

# 효율적인 feature map 추출 네트워크를 이용한 2D 이미지에서의 3D 포인트 클라우드 재구축 기법 3D Point Cloud Reconstruction Technique from 2D Image Using Efficient Feature Map Extraction Network

김 정 윤\*, 이 승 호\*\*

Jeong-Yoon Kim\*, Seung-Ho Lee\*\*

## Abstract

In this paper, we propose a 3D point cloud reconstruction technique from 2D images using efficient feature map extraction network. The originality of the method proposed in this paper is as follows. First, we use a new feature map extraction network that is about 27% efficient than existing techniques in terms of memory. The proposed network does not reduce the size to the middle of the deep learning network, so important information required for 3D point cloud reconstruction is not lost. We solved the memory increase problem caused by the non-reduced image size by reducing the number of channels and by efficiently configuring the deep learning network to be shallow. Second, by preserving the high-resolution features of the 2D image, the accuracy can be further improved than that of the conventional technique. The feature map extracted from the non-reduced image contains more detailed information than the existing method, which can further improve the reconstruction accuracy of the 3D point cloud. Third, we use a divergence loss that does not require shooting information. The fact that not only the 2D image but also the shooting angle is required for learning, the dataset must contain detailed information and it is a disadvantage that makes it difficult to construct the dataset. In this paper, the accuracy of the reconstruction of the 3D point cloud can be increased by increasing the diversity of information through randomness without additional shooting information. In order to objectively evaluate the performance of the proposed method, using the ShapeNet dataset and using the same method as in the comparative papers, the CD value of the method proposed in this paper is 5.87, the EMD value is 5.81, and the FLOPs value is 2.9G. It was calculated. On the other hand, the lower the CD and EMD values, the better the accuracy of the reconstructed 3D point cloud approaches the original. In addition, the lower the number of FLOPs, the less memory is required for the deep learning network. Therefore, the CD, EMD, and FLOPs performance evaluation results of the proposed method showed about 27% improvement in memory and 6.3% in terms of accuracy compared to the methods in other papers, demonstrating objective performance.

## 요 약

본 논문에서는 효율적인 feature map 추출 네트워크를 이용한 2D 이미지에서의 3D 포인트 클라우드 재구축 기법을 제안한다. 본 논문에서 제안한 기법의 독창성은 다음과 같다. 첫째로, 메모리 측면에서 기존 기법보다 약 27% 더 효율적인 새로운 feature map 추출 네트워크를 사용한다. 제안하는 네트워크는 딥러닝 네트워크의 중간까지 크기 축소를 수행하지 않아, 3D 포인트 클라우드 재구축에 필요한 중요한 정보가 유실되지 않았다. 축소되지 않은 이미지 크기로 인해 발생하는 메모리 증가 문제는 채널의 개수를 줄이고 딥러닝 네트워크의 깊이를 얇게 효율적으로 구성하여 해결하였다. 두 번째로, 2D 이미지의 고해상도 feature를 보존하여 정확도를 기존 기법보다 향상시킬 수 있도록 하였다. 축소되지 않은 이미지로부터 추출한 feature map은 기존의 방법보다 자세한 정보가 담겨있어 3D 포인트 클라우드의 재구축 정확도를 향상시킬 수 있다. 세 번째로, 촬영 정보를 필요로 하지 않는 divergence loss를 사용한다. 2D 이미지뿐만 아니라 촬영 각도가 학습에 필요하다는 사항은 그만큼 데이터셋이 자세한 정보를 담고 있어야 하며 데이터셋의 구축을 어렵게 만드는 단점이다. 본 논문에서는 추가적인 촬영 정보 없이 무작위성을 통해 정보의 다양성을 늘려 3D 포인트 클라우드의 재구축 정확도가 높아질 수 있도록 하였다. 제안하는 기법의 성능을 객관적으로 평가하기 위해 ShapeNet 데이터셋을 이용하여 비교 논문들과 같은 방법으로 실험한 결과, 본 논문에서 제안하는 기법의 CD 값이 5.87, EMD 값이 5.81 FLOPs 값이 2.9G로 산출되었다. 한편, CD, EMD 수치가 낮을수록, 재구축한 3D 포인트 클라우드가 원본에 근접하는 정확도가 향상된 결과를 나타낸다. 또한, FLOPs 수치가 낮을수록 딥러닝 네트워크에 필요한 메모리가 적게 소요되는 결과를 나타낸다. 따라서, 제안하는 기법의 CD, EMD, FLOPs 성능평가 결과가 다른 논문의 기법들보다 메모리 측면에서 약 27%, 정확도 측면에서 약 6.3% 향상된 결과를 나타내어 객관적인 성능이 입증되었다.

*Key words : Point Cloud, Feature Map, Reconstruction, Reparameterization Trick, Latent Vector, Deep Learning*

## I. 서론

최근 메타버스를 비롯한 3D 응용 분야의 시장 규모가 급격히 성장하고 있어, 3D 모델의 수요 또한 급격히 늘어나고 있는 추세이다. 가상세계에 현실과 똑같은 공간을 만들기 위해서 3D 모델 데이터의 표현을 최대한 자세하게, 최대한 효율적으로 나타내는 것이 더욱 중요해지고 있다. 기존의 3D 모델 데이터를 획득하는 방식으로는 대표적으로 LiDAR와 여러 각도에서 촬영한 여러 장의 이미지를 기반으로 영상처리 알고리즘을 이용한 방식이 있다. 그러나 단편적인 정보만으로 보이지 않는 부분을 재구축하는 것에 많은 어려움이 있었다. 이후, 인공지능이 등장하고 여러 분야에 응용되기 시작하면서 딥러닝 기반의 2D 이미지를 이용한 3D 모델 재구축 방식이 등장하였다. Choy 등[1]은 RNN(Recurrent Neural Network)으로 2D 이미지를 3차원 격자 형태의 3D 복셀 방식으로 재구축하여 딥러닝 기반의 2D 이미지를 이용한 3D 모델 재구축의 정확도가 기존 방식을 뛰어넘을 수 있다는 것을 입증하였다. Fan 등[2]은 2D 이미지를 3D 포인트 클라우드로 재구축하는 딥러닝 네트워크의 학습 방법을 발표하였다. 3D 포인트 클라우드는 사물의 윤곽을 형성하는 각각의 점들을 소수점 좌표로 표현하는 방법으로, 3D 격자에 0 또는 1로 형태를 나타낸다. 이후, Mandikal 등[3]은 입, 출력이 같아 정확도가 높은 VAE(Variational AutoEncoder)[4]의 latent vector와 2D 이미지의 feature map을 매칭하는 latent 매칭 기법을 제안하여 정확도 향상을 이루었다. 그러나 이 기법은 3D 재구축을 위한 딥러닝 네트워크 외에 VAE를 추가로 학습시켜야 하기 때문에 학습 속도가 느리고 학습 과정에서 메모리 소모가 크다는 단점이 존재하였다. 또한, latent vector의 값이 발산하지 않도록 하는 divergence loss에 2D 이미지의 x축 회전, y축 회전 등 촬영 정보가 학습에 추가적으로 필요하다는 단점도 존재

하였다. 이후 Bin 등[5]은 VAE를 사용하지 않고 2D 이미지의 feature map을 이용하여 직접 latent vector를 생성하여 더 향상된 정확도와 학습 속도, 그리고 학습에 필요한 메모리 양을 감소시켰다. 그러나, 여전히 촬영 정보가 추가적으로 필요하다는 단점이 있다.

따라서 본 논문에서는 효율적인 feature map 추출 네트워크를 이용한 2D 이미지에서의 3D 포인트 클라우드 재구축 기법을 제안한다. 본 논문에서 제안한 기법의 독창성은 다음과 같다. 첫 번째로, 메모리 측면에서 기존 기법보다 약 27% 효율적인 새로운 feature map 추출 네트워크를 사용한다. 두 번째로, 2D 이미지의 고해상도 feature를 보존하여 정확도를 기존 기법보다 약 6.3% 향상시킬 수 있다. 세 번째로, 촬영 정보를 필요로 하지 않는 divergence loss를 사용한다.

## II. 본론

### 1. 전체 개요도

본 논문에서 제안하는 효율적인 feature map 추출 네트워크를 이용한 2D 이미지의 3D 포인트 클라우드 재구축 기법의 전체 개요도는 그림 1과 같다.

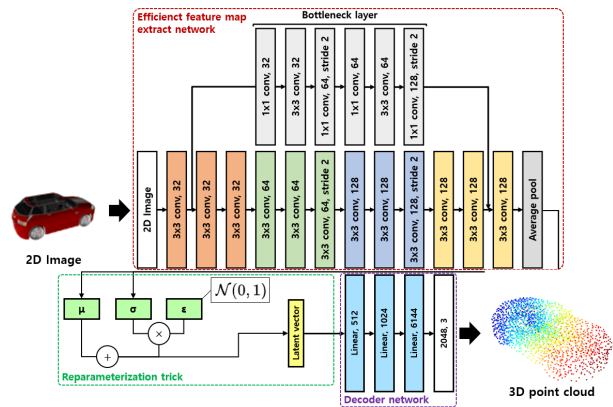


Fig. 1. A complete overview of the proposed method. 그림 1. 제안하는 기법의 전체 개요도

\* Dept. Electronic Engineering, Hanbat National University

★ Corresponding author

E-mail : shlee@cad.hanbat.ac.kr, Tel : +82-42-821-1137

※ Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT)(No. NRF-2022R1F1A1066371)

Manuscript received Sep. 01, 2022, revised Sep. 21, 2022, accepted Sep. 22, 2022.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

그림 2는 본 논문에서 제안하는 기법의 학습 과정을 나타낸다.

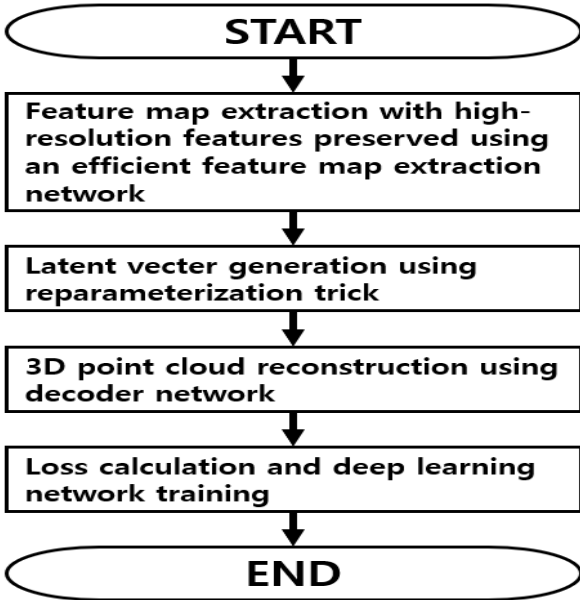


Fig. 2. The learning process of the proposed method in this paper.

그림 2. 본 논문에서 제안하는 기법의 학습 과정

2. 효율적인 feature map 추출 네트워크를 이용한 고해상도 feature가 보존된 feature map 추출

2.1 기존 ResNet(Residual Network)의 문제점

ResNet[6]은 2D 이미지에 포함된 사물을 분류하는 문제에서 기존의 심층 네트워크의 gradient vanishing으로 인한 정확도 감소를 해결한 네트워크이다. ResNet은 초기의 feature를 심층으로 효과적으로 전달하여 심층 네트워크에서 초기의 feature가 유실되는 문제를 해결하여 gradient vanishing을 방지하고 정확도를 획기적으로 향상시켰다. 이후에 발표된 대부분의 딥러닝 네트워크 구조는 대부분 ResNet의 초기 feature 전달 방법을 응용하였다. 한편, ResNet은 그 깊이로 인하여 메모리를 지나치게 사용하는 문제를 해결하기 위해 첫 층에 마스크 크기가 7×7이고 stride가 2인 convolutional 층을 적용하여 이미지 크기를 절반으로 축소한다. 또한, 바로 다음 층에 2×2 마스크 크기의 max pooling을 수행하여 마찬가지로 이미지의 크기를 절반으로 축소한다. 따라서 ResNet은 깊은 층을 만들고 유지하기 위해 초기에 많은 양의 feature가 유실된다는 문제가 있다. 그림 3은 ResNet 1~2층을 통과하는 입력 2D 이미지의 크기 축소과정의 문제점을 나타낸다.

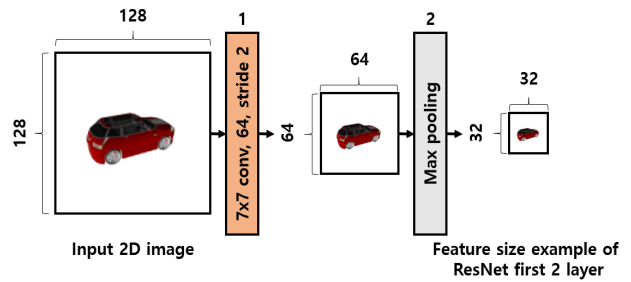


Fig. 3. Problems in the process of reducing the size of the input 2D image passing through ResNet 1st and 2nd layers.

그림 3. ResNet 1~2층을 통과하는 입력 2D 이미지의 크기 축소과정의 문제점

2.2 본 논문에서는 제안하는 효율적인 feature map 추출 네트워크

본 논문에서는 제안하는 네트워크를 통해 ResNet의 feature 유실 문제를 해결하고 메모리 측면에서도 효율적인 feature map 추출을 수행한다. 제안하는 네트워크는 딥러닝 네트워크의 중간까지 크기 축소를 수행하지 않아, 3D 포인트 클라우드 재구축에 필요한 중요한 정보가 유실되지 않았다. 또한, 이를 이용하여 정확도를 향상시킬 수 있다. 이미지 크기로 인해 발생하는 메모리 증가 문제는 채널의 개수를 줄이고 딥러닝 네트워크의 깊이를 얇게 구성하여 해결하였다. 그림 4는 입력 2D 이미지의 크기 축소가 없는 효율적인 feature map 추출 네트워크를 나타낸다.

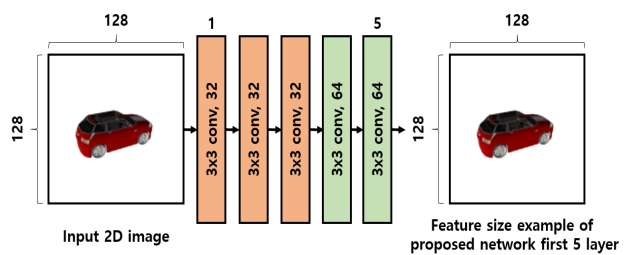


Fig. 4. Efficient feature map extraction network without reducing the size of the input 2D image.

그림 4. 입력 2D 이미지의 크기 축소가 없는 효율적인 feature map 추출 네트워크

한편, ResNet은 2D 이미지의 초기 feature를 심층으로 전달하는 방법으로 bottleneck 층을 여러 번에 걸쳐서 전달하는 방식을 사용하고 있다. 그러나 여러 층의 bottleneck 층도 gradient vanishing으로 인한 정확도 감소를 야기할 수 있다. 따라서 본 논문에서는 ResNet과 달리 여러 단계를 거치지 않고 feature map 네트워크

크의 마지막 층으로 초기 feature를 직접적으로 전달하여 초기 feature의 유실을 최소화하여 정확도를 더욱 향상시킬 수 있었다. 그림 5는 ResNet 34층과 제안하는 네트워크의 초기 feature 전달방식의 차이를 나타낸다.

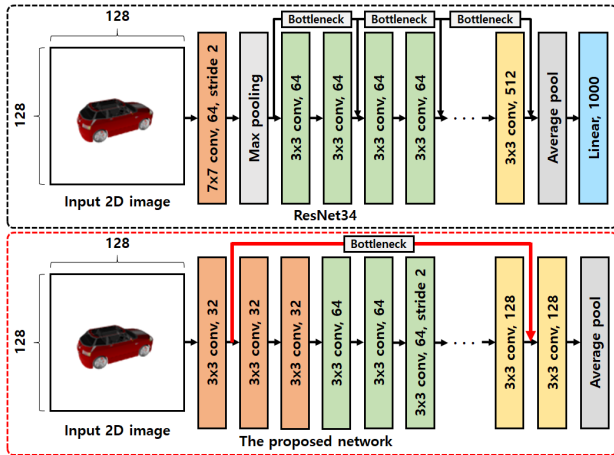


Fig. 5. Difference between ResNet 34 layer and the initial feature delivery method of the proposed network. 그림 5. ResNet 34층과 제안하는 네트워크의 초기 feature 전달방식의 차이

### 3. Reparameterization trick을 사용한 latent vector 생성

Reparameterization trick은 주로 VAE(Variational AutoEncoder)에서 사용되는 기법으로 입력 데이터로부터 추출한 feature map을 이용하여 정규 분포의 형태의 latent vector를 생성한다. 정규 분포의 형태로 제한된 데이터는 무작위성이 적어서 딥러닝 네트워크를 이용하여 어떠한 값을 추정하는 것이 편리하다. 본 논문에서는 2D 이미지로부터(1, 1000) 크기의 feature map을 추출하고 이를 linear 층의 입력 데이터로 이용하여(1, 100) 크기의 평균  $\mu$ 와 표준 편차  $\sigma$ 를 생성한다. 이후,  $\sigma$ 에 정규 분포를 따르는 (1, 100) 크기의 임의의 값인  $\epsilon$

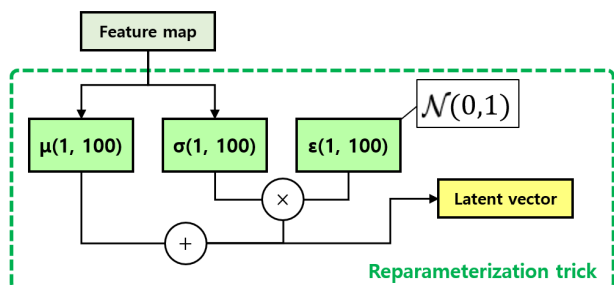


Fig. 6. Latent vector generation process using feature map.

그림 6. Feature map을 이용한 latent vector 생성과정

을 곱하여 표준 편차의 무작위성을 제한시킨다. 이후,  $\mu$ 와  $\sigma$ 를 더하여 latent vector를 생성한다. 그림 6은 feature map을 이용한 latent vector 생성과정을 나타낸다.

### 4. Decoder 네트워크를 이용한 3D 포인트 클라우드 재구축

Reparameterization trick을 통해 2D 이미지의 feature map으로부터 생성한 latent vector를, 3층의 linear layer로 이루어진 Decoder 네트워크를 사용하여 3D 포인트 클라우드로 재구축한다. LeakyReLU 활성화 함수를 여러 층에 걸쳐 적용하여 데이터를 비선형적으로 만들어 Decoder 네트워크의 추정 성능을 높이고 3D 포인트 클라우드의 재구축 정확도를 향상시킨다. (1, 100) 크기의 latent vector를 (1, 512), (1, 1024), (1, 6144), 최종적으로 (2048, 3) 크기의 2048개의 점에 대한 3차원 좌표 정보가 담긴 3D 포인트 클라우드의 형태로 재구축을 수행한다. 그림 7은 본 논문에서 사용된 Decoder 네트워크의 구조를 나타낸다.

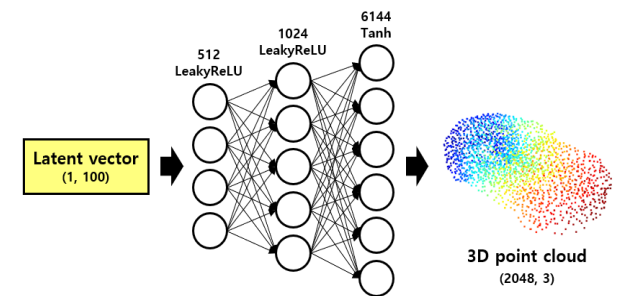


Fig. 7. Decoder network structure used in this paper. 그림 7. 본 논문에서 사용된 Decoder 네트워크의 구조

### 5. Loss 계산 및 딥러닝 네트워크 학습

본 논문에서는 여러 가지 loss 함수가 결합된 복합적인 loss 함수를 딥러닝 네트워크의 학습에 사용한다. 본 논문에서 제안한 Divergence loss는 latent vector 생성 시 분모가 0이 되어 값이 무한대로 발산하는 것을 방지한다. 또한, CD(Chamfer Distance) loss, EMD(Earth Mover's Distance) loss는 서로 다른 3D 좌표계 배열의 점들의 거리를 가까워지도록 만들어 재구축된 3D 포인트 클라우드의 정확도를 향상시킨다. 본 논문의 딥러닝 네트워크 학습에 사용된 loss 함수는 식 1과 같으며  $L_{DIV}$ 는 divergence loss를 나타낸다.

$$Loss = L_{DIV} + CD_{loss} + EMD_{loss} \quad (1)$$

### 5.1 Divergence loss

본 논문에서는 divergence loss로 KL(Kullback-Leibler) divergence를 사용하였다. KL divergence는  $\mu$ 와  $\sigma$ 를 제한하여 이를 이용하여 생성된 latent vector가 표준 편차의 형태를 따르도록 한다. 한편, Mandikal 등과 Bin 등이 사용한 식 2와 같은 divergence loss 2D 이미지가 촬영된 카메라의 회전각 정보를 필요로 하고, 이를 이용하여 촬영 각도가 유사한 2D 이미지들의 feature map들로부터 생성된 latent vector들이 서로 유사한 형태를 따르도록 한다.  $\eta$ 는 divergence loss의 학습 반영 정도를 결정하는 가중치를 나타내며,  $\phi_i$ 는 2D 이미지의 촬영 각도를,  $\delta$ 는 각도의 차이의 반영 정도를 결정하는 페널티 각도를 나타낸다.

$$(\sigma - \eta e^{-\frac{(\phi_i - \pi)^2}{\delta^2}})^2 \quad (2)$$

그러나, 2D 이미지뿐만 아니라 촬영 각도가 학습에 필요하다라는 사항은 그만큼 데이터셋이 자세한 정보를 담고 있어야 하며 데이터셋의 구축을 어렵게 만드는 단점이다. 따라서, 본 논문에서는 촬영 각도를 사용하지 않고 latent vector를 이루는  $\mu$ 와  $\sigma$ 만을 이용하여 divergence loss를 계산한다. Latent vector의 무작위성을 제한하는 기존의 논문[7]과 다르게, 무작위성을 통해 정보의 다양성을 늘려 3D 포인트 클라우드의 재구축 정확도가 높아질 수 있도록 하였다. 식 3은 본 논문의 딥러닝 네트워크 학습에 사용한 divergence loss를 나타낸다.

$$L_{DIV} = -0.05(1 + \sigma - \mu^2 - e^\sigma) \quad (3)$$

### 5.2 CD loss

CD loss는 서로 다른 두 3D 좌표 배열 사이에 가장 가까운 점을 구하는 함수이다. CD loss는 가장 가까운 거리에 있는 점을 양 방향 계산으로 찾는다. 그러나 가장 가까운 점을 공유하는 점들이 있는 경우 3개 이상의 점이 하나로 뭉치게 되고 사물의 윤곽을 구성하는 점 하나가 빠진 상태가 되어 손실이 일어난다. 따라서 이런 문제를 해결하기 위해 딥러닝 기반의 3D 포인트 클라우드 재구축에는 전체 형태를 손실 없이 유지할 수 있는 EMD loss와 함께 사용된다. 식 4는 CD의 계산식을 나타내며 그림 8은 CD의 양 방향 계산 예시를 나타낸다.

$$CD(S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2 + \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_2 \quad (4)$$

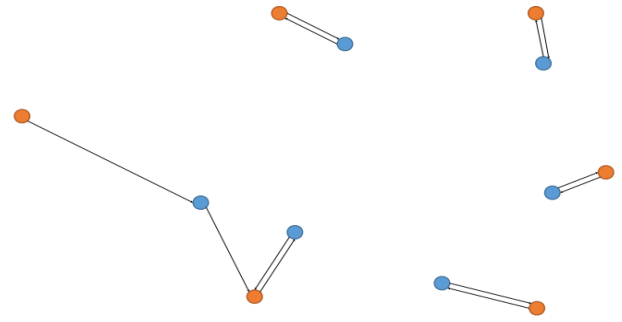


Fig. 8. Example of CD calculation.

그림 8. CD의 양 방향 계산 예시

### 5.3 EMD loss

EMD loss는 서로 다른 두 3D 좌표 배열 사이에 가장 가까운 점을 구하는 함수라는 점이 CD loss와 유사하다. 그러나 EMD loss는 CD loss와 다르게 단 방향 계산을 통해서 거리를 계산한다. 따라서 EMD loss는 2개보다 많은 점이 뭉치게 될 가능성이 현저히 적다. 따라서 EMD loss를 학습에 이용하면 사물 윤곽의 전체 형태가 손실 없이 잘 보존된다는 장점이 있다. 식 5는 EMD loss의 계산식을 나타내며 그림 9는 EMD loss의 단 방향 계산 예시를 나타낸다.

$$EMD(S_1, S_2) = \sum_{\phi: S_1 \rightarrow S_2} \min_{x \in S_1} \|x - \phi(x)\|_2 \quad (5)$$

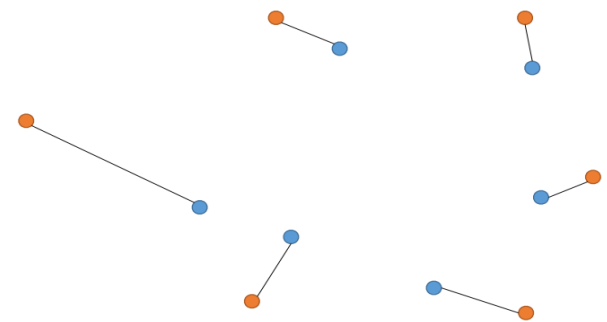


Fig. 9. Example of EMD calculation.

그림 9. EMD의 단 방향 계산 예시

## 6. 성능 실험

### 가. 실험 환경

실험에 사용된 OS 및 hardware는 Windows 10 64bit OS를 기반으로 Intel i7-10700k 3.8GHz CPU, 16GB RAM, NVIDIA GeForce RTX 2770(VRAM 8GB) GPU로 구성되어있다. 딥러닝 라이브러리는 Pytorch 1.9.1, CUDA 11.1, cuDNN 8.2.1 버전을 사용하였다. 본 논문에서는 비교 논문들과 동일한 실험 환경을 구성하기 위해 Princeton, Stanford 등의 연구원들이 구축한 사물

데이터셋인 ShapeNet 데이터셋[8]의 13가지 카테고리를 실험에 사용하였다. 표 1은 실험에 사용된 ShapeNet 데이터셋의 카테고리들을, 그림 10은 실험에 사용한 ShapeNet 데이터셋의 예시를 나타낸다.

Table 1. 13 categories of ShapeNet dataset used in the experiment.

표 1. 실험에 사용된 ShapeNet 데이터셋의 13가지 카테고리

Category	Contents
Object	bench, chair, lamp, speaker, gun, table, cabinet, monitor, long chair, cell phone
Vehicle	ship, airplane, car



Fig. 10. Example of the ShapeNet dataset used in the experiment.

그림 10. 실험에 사용한 ShapeNet 데이터셋의 예시

나. 실험 결과

성능평가 방법으로는 CD, EMD 및 네트워크의 연산에 필요한 메모리를 나타내는 Floating point Operations Per Second(FLOPs)를 비교하였다. 제안하는 기법의 성능을 객관적으로 평가하기 위해 비교 논문들과 같은 조건에서 13개의 카테고리에 대한 실험 결과의 평균을 산출하였다. 표 2는 본 논문에서 제안하는 기법과 다른 논문의 기법들의 비교 결과를, 그림 11은 ShapeNet 데이터셋의 2D 이미지에서의 3D 인트 클라우드 재구축 결과를 나타낸다. 표 2의 결과에서 본 논문에서 제안하는 기법의 CD 값이 5.87, EMD 값이 5.81 FLOPs 값이 2.9G로 산출되었다. 한편, CD, EMD 수치가 낮을수록, 재구축한 3D 포인트 클라우드가 원본에 근접하는 정확도가 향상된 결과를 나타낸다. 또한, FLOPs 수치가 낮을수록 딥러닝 네트워크에 필요한 메모리가 적게 소요되는 결과를 나타낸다. 따라서, 제안하는 기법의 CD, EMD, FLOPs 성능평가 결과가 다른 논문에서 발표한 기법들보다 우수한 결과를 나타내어 객관적인 성능이 입증되었다.

Table 2. Comparison result of the ShapeNet dataset reconstruction accuracy of the proposed method and the methods of other papers.

표 2. 제안하는 기법과 다른 논문의 기법들의 ShapeNet 데이터셋 재구축 정확도 비교 결과

Method	CD ↓	EMD ↓	FLOPs ↓
PSGN[2]	9.11	12.29	-
3D-lmnet[3]	8.91	9.70	4G
3D-reconstnet[5]	6.26	6.20	4G
The proposed method	<b>5.87</b>	<b>5.81</b>	<b>2.9G</b>

기존의 방식에서 초기의 이미지 축소를 통한 메모리 최적화 방법이 중요한 feature가 유실될 것이라 판단하였다. 따라서 본 논문에서 제안하는 기법은 해상도를 유지하고 채널 수를 줄이는 방법을 사용하여 중요한 feature를 보존하였다. 또한, 중요한 feature가 보존됨으로써 효율적인 네트워크 구조로도 정확도를 향상시킬 수 있게 되어 기존의 방법보다 낮은 FLOPs를 달성하였다. Reparameterization trick으로 생성된 latent vector는 KL divergence loss를 통해 제한하게 되면 지나치게 틀에 박힌 형태가 되는 것을 확인하였다. 또한, KL divergence의 반영 비율을 낮게 하였고 latent vector에 보다 다양한 정보가 담길 수 있게 되었다. 따라서, 이를 이용한 3D 포인트 클라우드 재구축 정확도가 향상된 결과를 나타내었다. 그림 11은 ShapeNet 데이터셋의 2D 이미지에서의 3D 포인트 클라우드 재구축 결과를 나타낸다.

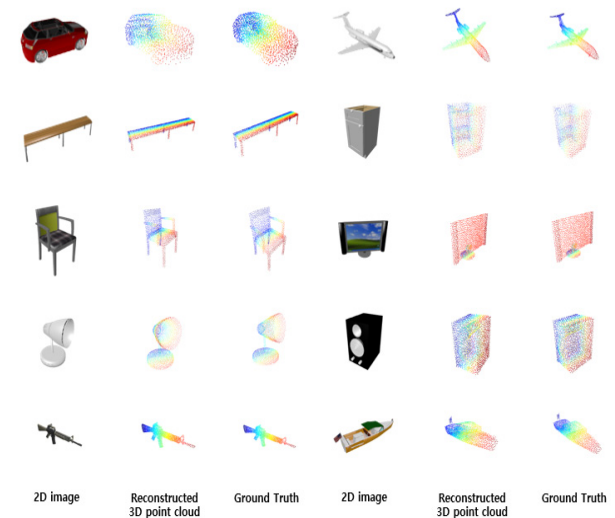


Fig. 11. 3D point cloud reconstruction results from a 2D image of the ShapeNet dataset.

그림 11. ShapeNet 데이터셋의 2D 이미지에서의 3D 포인트 클라우드 재구축 결과

### III. 결론

본 논문에서는 효율적인 feature map 추출 네트워크를 이용한 2D 이미지에에서의 3D 포인트 클라우드 재구성 기법을 제안하였다. 본 논문에서 제안한 기법의 독창성은 첫 번째, 2D 이미지의 고해상도 feature를 보존하여 정확도를 기존의 기법보다 약 6.3% 향상시킬 수 있었다. 두 번째, 메모리 측면에서 기존 기법보다 약 27% 효율적인 새로운 feature map 추출 네트워크를 사용하였다. 세 번째, 촬영 정보를 필요로 하지 않는 divergence loss를 사용하였다. 제안하는 기법의 성능을 객관적으로 평가하기 위해 ShapeNet 데이터셋을 이용하여 비교 논문들과 같은 방법으로 실험한 결과, 본 논문에서 제안하는 기법의 CD 값이 5.87, EMD 값이 5.81 FLOPs 값이 2.9G로 산출되었다. CD, EMD 수치가 낮을수록, 재구성한 3D 포인트 클라우드가 원본에 근접하는 정확도가 향상된 결과를 나타낸다. 또한, FLOPs 수치가 낮을수록 딥러닝 네트워크에 필요한 메모리가 적게 소요되는 결과를 나타낸다. 따라서, 제안하는 기법의 CD, EMD, FLOPs 성능평가 결과가 다른 논문의 기법들보다 메모리 측면에서 약 27%, 정확도 측면에서 약 6.3% 향상된 결과를 나타내어 객관적인 성능이 입증되었다.

향후 연구로는 범용적인 KL divergence 보다 latent vector에 3D 포인트 클라우드 재구축에 필요한 더 많은 정보를 담을 수 있는 새로운 divergence loss 도출에 관한 연구가 필요하다.

### References

- [1] Choy, C. B., Xu, D., Gwak, J., Chen, K., & Savarese, S.. "3d-r2n2: A unified approach for single and multi-view 3d object reconstruction," *In European conference on computer vision*, pp. 628-644. 2016. DOI: 10.48550/arXiv.1604.00449
- [2] Fan, Haoqiang, Hao Su, and Leonidas J. Guibas. "A point set generation network for 3d object reconstruction from a single image," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016. DOI: 10.48550/arXiv.1612.00603
- [3] Mandikal, P., Navaneet, K. L., Agarwal, M., Babu, R. V. "3D-LMNet: Latent embedding matching

- for accurate and diverse 3D point cloud reconstruction from a single image," *Proceedings of the British Machine Vision Conference(BMVC)*, 2018. DOI:10.48550/arXiv.1807.07796
- [4] KINGMA, Diederik P.; WELLING, Max. "Auto-encoding variational bayes," *International Conference on Learning Representations (ICLR)*, 2014. DOI: 10.48550/arXiv.1312.6114
- [5] Li, B., Zhang, Y., Zhao, B., & Shao, H. "3D-ReConstnet: a single-view 3d-object point cloud reconstruction network," *IEEE Access* 8 pp.83782-83790. 2020. DOI: 10.1109/ACCESS.2020.2992554
- [6] He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep residual learning for image recognition," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.770-778. DOI: 10.48550/arXiv.1512.03385
- [7] Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Lerchner, A. "beta-vae: Learning basic visual concepts with a constrained variational framework." 2016.
- [8] Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., ... & Yu, F.. "Shapenet: An information-rich 3d model repository," 2015. DOI: 10.48550/arXiv.1512.03012

### BIOGRAPHY

#### Jeong-Yoon Kim (Member)



2020 : BS degree in Electronic Engineering, Hanbat National University

2022 : MS degree in Electronic Engineering, Hanbat National University

2022~current : Ph. D degree course of Electronic Engineering Hanbat National University

**Seung-Ho Lee** (Member)



1986 : BS degree in Electronic Engineering, Hanyang University  
1989 : MS degree in Electronic Engineering, Hanyang University  
1994 : Ph. D degree in Electronic Engineering, Hanyang University

1994~current : Professor, Department of Electronic Engineering, Hanbat National University