Rendering a Dataset with Transparent Objects

- A transparent object takes its appearance not only from its backdrop, but from all its surroundings.
- A validated technique for photorealistic rendering of transparent objects is available.
- Collecting public domain High Dynamic Range (HDR) environment maps and shapes and generating random shapes, we can render a representative dataset including ground truth training data.
- We rendered 80,000 images of the 600 shapes observed from randomly chosen viewpoints. For each RGB-image, we also render segmentation mask, depth, and normal layers as ground truth.

Encoder-Decoder CNN Based on VGG16

- The output of our network is a segmentation mask, which separates the transparent object from the background, and also a normal map and a depth map. We train a separate encoder-decoder for each task.
- The following example predictions (when training with diffuse versus transparent object from the backdrop, but not fine tuned for specific devices or environments.

Results

- 0.9 threshold used to binarize the mask.
- Relative depth in [0, 1]: nearest to furthest visible part.
- Depth and normal predictions are masked.
- The error images below display pixelwise loss in the masked regions according to the loss functions, but mapped to a range from 0 (dark blue) to 1 (dark red).
- We used a fixed index of refraction of 1.5 for all shapes in our dataset. Nevertheless, we observe by testing that our network achieves low test errors across a wide range of refractive indices [1.3, 2.0].
- Testing "in the wild" using photographs of glass object.
- Note that our dataset currently does not include thin and hollow shapes.
- Our network was trained entirely with synthetic data, and not fine tuned for specific devices or environments.