Super-resolution of Voxel Model Using 3D ESRGAN

Ryosuke Ueda, Yuta Muraki,

***Abstract*— One of the techniques for 3D computer graphics representation is voxel representation. Voxels are a simple set-based operation, making them suitable for drawing irregular 3D models. Super-resolution is a method used to generate high-resolution data from low-resolution data. While numerous studies have been conducted on super-resolution for images, research in the context of 3D computer-generated models remains limited. Therefore, we aim to extend 2D super-resolution techniques to the realm of 3D, enabling high-precision super-resolution for voxel models.**

Keywords-voxel models, super-resolution, GAN

I. Introduction

One of the techniques for 3D computer graphics representation is voxel representation. Voxels are a simple set-based operation, making them suitable for drawing irregular 3D models. However, a high-definition 3D model requires a huge amount of data and requires time and effort to create. As method for converting low-resolution data to high-resolution data is super-resolution. It has been studied for 2D images, such as SRGAN [1] and ESRGAN [2]. Furthermore, Oka et al.'s method [3] extends SRGAN to the 3D domain, achieving super-resolution for voxel models. However, there are problems with the output models, such as areas with holes and missing voxels. Therefore, our proposed method extends the 3D adaptation of ESRGAN, which can achieve higher-precision super-resolution than SRGAN. This extension aims to achieve super-resolution for voxel models with greater accuracy compared to conventional methods.

II. Proposed Methods and Experimental Results

The proposed method is an extension of ESRGAN to a voxel model super-resolution technique, utilizing a similar structure to that of a GAN. First, we explain about 2D ESRGAN. In ESRGAN, by incorporating Residual in Residual Dense Block into the generator, unnecessary noise is reduced, allowing for the construction of deeper models. Additionally, the discriminator adopts the Relativistic Discriminator, enabling stable training. The proposed method extends the convolutional layers within the generator and discriminator, as well as the Batch Normalization layers within the discriminator, for 3D adaptation, is shown in Figure 1, Figure 2. The training data consists of pairs of 16×16×16 low-resolution models and 64×64×64 high-resolution models. In the experiments, the accuracy of the proposed method was compared with that of the conventional method. The evaluation method involves comparing the average values of four evaluation index precision, recall, Dice coefficient, and IoU, across 100 test data samples. Furthermore, an example of both output results is shown in Figure 3. The parameters configured for training were as follows: 500 epochs, batch size of 4, and a learning rate of 0.2**4. As shown in Table 1, the proposed method outperformed the conventional method. Some models even demonstrated accuracy improvements exceeding 10%. Furthermore, the proposed method successfully addressed the issue of hole-filled regions present in the conventional method, as shown in Figure 3(b).

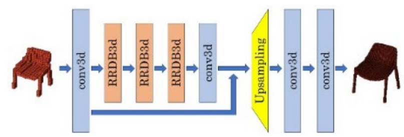

Figure1. Generator of flowchart

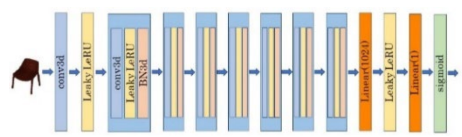

Figure2. Discriminator of flowchart

(a) True (b) 3D-SRGAN (c) Ours

Figure3. An example of the output result

TABLE I. EVALUATION INDEX (%)

Method	Average value			
	Precision	***Recall***	***F-score***	***IoU***
3D-SRGAN	65.70	55.85	59.97	45.26
Ours	65.92	58.74	61.77	47.57

References

[1] Ledig, Christian, et al. "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network,", Conference on Computer Vision and Pattern Recognition (CVPR) 2017.

[2] Xintao Wang, et al. "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks". Conference on Computer Vision and Pattern Recognition (CVPR) 2018.

[3] Kazutoshi Oka, et al. "3D-SRGAN: Application of SRGAN to Super-Resolution for 3-Dimensional Voxel Models", The journal of the Institute of Image Electronics Engineers of Japan: visual computing, devices & communications, vol. 48, no. 4, pp. 488-496, 2019-10.

Ryosuke Ueda is with the Graduate School of Information Science and Technology, Osaka Institute of Technology, Osaka, Japan, (e-mail: m1m22a07@oit.ac.jp).

Yuta Muraki is with the Graduate School of Information Science and Technology, Osaka Institute of Technology, Osaka, Japan, (e-mail: yuta.muraki@oit.ac.jp).