# Supplementary Material

- Dataset description:

  - Around 67 3D models for training and 10 models for testing were selected from Google warehouse 3D. The models consisted of surfaces with varying textures from textureless to repeated textured surfaces.

  - The shapes of models were carefully selected to include good variety in 3D shapes from House models , sofa sets, chairs, to action figures, cartoon characters. These models were of varied scales, hence to reduce complexity of the dataset, the models were rendered at almost the same scale.

  - The lighting conditions were kept simple with one light source illuminating the scene from the camera's perspective.

  - 900 samples of each model consisting of 30 random orientations about the horizontal axis and 30 random rotations about the vertical axis with angle less than or equal to 15 degrees. Since occlusions are limiting factors in computing correspondences for unrectified images, we had to constrain the rotation angle.

- Training Process:

  - With a constant learning rate of 0.001, the model was trained on Caffe framework on Nvidia Titan X Gpu for 10 epochs. We observed that the training and validation errors saturated at the end of 10 epochs.

  - The error function used was a simple pixel-wise L2 error used by the Flownet architecture. The fully convolutional architecture helps in predicting pixel-wise depth values similiar to flow values predicted by the Flownet architecture.

More Results
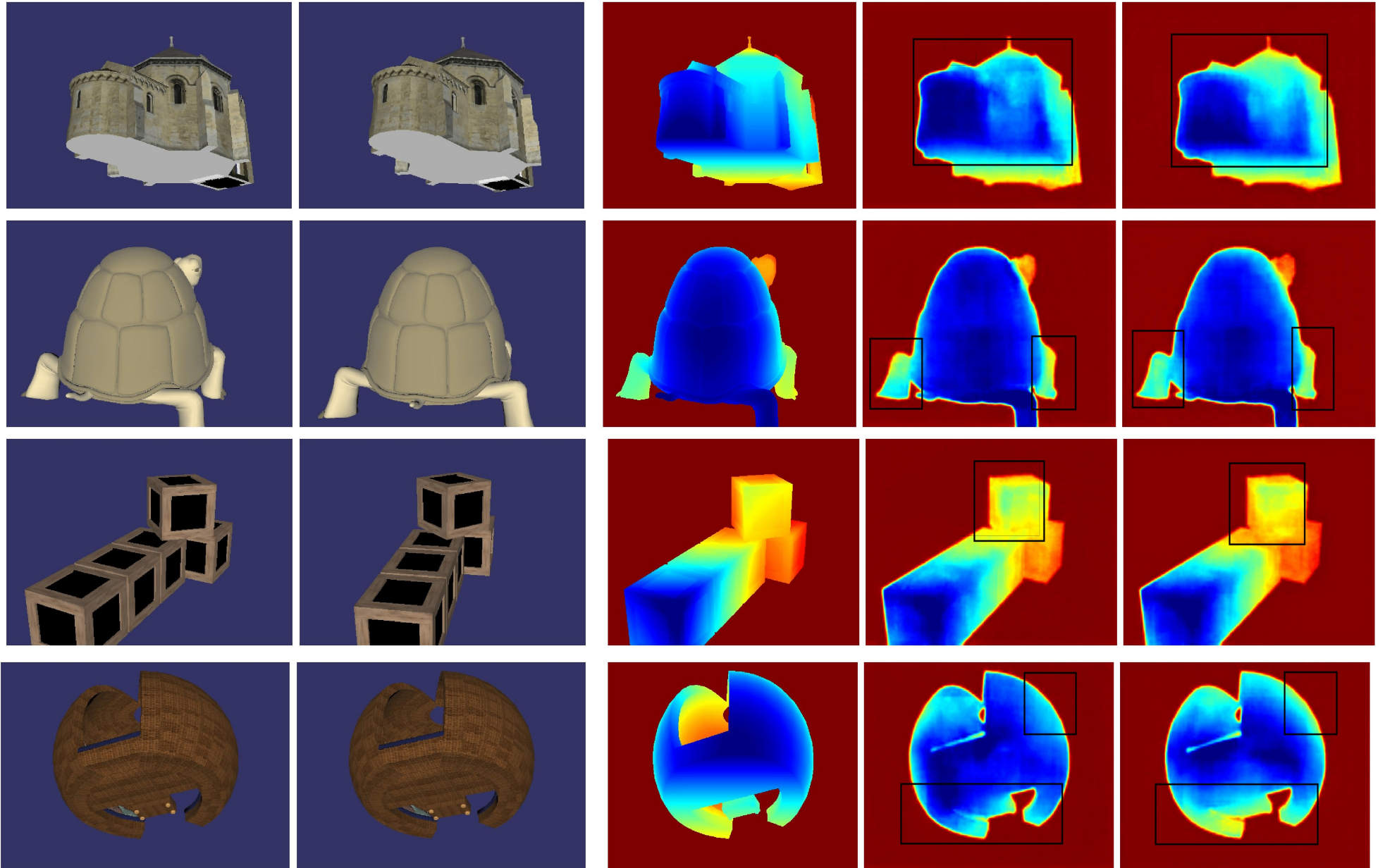
| Left Image | Right Image | Ground Truth | Our Model | FlownetCorr |