

The model takes last frame as input, how to deal with the first frame and why?

For current frame estimation the network takes the previous high resolution prediction as input. Since the previous frame is not available for the first frame sequence the initialization of the first previous high resolution input can be done as follows:

1. All entries set to zero,
2. Default values: mask=0, normal=[0,0,1], AO=1,
3. An upscaled version of current inputs.

The first approach is the most common approach. The paper stated that there is not much visual difference in the first frame between networks trained with these three options. So they used the simplest option one.

2. Why is pretrained VGG-19 possible to work as perceptual loss ?

VGG 19 is a neural network model trained to recognize objects in color images in the space . Perceptual loss is a metric to evaluate how two image have similar activations in the latent space of a pretrained network to judge how two image are similar. So VGG 19 is a neural network and perceptual loss is a metric, these are two different thing with different purpose.

But VGG 19 does use in perceptual loss because it has convolution layers which are able to extracts the layer activations when feeding the image into the network.

3. Loss functions like perceptual loss or GAN loss which works in other problems are not generating good results in the paper. Why do you think this happens? What is a good way to choose loss function?

I think only the GAN loss is typically not sufficient, dueto the plateaus if the discriminator fails to provide gradients or to stabilize the optimization in the beginning.

To combine the GAN loss with a perceptual loss on the normals and L1 loss on the mask and ambient occlusion with a small weighting factor.

A loss function is used to penalize differences between the high-resolution learned and ground-truth variants. It is a good way to apply multiple loss functions to different channels of the predicted and ground truth, and use a weighted sum of the losses functions

4. We can see that error increases when we add AO effect. In possible application of this work, is it beneficial to have estimated AO effect in sacrifice of precision?

My answer is yes, since AO is one of the most important global illumination effects for isosurfaces, it estimates for every surface point the attenuation of ambient light from the surrounding and uses this information to enhance cavities and locations closely surrounded by other surface parts.

5. Generally speaking, predicting the error of ANN predictions are impossible. How does this influence its application?

If we cannot know the uncertainty as well, sometimes it leads to some bad consequences, for example, self-driving, if the model cannot estimate how much likely the prediction is wrong, it may make some wrong decision.