Enhancing pose estimation through efficient patch synthesis

Pierre Rolin
https://members.loria.fr/PRolin/

Marie-Odile Berger
http://www.loria.fr/~berger/

Frédéric Sur
https://members.loria.fr/FSur/

Université de Lorraine

INRIA

Université de Lorraine

Estimating the pose of a camera from a scene model is a challenging problem when the camera is in a position not covered by the views used to build the model, because feature matching is difficult. Several viewpoint simulation techniques have been recently proposed in this context. They generally come with a high computational cost, are limited to specific scenes such as urban environments or object-centred scenes, or need an initial pose guess. This paper presents a viewpoint simulation method well suited to most scenes and query views. Two major problems are addressed: the positioning of the virtual viewpoints with respect to the scene, and the synthesis of geometrically consistent patches.
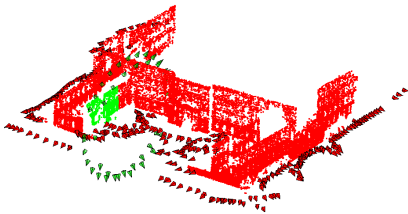

Figure 1: Virtual viewpoints (in green) positioning relative to one of the segmented patches (green points). The red viewpoints are the ones used to construct the model.

We propose a method to position the virtual viewpoints with respect to a segmentation of the scene in planar parts, the only assumption being that the scene is piecewise planar, which is not restrictive in most human-made environments. A set of virtual viewpoints is associated with each planar part. An adapted measure for viewpoint changes ensures that the existing viewpoints are completed with relatively few virtual viewpoints. Figure 1 illustrates the virtual viewpoints positioning relative to a segmented patch.

Viewpoint simulation techniques based on a scene model often generate images of the scene using a dense model [2] or generate local patches around every interest point [1]. The first approach fails in cases where some parts of the scene are not correctly densified, and the latter is computationally demanding without any pose guess. We propose an intermediate approach consisting in synthesizing semi-local planar patches of the scene and enriching the scene model with descriptors from these synthesized patches, using a visibility constraint.

Figure 2: In this scene the pose cannot be accurately estimated with a standard RANSAC-PnP approach (left). Patch synthesis improves the registration of the query view to the scene, as proved by the superposed scene edges (right).

The experiments show that a model enriched with the proposed synthesis method leads to more accurate poses, and even gives accurate poses when pose estimation simply fails without synthesis. In addition the time needed for pose computation is reduced, since the RANSAC step needs less iterations. Figure 2 shows an example where patch synthesis gives a dramatic improvement of pose estimation.

[1] P. Rolin, M.-O. Berger, and F. Sur. Viewpoint simulation for camera pose estimation from an unstructured scene model. In *Proc. International Conference on Robotics and Automation (ICRA)*, pages 6320–6327, 2015.

[2] A. Torii, R. Arandjelović, J. Sivic, M. Okutomi, and T. Pajdla. 24/7 place recognition by view synthesis. pages 1808–1817, June 2015.