

Dynamic Textured Billboards for Ever-Evolving Virtual Worlds

Supplementary Material

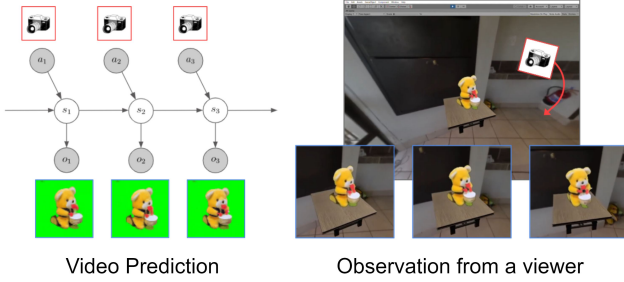


Figure 8

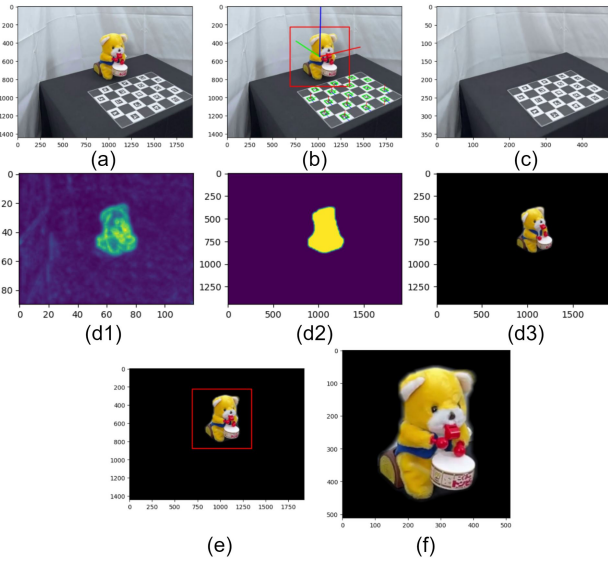


Figure 9

9. Implementation Details

9.1. Video Prediction and Synthesis

The overall image of our system is shown in Figure 8, and the flow of building our dataset is shown in Figure 9. For video prediction, we used PlaNet [26], one of the Deep State Space Models, which is a time series prediction model that extends the Variational Auto Encoder in the time sequence direction, and given an input series, predicts the following series one step at a time. PlaNet is used in the field of reinforcement learning and is designed to withstand video prediction. We then use DSSM to train a prediction of the video as the viewpoint changes and use it as a billboard renderer. In preparing the training data, a ChArUco marker is placed in front of the object to simultaneously detect the camera and the object's position. We use this method because it is

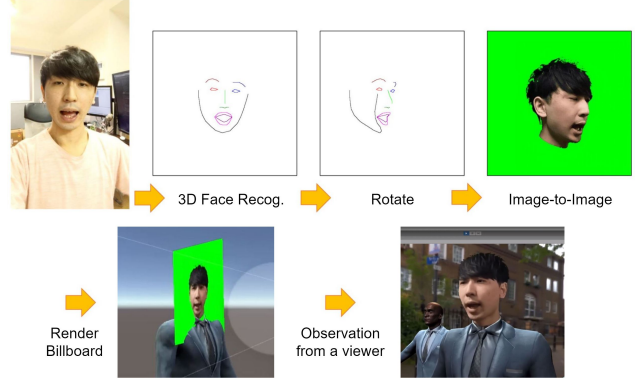


Figure 10

difficult for COLMAP to work well with data taken of moving objects. We also use NeRF to generate a background for each viewpoint, detect the object's region based on the subtraction from the background, and create data that are normalized to the object's size and position in the video. During training, the system predicts 10 steps of future video, but when used as a billboard renderer, it predicts infinite steps of video.

9.2. Pseudo 3D Face Avatar

The overall image of our system is shown in Figure 10. For Image-to-Image translation, we used Pix2PixHD [69]. Pix2Pix is a UNet-base GAN for style transfer, and Pix2PixHD improves it to withstand high-resolution generation. From 10 minutes of video data, we obtain normalized videos and facial landmarks for each frame. We train Pix2PixHD to generate the original image from the landmarks. When utilizing the model as a billboard renderer, we obtain 3D facial landmarks from the camera, prepare 2D landmark images from the observer's viewpoint by rotating 3D landmarks, apply Pix2Pix, and render the result on the billboard. As mentioned in the main text, please note that only this demo is not currently running in real-time and is a result of post-rendering.