Extreme View Synthesis

Abraham Jose

1 Summary

The paper introduces a novel view extrapolation technique using the depth probability volume based wrapping and CNN based image patch refinement. The depth probability volume allows the method to to leverage depth uncertainty in challenging regions such as depth discontinuity. It has a wrapping and fusing as well as a refinement stage of which the later is a novel proposal. The technique could extrapolate an view with least number of views(as low as two input camera views) and extrapolate the new view with baseline magnification of 30x. The refining network uses UNet to get the state of the art result.

2 Good points

The author's strongest point is that the view synthesis is possible with least number of input images/ views as low as 2 and with baseline magnification of 30x. The depth probability volume helps to carry more information about the depth-discontinuities and the back-to-front image reconstruction and overwriting the pixels if it falls in the same pixel location accounts for occlusion of objects and finely refined edges. Also, since the technique uses traditional wrapping and even if it uses CNN based refinement on the patches, the results are consistent with least ambiguity and does not requires domain adaptation compared to fully deep learning based solutions.

3 Weak points

The technique requires 1.40 minutes to generate the a synthesized view and the execution time can stretch depending on the size of the image and the number of patches in the image. Also, the system is not capable of reconstruction of pixel areas outside the frusta of the input camera also it struggles with artifacts that looks natural, like an entire reconstruction in wrong location.

4 Questions

The results does not shows qualitative results alongside with the existing techniques other than SM(like Soft3D). Also, the GTA-V MVS-synth dataset is a game simulation and is smoother and less complex compared to real life scenarios
Which other techniques can be used to reduce the execution time? Also, why did the author chose 64x64 patch for refinement network?

5 Ideas

We can replace the UNet with a GAN to generate high quality patches which can also account for the pixel regions that is outside of the frusta and to account for depth-quantization artifacts.