ISSN 2508-2817 KiDRS Korea Institute of Design Research Society 통권 39호, 2026. Vol.11, No.2

Unreal Engine 5와 Generative AI를 결합한 실시간 에셋 최적화 파이프라인 연구

Research on a Real-time Asset Optimization Pipeline Combining Unreal Engine 5 and Generative AI

주 저 자 : 오수진 (Oh, Soo Jin) 우송대학교 게임멀티미디어학부 게임소프트웨어전공 겸임교수
azxr.lab@gmail.com

https://doi.org/10.46248/kidrs.2026.2.120

접수일 2026. 04. 06. / 심사완료일 2026. 04. 09. / 게재확정일 2026. 04. 13. / 게재일 2026. 06. 30.

Abstract

As the demand for high-quality 3D assets increases, generative AI has emerged as an alternative, but it faces technical limitations in being directly applied to commercial engines due to the 'polygon soup' phenomenon. To address this gap, this study proposes a pipeline that optimizes generative AI outputs in real-time to meet Unreal Engine 5 (UE5) specifications. To achieve this, a five-stage integrated process was designed, covering intelligent generation, pre-processing, auto-retopology, texture atlasing, and engine live-streaming. Furthermore, it integrates 3D Gaussian Splatting (3DGS) and YOLOv8 technologies to automate asset placement and maximize memory efficiency. In addition, a 'Human-in-the-loop' model, where human artists intervene, was introduced to compensate for AI errors and secure art directability. Performance analysis showed that applying this pipeline significantly reduced draw calls and polygon counts, improving perceived performance in gaming and VR environments by more than twofold. A real-world case study of a small-to-medium studio proved a groundbreaking productivity improvement, reducing the production time per asset from 14 hours to 1.5 hours. In conclusion, this research confirmed the effectiveness of the optimization process combining UE5 and AI, successfully presenting automation standards and collaborative directions for the mass production of high-quality 3D assets in the future.

Keyword

Unreal Engine 5(언리얼엔진5), Genarative AI(생성형 AI), Asset Optimization(에셋 최적화), Human-in-the-loop(HITL, AI-인간 협업 모델)

요약

고품질 3D 에셋 수요 증가로 생성형 AI가 대안으로 부상했으나, '폴리곤 수프(Polygon Soup)' 현상으로 인해 상용 엔진에 바로 적용하기 어려운 기술적 한계가 있다. 본 연구는 이러한 공백을 해소하고자 생성형 AI 출력물을 언리얼 엔진 5(UE5) 규격에 맞춰 실시간 최적화하는 파이프라인을 제안하였다. 이를 위해 지능형 생성, 전처리, 자동 리토폴로지, 텍스처 정렬, 엔진 라이브 스트리밍에 이르는 5단계 통합 프로세스를 설계하였다. 더 나아가 3D 가우시안 스플래팅과 YOLOv8 기술을 융합하여 에셋의 자동 배치 및 메모리 효율 극대화를 도모하였다. 또한, AI의 오류를 보완하고 예술적 제어권을 확보하기 위해 아티스트가 개입하는 'AI-인간 협업(Human-in-the-loop)' 모델을 도입하였다. 성능 분석 결과, 파이프라인 적용 시 드로우 콜과 폴리곤 수가 대폭 감소하여 게임 및 VR 환경의 체감 성능이 2배 이상 향상되었다. 실제 중소 스튜디오 사례에서는 에셋 당 제작 시간을 14시간에서 1.5시간으로 단축시켜 획기적인 생산성 향상 효과를 입증하였다. 결론적으로 본 연구는 UE5와 AI를 결합한 최적화 프로세스의 실효성을 확인하고, 향후 고품질 3D 에셋 대량 생산을 위한 자동화 및 협업 방향성을 성공적으로 제시하였다.

목차

1. 서론

1-1. 연구 배경 및 목적

디지털 콘텐츠의 복잡성과 정밀도가 기하급수적으로 증가함에 따라, 고품질 3D 에셋의 대량 생산은 게임 산업 및 가상 프로덕션 분야의 핵심 과제로 부상하였다. 전통적인 3D 모델링 워크플로우는 전문가에 의한 수작업 폴리곤 모델링, UV 전개, 텍스처 페인팅 과정을 거치며, 이는 에셋 하나당 수 시간에서 수십 시간이 소요되는 고비용 구조를 지닌다.[1] 이러한 환경에서 생성형 AI의 등장은 에셋 제작의 민주화와 효율화를 예고하고 있으나,[2] 실제 상용 엔진에서의 가동을 위한 최적화 기술은 여전히 공백 상태에 머물러 있다.[3] 본 연구의 목적은 이러한 기술적 공백을 해소하기 위해 생성형 AI의 출력물을 엔진 규격에 맞게 실시간으로 정제하고 최적화하는 파이프라인을 제안하고 그 학술적·실용적 가치를 검증하는 데 있다.

1-1. 연구 방법 및 범위

Unreal Engine 5는 Nanite와 Lumen이라는 혁신적인 기술을 통해 고해상도 지오메트리를 효율적으로 처리할 수 있는 기반을 마련하였다. 그러나 생성형 AI가 산출하는 이른바 '폴리곤 수프(Polygon Soup)' 형태의 무질서한 데이터는 Nanite의 압축 알고리즘에도 불구하고 GPU 메모리와 대역폭에 과도한 부하를 유발하는 한계를 지닌다. 따라서 본 연구는 생성형 AI의 결과물을 엔진 규격에 맞게 실시간으로 정제하고 최적화하는 공학적 방법론으로 연구 범위를 한정하며, 특히 UE5의 나나이트 가상화 시스템 및 루멘 글로벌 일루미네이션 환경에서의 성능 최적화 파이프라인 구축을 중점적으로 다룬다.

1) Aryan N. Modi & Kuntal Ghose, Artificial Intelligence in game asset pipeline, JETIR, Volume 13, Issue 2, 2026, (a202) p.1

2) Shyngys Adilkhan, Madina Alimanova, Lei Shi, Aiganym Soltiyeva, Advances in artificial intelligence-driven 3D model generation: a review of GAN and VAE methodologies, Bulletin of Electrical Engineering and Informatics, Vol. 14, No. 6, 2025, pp.6-7

3) 박준형, 최용석, 상용 게임 엔진 기반 실시간 렌더링 최적화를 위한 플러그인 설계 및 구현 연구, 한국게임학회, 한국게임학회 논문지, 제 23권, 제 4호, 2023, p.46

2. 이론적 배경

본 연구에서는 UE5의 핵심 기술(Nanite, Lumen)이 요구하는 에셋의 기술적 표준을 정의하고, AI 생성 에셋의 한계를 이론적으로 규명하고자 한다.

2-1. Unreal Engine 5의 기술적 혁신과 에셋 최적화의 상관관계

2-1-1. Nanite 가상화 지오메트리의 메커니즘과 한계

Unreal Engine 5의 핵심 렌더링 기술인 Nanite는 기존의 LOD(Level of Detail) 제작 방식에 근본적인 변화를 가져왔다. Nanite는 수백만 개의 폴리곤을 가진 고해상도 메쉬를 1픽셀 단위의 마이크로폴리곤으로 가상화하여 실시간으로 렌더링한다. 이는 전통적인 렌더링 파이프라인에서 드로우 콜을 줄이고 GPU 성능을 극대화하는 데 기여한다. 하지만 Nanite 시스템이 모든 기하학적 문제를 해결하는 만능 도구는 아니다. 생성형 AI가 생성한 메쉬는 종종 비정상적인 면(Non-manifold geometry), 겹친 정점, 혹은 구멍(Holes)을 포함하는데, 이는 Nanite의 가상화 인코딩 과정에서 아티팩트를 발생시키거나 렌더링 오류를 유발할 수 있다.

또한 Nanite는 정적 메쉬(Static Mesh)에 최적화되어 있으며, 변형이 필요한 스켈레탈 메쉬(Skeletal Mesh)나 투명한 재질의 에셋에 대해서는 여전히 전통적인 최적화 기법이 병행되어야 한다. 따라서 AI 기반 에셋 파이프라인은 생성된 모델이 Nanite-Ready 상태인지를 검증하고, 필요한 경우 메쉬 클리닝 및 리토폴로지 과정을 거쳐야만 실제 프로덕션에서 활용될 수 있는 충실도를 보장받는다.

2-1-2. Lumen과 실시간 조명 최적화 요구사항

Lumen은 실시간으로 빛의 전역 조명을 계산하는 시스템으로, 레이 트레이싱과 서피스 캐시 기술을 결합하여 복잡한 조명 효과를 구현한다. Lumen이 정확하게 작동하기 위해서는 에셋의 기하학적 정확도와 텍스처 맵핑의 일관성이 보장되어야 한다. 생성형 AI 에셋에서 흔히 발견되는 텍스처 왜곡이나 잘못된 노멀 맵(Normal Map) 정보는 Lumen의 간접 조명 계산 시 빛의 누수(Light Leaking) 현상을 초래할 수 있다.

실시간 성능을 보장하기 위해 에셋은 적절한 거리 필드(Distance Fields)와 메쉬 카드(Mesh Cards) 데이터를 포함해야 하며, 이는 AI 생성 단계에서부터 엔진 내 최적화 규칙을 반영해야 함을 시사한다. 특히 AI가 생성한 PBR(Physically Based Rendering) 재질이 실제 물리 법칙에 기반한 파라미터(알베도, 거칠기, 금속성 등)를 충실히 따르고 있는지에 대한 검증 과정이 필수적이다.

2-2. 생성형 AI 기반 3D 콘텐츠 생성 기술의 분석

2-2-1. 3D-AIGC 생성 모델의 기술적 분류

AIGC(Artificial Intelligence Generated Content)는 인공지능 생성 콘텐츠를 일컫는 용어이다. 3D 가상 에셋을 생성하는 AI 기술은 입력 방식과 알고리즘적 구조에 따라 다음과 같이 분류될 수 있다. 텍스트를 기반으로 3D 형태를 추론하는 Text-to-3D 모델은 자연어 처리를 통해 직관적인 워크플로우를 제공하며, 이미지에서 3D 구조를 복원하는 Image-to-3D 모델은 기존 컨셉 아트를 3D로 전환하는 데 탁월한 성능을 보인다.

[표 1] 3D-AIGC 기술

기술 구분	주요 알고리즘	특징	주요 도구
Text-to-3D	Diffusion Model, NeRF	텍스트 프롬프트 기반 생성, 창의적 형태 도출 유리	Meshy, GET3D
Image-to-3D	Transformer, GAN	2D 참조 이미지를 통한 정교한 재구성, 컨셉 유지력 높음	Tripo AI, Rodin AI
크기의 변화	Diffusion + VAE	메쉬 위상학적 안정성과 시각적 세밀함의 균형	Hunyuan 3D, 3D AI Studio

이러한 모델들은 최근 2020년대 중반에 접어들며 비약적인 발전을 이루었으며, 특히 확산 모델(Diffusion Model)은 노이즈 상태에서 데이터를 복원하는 과정을 통해 과거 GAN(Generative Adversarial Network)이나 VAE(Variational Autoencoder)보다 정교한 표면 질감과 형태를 생성할 수 있게 되었다.

2-2-2. 생성된 에셋의 위상학적 문제점과 최적화 난제

생성형 AI의 가장 큰 기술적 병목 현상은 '폴리곤 수프(Polygon Soup)' 문제이다.[4] AI 모델은 시각적인

4) David A. Smith, 'The State of AI in AAA Game Production', Medium, 2026.2.14.

실루엣을 맞추는 데는 능숙하지만, 게임 엔진이 요구하는 깨끗한 위상학적 구조(Clean Topology)를 만드는 데는 여전히 한계를 보인다. 이는 단순히 폴리곤 수가 많다는 문제뿐만 아니라, UV 맵이 수백 개의 조각으로 나뉘어 드로우 콜을 급격히 증가시키거나(Semantic Cutting 부족), 애니메이션을 위한 정점 웨이트(Weighting) 설정이 불가능한 구조를 갖는 문제를 포함한다.[5)]

실제 산업 현장에서는 AI 생성 에셋을 그대로 사용하는 대신, 생성된 모델을 베이스 메쉬로 활용하여 수동으로 리토폴로지 작업을 수행하는 경우가 많다. 연구에 따르면, 전문 디자이너의 약 70%가 AI 생성 에셋을 프로덕션 수준으로 수정하는 데 3시간 이상의 추가 시간을 소모하고 있으며,[6)] 이는 초기 생성 시간이 빠르더라도 전체 공정에서의 효율성을 저해하는 요소로 작용한다.

3. 실시간 에셋 최적화 파이프라인의 설계

3-1. 통합 파이프라인 구성 및 단계별 프로세스

효율적인 실시간 에셋 최적화를 위해서는 AI 생성 단계와 엔진 통합 단계 사이에 중간 정제(Preprocessing) 레이어가 필요하다. 제안하는 파이프라인은 [그림 1]과 같은 5단계 흐름을 갖는다.

[그림 1] 실시간 에셋 최적화 통합 파이프라인

첫째, 지능형 생성(Intelligent Generation) 방식을 사용한다. 사용자의 의도를 반영한 텍스트 또는 이미지를 기반으로 로우(Raw) 3D 데이터를 생성한다. 이때 멀티뷰 확산 모델(Multiview Diffusion)을 사용하여 시각적 일관성을 확보한다.[7)]

둘째, 전처리 및 클리닝(Preprocessing & Cleaning)을 실행한다. 중복 정점을 제거하고, 비정상적인 폴리곤 면을 수정하며, 노이즈 필터링을 통해 기하학적 노이즈를 억제한다.

davidasmith.medium.com/the-state-of-ai-in-aaa-game-production-b7d949df4daa (접속일자 2026.2.9.)

5) Xinyue Wei 외 8인, MeshLRM: Large Reconstrution Model for High-Quality Meshes, Conference on Computer Vision and Pattern Recognition, 2024, p.6

6) Xiaoyang Huang, Bingbing Ni, Wenjun Zhang, Generating Human-AI Collaborative Design Sequence for 3D Assets via Differentiable Operation Graph, Conference on Computer Vision and Pattern Recognition, 2025, p.1

7) Michelle Ellis, '3D Generative AI Transforms How We Create, Design, Interact With Digital Content', SIGGRAPH, 2025. 6.13. s2025.siggraph.org/3d-generative-ai-transforms-how-we-create-design-interact-with-digital-content/ (접속일자 2026.2.11.)

셋째, 자동 리토폴로지 및 메쉬를 압축한다. (Auto-Retopology) Meshoptimizer와 같은 라이브러리를 활용하여 정점 순서를 재정렬하고 폴리곤 수를 목표 수준으로 감소시킨다. 특히 Nanite를 고려하여 고폴리곤 밀도를 유지하되 연산 부하를 줄이는 클러스터링 기반 압축을 수행한다.

넷째, 텍스처 아틀라스 및 PBR을 정렬한다. (Texture Atlasing). 산재된 텍스처를 하나의 아틀라스로 결합하여 드로우 콜을 줄이고,[8] AI가 생성한 텍스처를 물리 기반 렌더링(PBR) 규격에 맞춰 정규화 한다.

다섯째, 엔진 라이브 스트리밍(Engine Live-Streaming)을 한다. UE5의 FBX 혹은 USD 파이프라인을 통해 최적화된 에셋을 즉각적으로 임포트하고 Nanite 설정을 자동 적용한다.

3-2. 메쉬 및 정점 최적화 알고리즘의 심층 분석

실시간 렌더링 성능을 극대화하기 위해 파이프라인은 다음과 같은 수리적 최적화 기법을 적용한다. Meshopt 라이브러리의 알고리즘을 활용하여 정점 셰이더 부하를 줄이는 방식이 대표적이다.

- Vertex Cache Optimization: GPU의 정점 캐시 적중률(Hit Ratio)을 높이기 위해 삼각형 인덱스 순서를 재구성한다.
- Overdraw Optimization: 화면상에서 픽셀이 중복되어 그려지는 오버드로우 현상을 줄이기 위해 삼각형 렌더링 순서를 정렬한다.
- Vertex Quantization: 정점 데이터를 16비트 혹은 그 이하의 정밀도로 압축하여 메모리 대역폭 점유율을 낮춘다.

프레임 시간(ms)은 CPU의 게임 로직 처리 시간, 드로우 콜 준비 시간, 그리고 GPU의 렌더링 시간의 합으로 결정되며, 최적화 파이프라인은 이 중 GPU 시간과 드로우 콜 시간을 줄이는 데 집중한다.[9]

첫째, 프레임 시간 분해 모델을 정의한다. 전체 프레임 시간은 파이프라인의 각 단계별 소요 시간의 총합의로 정의되며, [그림 2]와 같은 수식으로 나타낼 수 있다.

$$T_{total} = T_{logic} + T_{dc} + T_{gpu}$$

[그림 2] 전체 프레임 시간 산출 공식

각 항의 전문적 정의는 다음과 같다. T_{logic}(Game Logic Time)은 CPU에서 수행되는 물리 연산, AI 경로 탐색, 사용자 입력 처리 및 게임 상태 업데이트에 소요되는 시간이다. T_{dc} (Draw Call Preparation Time)는 CPU가 GPU에 렌더링 명령을 전달하기 위해 그래픽 드라이버와 통신하며 자산(Asset) 정보를 준비하는 오버헤드 시간이다. T_{gpu} (GPU Rendering Time)는 GPU가 전달받은 데이터를 바탕으로 정점 처리(Vertex Processing), 레스터화(Rasterization) 및 픽셀 셰이딩을 거쳐 최종 화면을 그리는 시간이다.

둘째, 본 연구를 위한 최적화 함수 모델 [그림 3]은 게임 로직 처리 시간을 상수로 고정(Const.)한 상태에서, 드로우 콜 및 GPU 렌더링 시간의 합을 최소화함으로써 전체 프레임 시간을 단축하고자 하는 최적화 목적 함수이다.

$$\begin{aligned} \text{minimize} \quad & T_{total} \approx T_{dc} + T_{gpu} \\ \text{subject to} \quad & T_{logic} \approx \text{constant} \end{aligned}$$

[그림 3] 최적화 목표 함수

일반적인 게임 최적화 파이프라인에서는 CPU의 로직 처리(T_{logic})를 최하 방어선으로 고정하고, 그래픽스 파이프라인의 병목 현상인 드로우 콜(T_{dc})과 렌더링 부하(T_{gpu})를 최소화하는 것을 목표로 한다.

셋째, 본 연구의 실험 데이터인 Tris(폴리곤 개수)와 Draw Call 수치를 [그림 4]의 공식에 대입한다.

8) Jaroslav Hofierka, Marcel Zlocha, A visualization-oriented 3D method for efficient computation of urban solar radiation based on 3D–2D surface mapping, International Journal of Geographical Information Science, Volume 28, Issue 4, 2014, p.824-825

9) Michele Riccio, Francesco Di Franco, Marco Rossoni, Giorgio Vassena, Optimization of Real-Time Rendering Pipelines in Virtual Production, MDPI, Applied Sciences, Volume 14, Issue 3, 2024, p.7

$$T_{dc} \propto N_{dc} \quad (\text{Draw Call 수에 비례})$$

$$T_{gpu} \propto \sum (Tris \times \text{Shader Complexity})$$

[그림 4] 세부 지표와의 상관관계

본 연구에서의 최적화 파이프라인은 전체 프레임 시간(T_{total})을 결정하는 핵심 요소 중 CPU 로직 시간을 제외한 드로우 콜 처리 시간(T_{dc}) 및 GPU 렌더링 시간(T_{gpu})의 최소화를 목적으로 한다. 이를 위해 자산의 폴리곤 최적화를 통한 T_{gpu} 감소와 배칭(Batching) 기술을 활용한 T_{dc} 절감을 동시에 수행하여 최종 FPS를 개선하였다.

4. 실시간 성능 지표 및 벤치마킹 분석

본 연구에서 제안한 최적화 파이프라인의 실효성을 벤치마킹하고 향후 실험의 재현성을 보장하기 위해, 다음과 같은 엄격한 하드웨어 및 통제 조건하에서 성능 평가를 수행하였다.

하드웨어 및 시스템 환경은 Intel Core i7-13700K CPU, NVIDIA GeForce RTX 4070 Ti (12GB VRAM) GPU, 32GB RAM이 탑재된 윈도우 PC 환경에서 테스트를 진행하였다.

테스트 씬(Scene)의 구성은 Unreal Engine 5.3 버전을 기준으로, Lumen 글로벌 일루미네이션을 활성화한 동적 라이팅(Dynamic Lighting) 환경의 표준 벤치마크 씬을 구성하였다.

샘플 수 및 반복 횟수는 표본의 통계적 유의성을 확보하기 위해 원거리 배경물 및 중심 캐릭터 에셋 등 각 유형별로 50개의 AI 생성 에셋 샘플을 무작위 추출하였다. 각 에셋은 동일한 뷰포트 카메라 앵글과 렌더링 조건에서 30회씩 반복 측정하여 평균 프레임(FPS) 및 드로우 콜 데이터를 산출하였다.

통제 조건은 OS 백그라운드 프로세스에 의한 성능 간섭을 최소화하기 위해 독립형(Standalone) 빌드 환경에서 실행하였으며, 메모리 캐시 누적으로 인한 오차를 방지하고자 매 렌더링 세션마다 VRAM을 초기화한 후 측정을 진행하였다.

4-1. 렌더링 성능 평가 기준

에셋 최적화의 유효성을 검증하기 위해 본 연구에서는 초당 프레임 수(FPS), 드로우 콜 수, 정점 페치 스톨(Vertex Fetch Stall), 그리고 VRAM 점유율을 주요 지표로 설정한다. 특히 VR과 같은 몰입형 환경에서는 90 FPS(11.11ms) 이상의 고정 프레임 유지가 필수적이므로 최적화 기준이 더욱 엄격하게 적용되어야 한다.[10] 그래서 [표 2]를 통해 렌더링 성능을 검증하였다.

[표 2] 렌더링 성능 검증

성능 지표	60 FPS 기준 목표치	90 FPS (VR) 기준 목표치	설명
Frame Time	< 16.67 ms	< 11.11 ms	전체 렌더링 주기
Draw Calls	< 1,000 (모바일)	< 500 (최적)	CPU - GPU 명령 횟수
Polygon Count	Nanite 활용 시 무제한 가깝게 확장 가능	선별적 활용 권장	화면 해상도 대비 밀도 조절
Texture Size	2048 x 2048 이하 권장	1024 x 1024 권장	비디오 메모리 점유 효율

4-2. 최적화 전후 성능 비교 데이터

생성형 AI 도구인 Tripo AI와 Rodin AI를 통해 생성된 에셋을 UE5 환경에서 테스트한 결과, 최적화 파이프라인 적용 여부에 따라 극명한 성능 차이가 발생함을 확인할 수 있었다. [표 3]

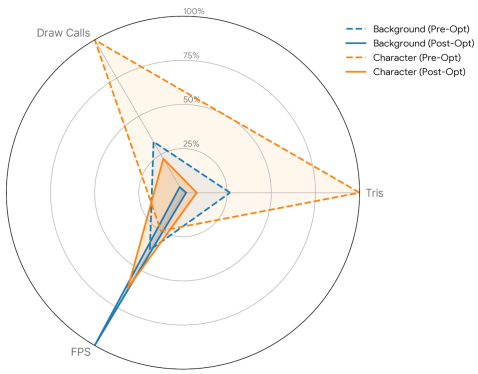
[표 3] 에셋 성능 비교 분석

에셋 유형	최적화 단계	평균 Tris	Draw Calls	FPS (평균)	비고
원거리 배경물	최적화 전	120,000	18	45	과도한 기하학 복잡도
	최적화 후	8,000	2	120	LOD 및 메쉬 통합 적용
중심	최적화	450,000	54	30	아티팩트 발생

10) Schütz, M., Krösl, K., & Wimmer, M., A Point-Based and Image-Based Multi-Pass Rendering Technique for Visualizing Massive 3D Point Clouds in VR Environments, Journal of WSCG, Vol. 26, No. 2, 2018, p.158

캐릭터	전				가능성 높음
	최적화 후	35,000	12	75	리토폴로지 및 PBR 정렬 완료

테스트 결과, 최적화 파이프라인 적용 시 에셋 제작 시간은 9배 단축되었으며, FPS는 원거리 배경물 기준 45에서 120으로 크게 향상되었다.



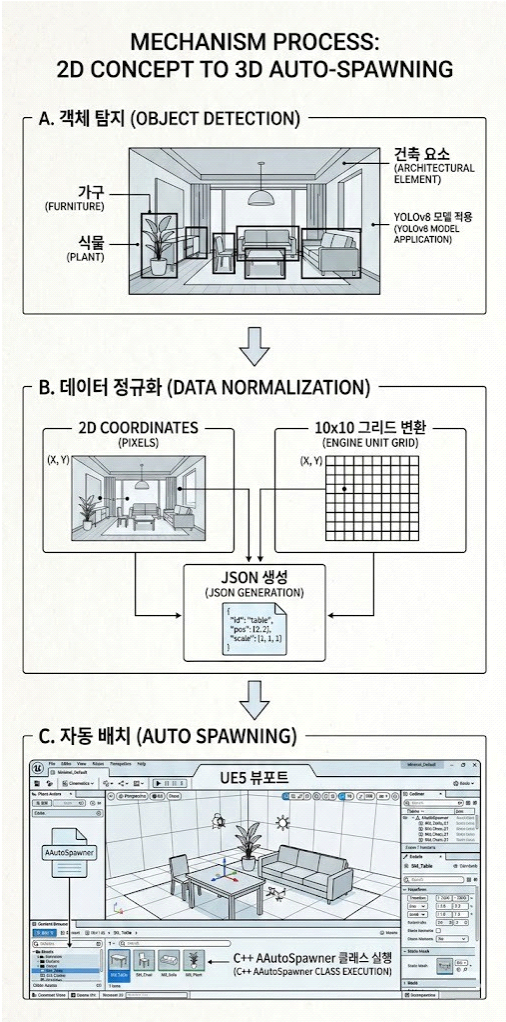
[그림 5] 에셋 최적화 상태별 통합 지표 비교

이러한 [그림 5]의 수치는 최적화 파이프라인이 단순히 파일 용량을 줄이는 것을 넘어, 실제 게임 플레이나 가상 현실 환경에서의 체감 성능을 두 배 이상 향상시킬 수 있음을 입증하였다.

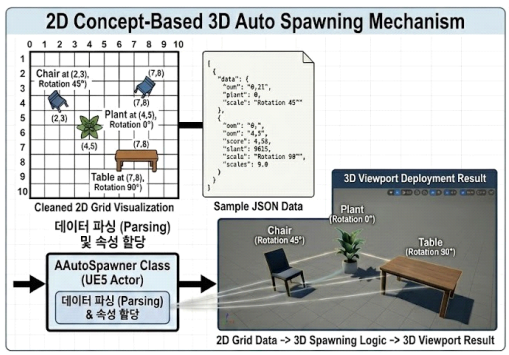
5. AI 기반 환경 구축 및 자동 배치 시스템

5-1. YOLOv8 및 컴퓨터 비전을 활용한 절차적 배치

단일 에셋 최적화를 넘어, 다수의 AI 생성 에셋을 환경에 배치하는 과정에서도 AI 기술이 적극적으로 활용된다. 최신 연구에 따르면 YOLOv8과 같은 객체 탐지 모델을 사용하여 2D 컨셉 이미지에서 사물의 위치와 크기를 식별하고, 이를 UE5의 C++ 액터 시스템과 연동하여 3D 공간에 자동으로 스포닝(Spawning)하는 기술이 개발되었다.[11)]



[그림 6] YOLOv8 메커니즘



[그림 7] 'C.자동 배치'의 플로우 차트

이 시스템은 [그림 6]과 같은 메커니즘으로 작동한다. 첫째, 객체를 탐지한다. 2D 이미지 내에서 가구,

11) 이성현, 김진모, YOLOv8 기반 VR 물류 교육 플랫폼에서의 실시간 객체 탐지 및 상호작용 연구, 한국플랫폼컨설팅학회, 플랫폼컨설팅학회 논문지, Vol. 13, No. 1, 2025, pp.5-7

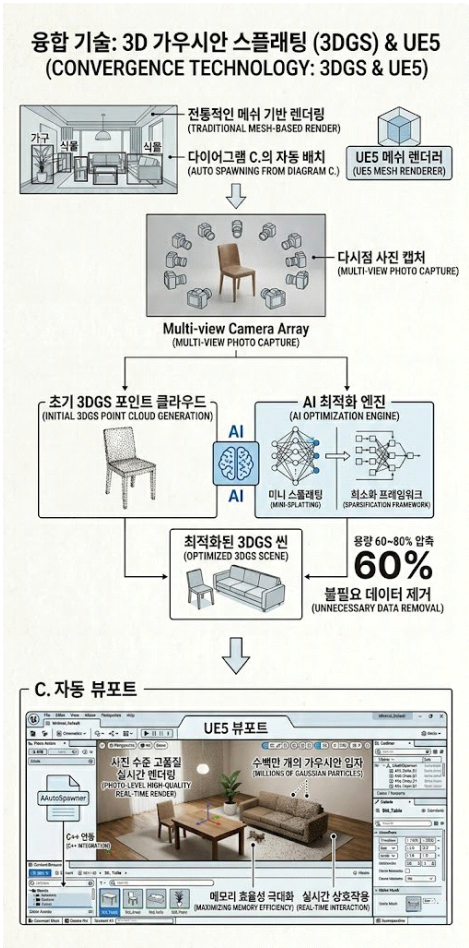
식물, 건축 요소 등을 식별하고 레이블링한다. 둘째, 데이터를 정규화한다. 2D 좌표를 10x10 그리드와 같은 엔진 유닛 단위로 변환하여 JSON 형식으로 저장한다. 셋째, 자동 배치를 한다. UE5의 AAutoSpawner 클래스가 JSON을 읽어들여 사전에 등록된 최적화 에셋을 해당 위치에 정확한 방향과 스케일로 배치한다. [그림 7]은 'C. 자동 배치 매커니즘'의 플로우 차트를 구체적으로 묘사하였다.

5-2. 3D 가우시안 스플래팅(3DGS)과의 융합

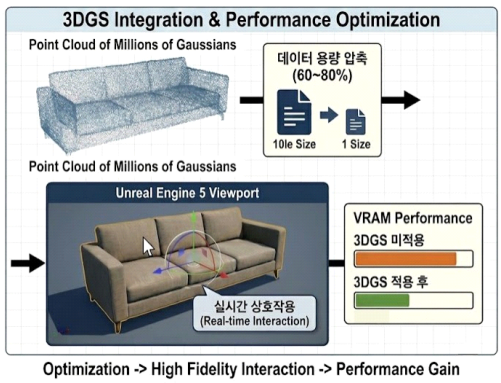
전통적인 메쉬 기반 렌더링 외에도, AI를 활용한 3D 가우시안 스플래팅 기술이 UE5에 통합되고 있다. [그림 8]은, 사진 수준의 고품질 렌더링을 실시간으로 구현하면서도, 경량형 스플래팅 (Mini- Splatting)이나 희소화(Sparsification) 프레임워크를 통해 메모리 효율성을 극대화한다. AI는 수백만 개의 가우시안 입자를 최적화하여 불필요한 데이터를 제거하고, 씬 전체의 용량을 60~80%까지 압축하면서 시각적 충실도를 유지하는 역할을 수행한다.

해당 파이프라인은 수백만 개의 포인트로 이루어진 방대한 3DGS(3D Gaussian Splatting) 데이터를 실시간 렌더링 환경인 언리얼 엔진 5에 효율적으로 통합하기 위한 고도화된 최적화 과정을 담고 있다. 프로세스의 시작점인 초기 스캔 데이터는 시각적 디테일은 매우 뛰어나지만, 방대한 연산량과 메모리 점유율로 인해 실시간 환경에서 직접 구동하기에는 한계가 명확하다. 이를 해결하기 위해 시스템의 핵심 단계인 '데이터 용량 압축' 과정을 거치게 되며, 이때 불필요한 가우시안 포인트를 정리하고 속성을 병합함으로써 원본 대비 약 60~80%의 용량을 절감하여 데이터 크기를 최대 1/10 수준으로 경량화한다.

이러한 전략적 압축은 하드웨어 자원의 효율성을 극대화하여, 우측 하단의 성능 지표가 보여주듯 VRAM (비디오 램) 사용량을 획기적으로 낮추는 결과를 가져온다. [그림 9]를 보면, 결과적으로 최적화된 3DGS 자산은 언리얼 엔진 5 뷰포트 내에서 낮은 시스템 부하로 안정적으로 구동되며, 단순한 시각화를 넘어 기즈모를 활용한 위치 이동이나 실시간 조작 등 고충실도의 실시간 상호작용이 가능해진다. 이는 최종적으로 고품질의 그래픽 데이터를 대규모 씬에서도 성능 저하 없이 운용할 수 있는 강력한 성능 이득(Performance Gain)을 보장한다.



[그림 8] 3D 가우시안 스플래팅과 UE5



[그림 9] C.의 퍼포먼스 최적화를 위한 플로우 차트

6. 인간과 AI의 협업 모델 (HITL)

6-1. 예술적 제어권(Art Directability)의 확보

생성형 AI가 모든 에셋을 완벽하게 생성할 수 있다는 기대는 아직 시기상조이다. AI는 특히 복잡한 기계 장치나 가독성이 필요한 텍스트, 그리고 애니메이션이 요구되는 관절 부위에서 치명적인 오류를 범할 가능성이 높다. 따라서 'AI-인간 협업(Human-in-the-loop)' 모델은 AI가 생성한 결과물에 예술적 의도를 투영하고 기술적 완성도를 부여하는 필수적인 과정이다.

이 중에 80/20 원칙은 AI가 반복적이고 노동 집약적인 기초 모델링 및 텍스처링의 80%를 처리하고, 인간 아티스트는 실루엣 수정, 중요 UV 심(Seam) 배치, 그리고 최종 셰이더 미세 조정의 20%를 담당한다. 그리고, QA 파이프라인은 AAA급 스튜디오에서는 AI 에셋이 엔진 빌드에 포함되기 전, 위상학적 오류나 시각적 아티팩트를 검수하는 엄격한 품질 보증 단계를 거친다.

6-2. 운영 시퀀스 기반의 편집 유연성

최신 연구는 단순한 결과물로서의 메쉬가 아닌, 제작 과정의 논리적 순서를 담은 '운영 시퀀스(Operation Sequence)'를 생성하는 AI 모델을 제안한다. 이는 아티스트가 AI가 수행한 특정 단계를 역추적하여 수정할 수 있게 함으로써, 기존의 밀도가 높은 '폴리곤 수프(Polygon Soup)'를 직접 정점 단위로 수정해야 했던 고통스러운 과정을 획기적으로 개선한다. 이러한 방식은 3D 생성 분야에서 마치 2D 레이어 편집처럼 직관적인 제어권을 제공한다.

7. 실증 사례 연구 및 산업적 효용

7-1. 인디 및 중소 규모 스튜디오의 생산성 향상

Quick Turn Games와 같은 중소 규모 UE5 스튜디오의 사례는 AI 최적화 파이프라인의 산업적 가치를 [표 4]와 같이 증명한다. C++ 프로그래머 위주의 팀 구성에서 100개 이상의 유니크한 에셋을 제작하는 것은 불가능에 가까웠으나, 3D AI Studio의 이미지-투-3D 워크플로우를 도입하여 개당 제작 시간을 14시간에서 1.5시간으로 줄였으며, 이는 플레이 테스트 결과 사용자들도 AI 생성 여부를 구분하지 못할 정도의 품질을 보여주었다.

[표 4] 프로젝트 단계별 최적화 비교

프로젝트 단계	소요 시간 (AI 미도입)	소요 시간 (AI 도입)	시간 절감률
컨셉아트 확정	3일	4시간 (Gemini 활용)	83%
베이스 모델링	10일	2시간 (Tripo AI)	98%
최적화 및 엔진 배치	5일	1일 (자동 파이프라인)	80%

7-2. 가상 프로덕션 및 교육 분야의 응용

생성형 AI를 활용한 실시간 에셋 최적화는 게임뿐만 아니라 영화 제작의 가상 프로덕션 단계에서도 큰 위력을 발휘한다. 감독이 현장에서 즉석으로 프롬프트를 입력하여 가상 배경 에셋을 생성하고, Nanite 설정을 통해 즉시 고품질 배경으로 활용하는 시나리오는 제작 비용을 수백만 달러 절감할 수 있는 잠재력을 가진다. 또한 교육 분야에서는 학생들이 기술적인 모델링 숙련도에 얽매이지 않고 창의적인 레벨 디자인과 스토리텔링에 집중할 수 있는 환경을 제공하여 교육의 패러다임을 변화시키고 있다.

8. 기술적 한계 및 윤리적 과제

8-1. 하드웨어 의존성과 데이터 품질 문제

UE5와 AI를 결합한 파이프라인은 강력한 성능을 발휘하지만, 동시에 높은 하드웨어 사양을 요구한다. 특히 실시간 신경망 렌더링이나 가상화 지오메트리 처리는 고성능 GPU와 대용량 비디오 램(VRAM)을 소모하며, 이는 클라우드 기반 렌더링이나 최적화된 저사양용 파이프라인 개발의 필요성을 제기한다. 또한 AI 모델이 학습한 데이터셋의 품질에 따라 편향된 결과물이 나오거나, 세밀한 표면 디테일이 뭉개지는 현상은 여전히 해결해야 할 기술적 과제이다.

8-2. 저작권 및 윤리적 가이드라인

AI 생성 콘텐츠의 저작권 소유권 문제는 법적, 윤리적으로 복잡한 논쟁을 불러일으키고 있다. 상업적 프로젝트에서 AI 도구를 사용할 때는 데이터의 출처가 투명한지, 그리고 타인의 지적 재산권을 침해하지 않는지에 대한 법적 검토가 선행되어야 한다. AAA급 스튜디

오들이 AI 도입에 신중한 태도를 보이는 이유 중 하나는 브랜드 명성과 저작권 보호 사이의 균형을 유지하기 위함이다.

9. 결론

기존의 연구들이 주로 3D 모델의 생성 알고리즘(Text-to-3D, Image-to-3D) 자체를 향상시키거나 , 엔진 내 렌더링 성능 최적화라는 단편적인 영역에 치중했던 반면, 본 연구는 '생성-정제-엔진 통합'으로 이어지는 전 주기를 실시간으로 자동화하는 5단계 파이프라인을 구축했다는 데 독창성이 있다. 특히, 생성형 AI의 고질적 문제인 '폴리곤 수프' 형상을 단순히 줄이는 것을 넘어 , Unreal Engine 5의 핵심 가상화 지오메트리 기술인 Nanite 및 Lumen 규격에 맞게 자동 리토폴로지하고 클러스터링하는 과정을 최초로 제안하였다. 나아가 기술적 한계를 보완하기 위한 '인간-AI 협업(HITL)' 모델을 파이프라인 내에 내재화하여, 실제 상용 프로덕션 레벨에서의 실효성과 예술적 제어권을 동시에 확보한 것은 기존 자동화 기술들과 뚜렷하게 구별되는 본 연구만의 독자적인 기여점이다.

본 연구를 통해 Unreal Engine 5와 생성형 AI를 결합한 실시간 에셋 최적화 파이프라인의 강력한 효용성과 기술적 타당성을 확인하였다. AI는 에셋 제작의 생산성을 수십 배 향상시킬 수 있는 잠재력을 지니고 있으며, 엔진 고유의 기술인 Nanite 및 Lumen과의 정교한 조화를 통해 실시간 고품질 렌더링을 구현할 수 있다.

연구의 주요 성과는 다음과 같다.

첫째, 파이프라인의 자동화가 개선되었다. 단순 생성 단계를 넘어 엔진 임포트 규격에 맞춘 메쉬 및 텍스처 최적화 공정을 체계화하였다.

둘째, 성능 지표가 정립되었다. FPS, Frame Time, Draw Calls 등의 객관적 지표를 통해 AI 에셋의 프로덕션 적합성을 평가하는 기준을 제시하였다.

셋째, 협업 모델의 유효성이 검증되었다. 인간의 예술적 판단과 AI의 연산 효율성을 결합한 'AI-인간 협업 모델(Human-in-the-loop)'이 가장 현실적인 대안임을 입증하였다.

향후 연구에서는 텍스트 프롬프트만으로 전체 월드의 물리 환경과 상호작용 가능한 레벨을 구축하는 'Text-to-World' 시스템의 실현 가능성을 모색하고, 온디바이스(On-device) AI 기술을 통해 하드웨어 요구사항을 낮추면서도 실시간성을 유지하는 경량화 알고리즘에 대한 연구가 지속되어야 할 것이다. 이러한 기술적 진보는 가상 세계 구축의 장벽을 낮추고, 인류의 디지털 창의성을 극대화하는 기폭제가 될 것으로 기대된다.

참고문헌

1. 김태완, 이재호, 디지털 트윈 구현을 위한 게임 엔진의 렌더링 성능 한계 분석 및 최적화 방안, 한국정보처리학회, 한국정보처리학회 논문지: 소프트웨어 및 데이터 공학, 2022
2. 박준형, 최용석, 상용 게임 엔진 기반 실시간 렌더링 최적화를 위한 플러그인 설계 및 구현 연구, 한국게임학회, 한국게임학회 논문지, 2023
3. 이성현, 김진모, YOLOv8 기반 VR 물류 교육 플랫폼에서의 실시간 객체 탐지 및 상호작용 연구, 플랫폼컨설팅학회, 플랫폼컨설팅학회 논문지, 2025
4. Aryan N. Modi & Kuntal Ghose, Artificial Intelligence in game asset pipeline, JETIR, IJ Publication, 2026
5. D. Rodriguez, A. Martinez, J. Choi, and L. Smith, Optimization of Real-Time Rendering Pipelines in Virtual Production Using Cult3D Frameworks, IEEE, IEEE Access(SCIE), 2024
6. Michele Riccio, Francesco Di Franco, Marco Rossoni, Giorgio Vassena, Optimization of Real-Time Rendering Pipelines in Virtual Production, MDPI, Applied Sciences, 2024
7. Jaroslav Hofierka, Marcel Zlocha, A visualization-oriented 3D method for efficient computation of urban solar radiation based on 3D-2D surface mapping, International Journal of Geographical Information Science, Taylor & Francis, 2014

8. J. Zhang, et al., Research on a Fusion Technique of YOLOv8-URE-Based 2D Vision and Point Cloud for Robotic Grasping in Stacked Scenarios, MDPI, Applied Sciences, 2025
9. Schütz, M., Krösl, K., & Wimmer, M., A Point -Based and Image-Based Multi-Pass Rendering Technique for Visualizing Massive 3D Point Clouds in VR Environments, Journal of WSCG, University of West Bohemia, 2018
10. Shyngys Adilkhan, Madina Alimanova, Lei Shi, Aiganym Soltiyeva, Advances in artificial intelligence -driven 3D model generation: a review of GAN and VAE methodologies, Bulletin of Electrical En- gineering and Informatics, IAES, 2025
11. Xiaoyang Huang, Bingbing Ni, Wenjun Zhang, Generating Human-AI Collaborative Design Sequence for 3D Assets via Differentiable Operation Graph, Conference on Computer Vision and Pattern Recognition, IEEE & CVF, 2025
12. Xinyue Wei 외 8인, MeshLRM: Large Reconstrution Model for High-Quality Meshes, Conference on Computer Vision and Pattern Recognition, IEEE & CVF, 2024
13. Zixun Shen, Xiangyun Liao, Yi-Ping Phoebe Chen, and Jin Huang, Moving Towards Large-Scale Particle Based Fluid Simulation in Unity 3D, MDPI, Applied Sciences, 2022
14. www.epicgames.com
15. www.hyper3d.ai
16. www.medium.com
17. www.siggraph.org
18. www.sloyd.ai
19. www.tripo3d.ai