

# 3D 콘텐츠 생성 드림가우시안의 한계와 개선방안에 관한 연구

이기성, 임채진, 현준희, 박선우, 황영희, \* 한동일

세종대학교 컴퓨터공학과

[leekisung0212@gmail.com](mailto:leekisung0212@gmail.com), [papou6230@naver.com](mailto:papou6230@naver.com), [hyeonjunhee@gmail.com](mailto:hyeonjunhee@gmail.com),  
[seonu123@naver.com](mailto:seonu123@naver.com), [nsun14363@gmail.com](mailto:nsun14363@gmail.com), \* [dihan@sejong.ac.kr](mailto:dihan@sejong.ac.kr)

## A Study on the Limitation and Improvements of 3D Content Creation DreamGaussian

Kiseong Lee, Chaejin Lim, Junhee Hyeon, Seonu Park, Yeonghui Hwang, \* Dongil Han

Department of Computer Engineering Sejong University

### 요약

최근 몇 년 동안, 3D 디지털 콘텐츠의 중요성이 높아지면서 이미지-투-3D 변환 기술은 수작업의 번거로움을 줄이고 사용자 참여를 촉진하는 중요한 역할을 하고 있다. 본 연구는 최적화 기반 파이프라인을 개선하고, 명시적인 메쉬 및 텍스처 맵을 활용하여 3D 자산 생성의 효율성을 향상시킨 논문인 DreamGaussian을 참조하고 있다. 그러나 이 적용 결과로 발생한 흐릿함 문제를 확인하고, 새로운 해결 방법을 모색 중에 있다. 해당 연구는 Image-to-3D 분야에서 해당 프레임워크의 성능 향상을 목표로 하며, 논문의 문제 부분을 파악하여 극복하는 방안을 제시한다.

### I. 서론

최근 몇 년 동안, 디지털 콘텐츠 생성 분야에서의 발전은 다양한 응용 프로그램에서 3D 디지털 콘텐츠의 중요성을 높이고 있다. 이미지-투-3D의 기술은 수작업의 번거로움을 줄이고 비전문 사용자의 참여를 촉진하여 혁명적인 변화를 가져오고 있다. [1] 최근 2D 콘텐츠 생성의 발전에서 영감을 받은 3D 콘텐츠 생성 분야는 추론 전용 3D 네이티브 방법과 최적화 기반 2D 리프팅 방법으로 나뉜다. 본 연구에서는 최적화 기반 파이프라인을 개선하여 3D 콘텐츠 생성의 효율성을 크게 향상시킨 DreamGaussian 프레임워크를 소개한다. [2]

DreamGaussian은 명시적인 메쉬 및 텍스처 맵을 사용하여 단일 뷰 이미지에서 현실적인 3D 자산을 몇 분 내에 생성할 수 있는 기술적인 혁신을 제공한다. 핵심 설계는 3D Gaussian Splatting [4]을 생성 설정에 적용하여 효율성을 극대화한다. 이를 통해 NeRF(Neural Radiance Field) [3] 표현의 단순화와 함께 빈 공간의 효과적인 가지치기를 달성했다. 연구 결과로 이미지-투-3D 파이프라인은 몇 초 내에 빠르게 거친 모양을 생성하며, 단일 GPU에서 효율적으로 수렴하는 것을 확인할 수 있다.

그러나, 우리는 DreamGaussian을 직접 적용한 결과, 물체의 뒷부분에서 발생한 흐릿함 문제를 확인했다. 저자들은 이 문제의 해결방법으로 Stage 2 알고리즘의 Iteration 증가를 제안했지만, 실제로는 효과가 없었다. 특히, 생성된 물체의 알려진 부분은 고품질이었지만, 알려지지 않은 부분의 품질은 그만큼 좋지 않은 것을 발견했다. 우리는 이 문제를 해결하려는 새로운 방법을 모색하고 있으며, 어려운 텍스처, 예를 들면 반짝거림이나 굴곡이 많은, 처리에서 DreamGaussian의 성능이 떨어지고 있음을 확인했다. 이에 대한 해결책을 개발하기 위해 추가 연구가 필요한 시점이다.

본 연구는 Image-to-3D 분야에서의 DreamGaussian의 성능 향상을 목

표로 하며, 현재 알고리즘의 문제 부분을 파악하여 명시하고 이를 극복하는 방안을 제시한다. 논문의 후반부에서는 이러한 문제에 대한 자세한 내용을 다루며, 향후 연구 방향과 개선 전략을 논의할 것이다.

### II. DreamGaussian의 한계점

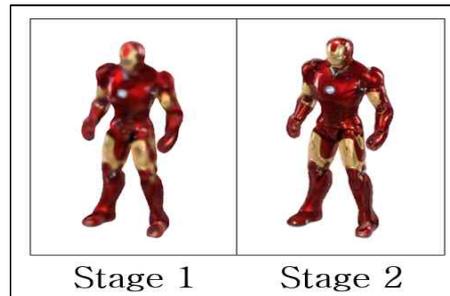


그림 1 DreamGaussian  
각 단계에서의 렌더링 결과

DreamGaussian 모델은 먼저 Stage 1에서 SDS를 활용하여 3D Gaussian Splatting을 초기화하고, 이어서 Stage 2에서는 3D 가우시안에서 텍스처 메쉬를 효과적으로 추출하며, 래스터화에 사용되는 밍맵 텍스처 샘플링 기술을 통해 텍스처를 미세 조정합니다(그림 1). 실험에서는 더 어려운 이미지 데이터를 사용하여 Stage 2 결과물을 개선하기 위해 여러 파라미터를 조절했지만 큰 효과를 얻지 못했습니다(그림 2). Stage 2에서는 두 가지 Loss를 계산하는데, 첫 번째는 Stage 1에서 생성된 Gaussian Splatting과 알려진 시점의 입력 이미지 간의 계산이며, 두 번째는 2D Diffusion Model에서 조정된 무작위 시점의 이미지와 최적화된 Gaussian Splatting 간의 계산입니다.

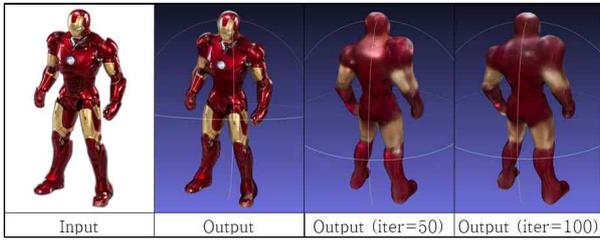


그림 2 미세 조정 횟수에 따른 DreamGaussian 최종 렌더링 결과

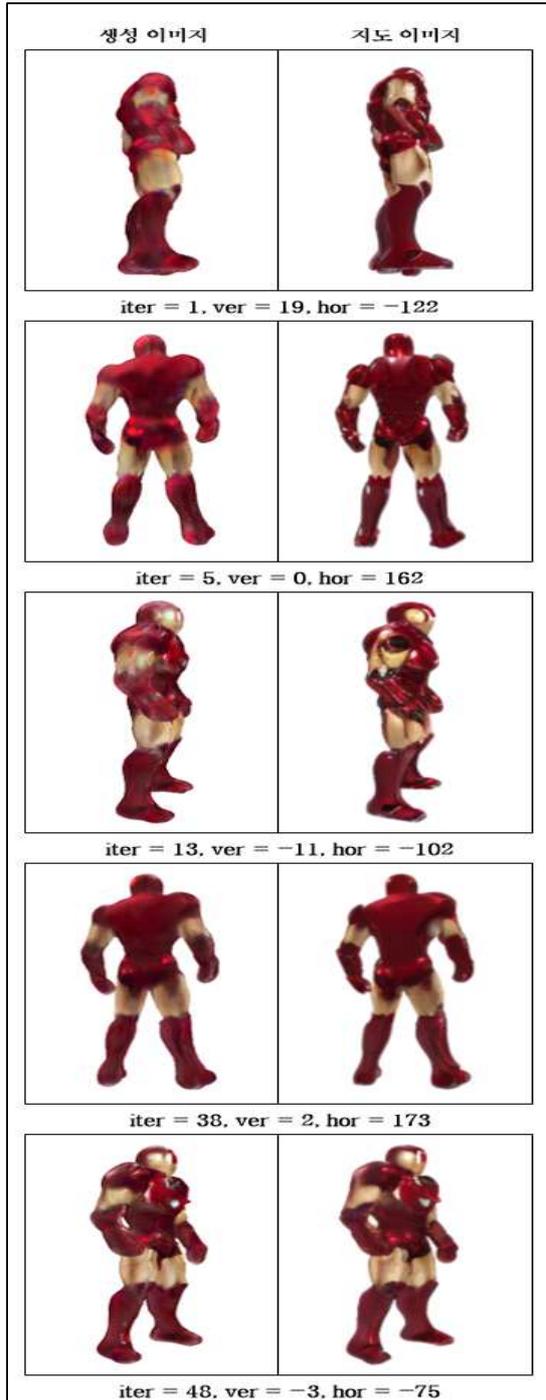


그림 3 반복 횟수와 시점에 따른 DreamGaussian의 Stage 2의 최적화 과정 시각화

그러나 이 최적화 과정에서는 2D Diffusion Model이 지속적으로 원본 이미지의 텍스처 성질을 상실하는 문제가 발견되었습니다. PSNR 및 SSIM 값은 일정하지만 시각적인 특성 중에서 주로 텍스처가 변화합니다. 훈련 후반부로 갈수록 이미지의 텍스처가 부족하거나 변형되어 보이며, 훈련 막바지에는 생성 이미지가 지도 이미지보다 품질이 우수해지는 상황이 발생합니다(그림 3). 올바른 방향으로 최적화하려면 계속해서 원본 이미지의 텍스처가 유지된 고품질 이미지를 제공하는 것이 중요하며, 이는 모델이 원하는 결과를 효과적으로 학습하고 품질을 향상시킬 수 있을 것으로 기대됩니다.

### III. 결론

DreamGaussian 모델은 초기 단계에서 3D Gaussian Splatting을 생성하고, Stage 2에서는 텍스처 매쉬 추출 및 맵 텍스처 샘플링을 통해 최종 렌더링을 수행한다. 실험 결과, 두 번째 Loss 계산에서 2D Diffusion Model이 원본 이미지의 텍스처를 상실하고 있음을 확인했다. 훈련 알고리즘을 올바른 방향으로 최적화하기 위해서는 계속해서 원본 이미지의 텍스처를 유지한 고품질 이미지를 사용해야 한다. 이러한 지속적인 고품질 이미지 제공이 품질 개선에 도움이 될 것으로 기대되며, 추가적인 실험과 조치가 필요하다.

### ACKNOWLEDGMENT

This research was supported in part by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2024-RS-2022-00156354) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation), in part by the Korea Institute for Advancement of Technology (KIAT) grant funded by the Ministry of Trade, Industry and Energy (MOTIE) (P0012725).

### 참고 문헌

- [1] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. Highresolution image synthesis with latent diffusion models. In CVPR, pp. 10684 - 10695, 2022.
- [2] Jiayang Tang, Jiawei Ren, Hang Zhou, Ziwei Liu and Gang Zeng. DreamGaussian: Generative Gaussian Splatting for Efficient 3D Content Creation. arXiv, 2023.
- [3] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In ECCV, 2020.
- [4] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. ToG, 42(4):1 - 14, 2023.
- [5] Samuli Laine, Janne Hellsten, Tero Karras, Yeongho Seol, Jaakko Lehtinen, and Timo Aila. Modular primitives for high-performance differentiable rendering. ToG, 39(6), 2020.
- [6] Ruoshi Liu, Rundi Wu, Basile Van Hoorick, Pavel Tokmakov, Sergey Zakharov, and Carl Vondrick. Zero-1-to-3: Zero-shot one image to 3d object. arXiv preprint arXiv:2303.11328, 2023b.