Project Proposal: Enhanced View-Synthesis by Appearance Flow

Shradha Agrawal
A53105993
sha015@eng.ucsd.edu

Sudhanshu Bahety
A53209213
sbahety@eng.ucsd.edu

Jean Choi
A53214201
jsc078@eng.ucsd.edu

Ajitesh Gupta
A53220177
ajgupta@eng.ucsd.edu

Nimish Srivastava
A53222968
n2srivas@eng.ucsd.edu

1 Idea Proposal

Given an input image, we plan to synthesize new images of same object or scene observed from arbitrary viewpoints. This can be helpful for 3D reconstruction where we do not have abundance of images of objects from various viewpoints. This could also enable photo editing programs like Photoshop to manipulate objects in 3D instead of 2D and help create full virtual reality environments based on historic images or video footage.

Catering to this task, we wish to extend the recent work of Zhou et al[1] on View-Synthesis from Appearance flow, drawing inspiration from Single-view to multi-view paper by Tatarchenko et al[2]. We believe that the current appearance flow system[1] lacks the capability to hallucinate pixels which are not present in the input view, as pointed out by the authors themselves. We first intend to pre-train the appearance flow network (Figure 2) in an autoencoder[1] style using only 3 RGB channels of output layer. Thereafter, while fine-tuning we plan to use the reconstructed image (as it may contain hallucinated information for the new viewpoint) instead of the original input for warping with generated appearance flow, to produce synthesized view. This would help the system to learn better representations of the objects and we hope this will enable it to hallucinate parts of the object (required in output view) which were not present in the input image.

2 Dataset

In original work[1], the authors use 7499 cars and 700 chairs from ShapeNet[4] dataset. However, 2D rendered images of ShapeNet 3D models are not available online. As we will have to render these images from 3D models which might require additional AWS resources so we plan to use Chairs dataset by Aubry et. al [3]. At present we plan to show a comparative study between baseline and our approach on the Chairs dataset[3] only.

*If we get more AWS resources in the future, we plan to render the ShapeNet[4] data as used in [1] and compare our results with the reported results in [1].

Description of Data

The data contains 1393 chair styles/models, each rendered from 62 viewpoints: 31 azimuth angles (with step of 11 degrees) and 2 elevation angles (20 and 30 degrees), with a fixed distance to the chair. The resulting chair dataset contains a total of 1, 393 * 62 = 86, 366 rendered images, each annotated

[1] autoencoder style: combination of encoder decoder network which have proven efficiency in reconstructing input.

with the chair ID indicating the different style, as well as the viewing orientation. The original dataset can be downloaded from: `https://www.di.ens.fr/willow/research/seeing3Dchairs/`

3 Network Architecture

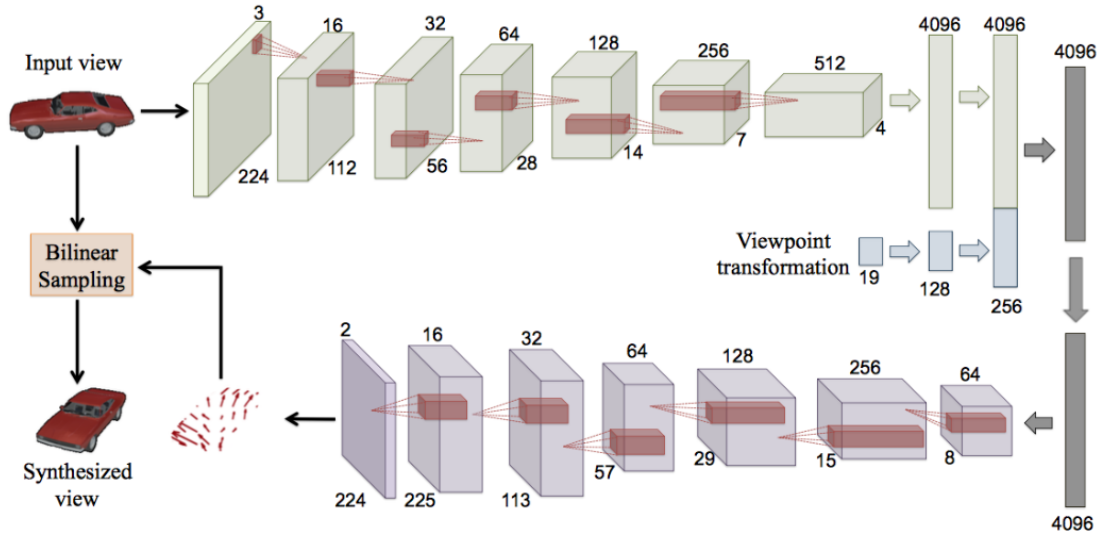

Figure 1: This figure illustrates the present architecture of the network[1]. We plan to extend this network to include 3 more channels for reconstructed image in output layer. And to produce synthesized view, we plan to warp generated appearance flow with reconstructed image (3 extra channels of output layer) instead of input image view. We hope that with reconstructed image view, we will be able to hallucinate pixels which aren't available in input image view.

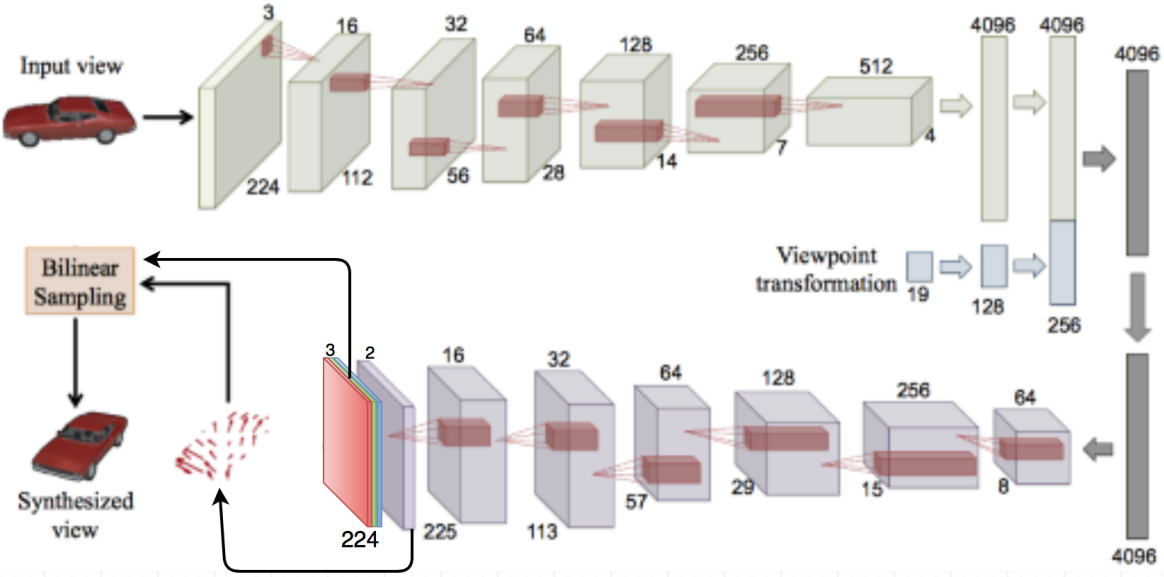

Figure 2: This figure illustrates our proposed architecture. The architecture shows 3 added RGB channels for reconstructed image in output layer. The generated appearance flow (2 channels of output layer) will be warped with these 3 channels to produce output view.

References

1. Tinghui Zhou, Shubham Tulsiani, Weilun Sun, Jitendra Malik, Alexei A. Efros - View Synthesis by Appearance Flow - In ECCV 2016
2. Tatarchenko, M., Dosovitskiy, A., Brox, T.: Single-view to multi-view: Reconstructing unseen views with a convolutional network. arXiv preprint arXiv:1511.06702 (2015)
3. M. Aubry, D. Maturana, A. Efros, B. Russell and J. Sivic - Seeing 3D chairs: exemplar part-based 2D-3D alignment using a large dataset of CAD models - CVPR, 2014: Download Link: https://www.di.ens.fr/willow/research/seeing3Dchairs/
4. Chang, Angel X., et al. "Shapenet: An information-rich 3d model repository." arXiv preprint arXiv:1512.03012 (2015).