TeRF: Thermal Radiance Field for Rendering Depth in Transparent Objects

Jeongyun Kim[1], Yunho Choi[1] and Ayoung Kim[1*]

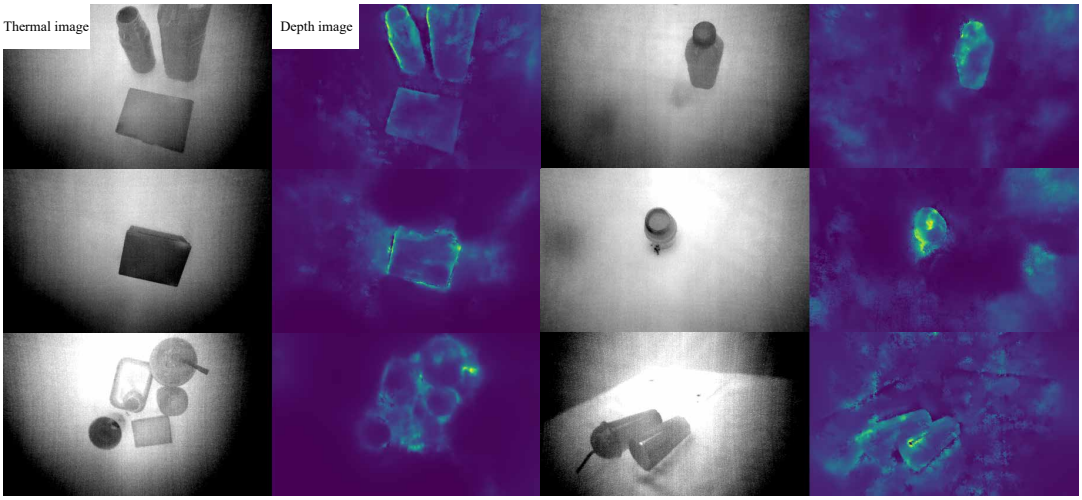

Fig. 1: TeRF estimates depth for various transparent objects made of plastic, glass. The left images show the normalized thermal image, and the right images show the depth rendering from the image view.

***Abstract*— Research on object manipulation using robots has been steadily progressing. Transparent objects hardly appear clearly in RGB images due to reflection and refraction; therefore, most of studies have targeted opaque objects. In particular, the depth sensor has a limitation on recovering depth of transparent objects due to noise. In this paper, we propose Thermal Radiance Field (TeRF) to obtain depth on transparent objects using a powerful thermal imaging camera.**

I. Introduction

Object manipulation using a robot equipped with a sensor system is being used in various places, such as logistics centers and factories. Transparent objects made of glass and plastic refract and reflect light, and they rarely appear clearly on RGBD sensors. In particular, the depth sensor generates incorrect depth in the case of transparent objects. Hence, it is challenging to manipulate transparent objects, so object manipulation objects are mostly opaque. Using transparent objects such as glasses and plastic bottles is unavoidable. Consequently, it is essential to operate transparent objects. In this paper, we solve this problem by using a thermal camera that can ignore the effect of reflection and refraction of light, which is the limitation of transparent objects. The advantage of using a thermal camera is that transparent objects appear clearly and robust, even in an environment without light.

Even with a thermal camera, one problem remains. Depth information is necessary to manipulate an object, but the depth sensor cannot obtain a transparent body's depth. Neural Radiance Fields (NeRF) [1] successfully estimated the depth using a supervised learning-based method. Thus, we used the NeRF because most environments have no ground truth depth.

[†]This work was supported by IITP grant (No.2022-0-00480).

[1]J. Kim, Y. Choi and A. Kim are with the Department of Mechanical Engineering, SNU, Seoul, S. Korea [jeongyunkim, potatohead12, ayoungk]@snu.ac.kr

II. Method and Experiment

TeRF based on Plenoxel [2] without a neural network utilizes the thermal images and the camera pose as inputs. The thermal image shows the temperature, and because the temperature difference between the object and the background is small, the performance deteriorates when raw data is used. Therefore, to increase the temperature difference, we normalized it based on the section where the temperature is concentrated. We performed depth rendering by optimizing temperature and density using camera pose and thermal images.

We coupled a thermal camera to a robotic arm to acquire data sets of transparent objects in various environments, such as one or more objects in a low-light environment. Fig. 1 shows the results of depth estimation using TeRF on 5 datasets. Objects can be manipulated based on the estimated depth of objects.

References

[1] Mildenhall and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," in *Proc. European Conf. on Comput. Vision*. Springer, 2020, pp. 405–421.

[2] Fridovich-Keil and A. Kanazawa, "Plenoxels: Radiance fields without neural networks," in *Proc. IEEE Conf. on Comput. Vision and Pattern Recog.*, 2022, pp. 5501–5510.