EEL 5406 - Computational Photography, Fall 2020

Mert Canatan
Department of Electrical and Computer Engineering
Master of Science Student

December 14, 2020

In this project, my group and I selected the article "Revealing Scenes by Inverting Structure from Motion Reconstructions" [1], because we were interested in doing Stucture from Motion (SfM) and we found interesting results in the paper. In the article, authors showed that 3D point clouds retain enough information to reveal scene apperance and enough to compromise privacy. They reconstructed scenes using sparse 3D point clouds. They also tried different variations of the different input features (such as SIFT descriptor, Depth, and RGB; SIFT descriptor, and Depth; etc.) and sparsity levels to show the effect on the reconstruction quality.

In the article, the method is based on a cascaded U-Net that takes as input, a 2D multichannel image of the points rendered from a specific viewpoint containing point depth and optionally color and SIFT descriptors and outputs a color image of the scene from that viewpoint. The novelty in this paper is that authors deal with highly sparse and irregular 2D point distributions and inputs where many point attributes are missing, namely keypoint orientation and scale, descriptor image source and the 3D point visibility.

Proposed neural network architecture to train the input features consists 3 different architecture namely VisibNet, CoarseNet and RefineNet. VisibNet is used to identify a occluded part of an image (when it is transformed from 3D to 2D) and remove those parts for visibility estimation. CoarseNet is used to produce RGB images of same dimensions as the input feature map for image reconstruction, and RefineNet is used to produce a refined RGB images for higher quality images.

Results of the generated images from the proposed model are extremely realistic in terms of visual quality.

I personally learnt how to read and understand a research paper and implement the paper. In my opinion, 3D point clouds can be dangerous to keep since it can be used to reveal the original scene as we have seen in the paper. More secure ways (cryptography techniques might be useful) to save these 3D point clouds need to be developed to secure the privacy.

Reference: [1] Pittaluga, F., Koppal, S.J., Kang, S.B., Sinha, S.N.: Revealing scenes by inverting structure from motion reconstructions. In: Proc. CVPR. (June 2019)