in the proposed embedding. Furthermore, the proposed approach operates robustly in the presence of input depth noise, even that of low-cost commodity depth sensors, and well beyond.

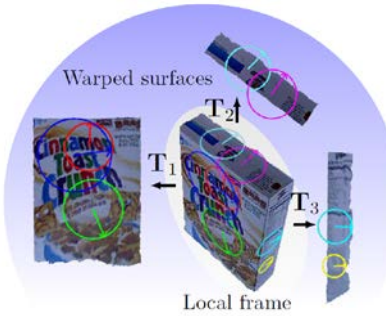

Figure 1: Keypoints are detected on the warped surfaces are remapped back to the original input local frame.

Figure 2: Before improving the viewpoint invariance of SIFT features.

Figure 3: After improving the viewpoint invariance of SIFT features.

Keywords: viewpoint invariance; local image feature embedding; wide baseline matching.

References

[1] Sahloul,Hamdi, Shirafuji,Shouhei, & Ota,Jun. (2019). 3D Affine: an embedding of local image features for viewpoint invariance using RGB-D sensor data. Sensors, 19(2), 291, 1-32.