한국컴퓨터그래픽스학회
Korea Computer Graphics Society Vol. 32, No. 3, P. 85~96

# 사용자 상호작용을 지원하는 SAM3D 기반 단일 뷰 실내 공간 및 가구 3D 재구성 시스템[1]

심우현 ○ 위영림 손예진 김성기*
조선대학교 컴퓨터공학과
{swh, wyr3354, ths9316, skkim}@chosun.ac.kr

# User-Interactive Single-View 3D Reconstruction of Indoor Environments and Furniture based on SAM3D

Woo-hyun Sim ○ Young-rim Wee Ye-jin Son SeongKi Kim*
Dept. of Computer Science and Engineering, Chosun University

## 요 약

기존 상용 인테리어 앱들은 자체 제공 3D 카탈로그만을 활용하여 미리 정해진 가구들만 배치할 수 있고 사용자가 실제 보유한 가구를 배치할 수 없으며 배치한 가구의 위치를 변경해볼 수 없다는 근본적 한계를 가진다. 본 논문은 단일 실내 사진으로부터 가구를 자동 분리·복원·3D 변환하여 인터랙티브하게 재배치할 수 있는 end-to-end 인테리어 재구성 시스템을 제안한다. 본 논문에서 제안한 시스템은 SAM2 기반 인스턴스 분할(Segmentation), LaMa 및 Stable Diffusion (SD) ControlNet(Canny)을 이용한 2 단계 빈 방 재구성 파이프라인, SAM3D 기반 3D 가구 복원 파이프라인, 그리고 Three.js 기반 인터랙티브 3D 배치 시스템으로 구성된다. 본 논문에서는 실험을 통해 버텍스 컬러 방식(3D 메쉬의 각 꼭짓점에 직접 색상값을 할당하는 방식)이 텍스처 베이킹 방식(UV 좌표 공간에 색상 정보를 투영·저장하는 후처리 방식)대비 색상 번짐이 없고, 가구 형상 복잡도와 무관하게 처리 시간이 안정적이어서 실시간 시스템에 우수함을 증명하였다. 또한, SAM3D 를 방 전체 3D 재구성에 적용하는 탐색적 실험을 수행하여 볼록한 단일 객체를 가정한 shape prior 로 인해 오목한 실내 공간에서 벽·바닥 메쉬 생성 실패 및 색상 일관성 저하 문제가 발생함을 확인하고 trimesh 후처리 및 평면 snap 으로 완화하였다. 이후 실제 방 이미지에서 색상·텍스처를 추출하여 Three.js 기반 절차적 3D 방을 구성하는 방식을 사용하였다. 마지막으로 20 명 대상 사용자 평가를 통해 시스템의 실용성을 정량적으로 검증하였다.

## Abstract

This paper presents an end-to-end interactive interior redesign system that automatically segments, completes, and reconstructs furniture from a single indoor photograph for interactive 3D placement. Existing commercial interior applications are fundamentally limited to their own 3D furniture catalogs, preventing users from incorporating their existing furniture. To overcome this limitation, the proposed system consists of a SAM2-based instance segmentation pipeline, a two-stage empty room reconstruction pipeline using LaMa and Stable Diffusion ControlNet (Canny edge), a SAM3D-based furniture 3D reconstruction pipeline, and an interactive placement system built on Three.js. Experiments demonstrate that the vertex color rendering method produces no color bleeding artifacts and maintains stable processing time regardless of furniture geometric complexity, proving far superior for real-time systems compared to UV texture baking. Furthermore, we conducted an exploratory experiment applying SAM3D to full-room 3D reconstruction. However, due to the convex single-object shape prior, mesh generation failures on walls and floors as well as color inconsistency issues occurred in the concave indoor space. While trimesh post-processing and plane snap partially mitigated these problems, a complete solution remained elusive. As an alternative, we adopted a Three.js-based procedural 3D room construction approach that extracts color and texture directly from the input room image; complete SAM3D-based 3D room reconstruction remains a direction for future work. Finally, a user study with 20 participants was conducted to quantitatively validate the practical utility of the proposed system.

**키워드:** 인테리어 재구성, SAM3D, 인스턴스 분할, 이미지 인페인팅, 버텍스 컬러

**Keywords:** Interior Redesign, SAM3D, Instance Segmentation, Image Inpainting, Vertex Color

## 1. 서론

인테리어 디자인은 가구 배치와 공간 구성에 따라 생활 환경의 질이 크게 달라지는 중요한 활동이다. 그러나 실제 가구를 구매하거나 이동하기 전에 배치를 미리

[1] 학부생 주저자 논문임

*corresponding author: SeongKi Kim / Dept. of Computer Science and Engineering, Chosun University (skkim@chosun.ac.kr)

Received : 2026.05.28./ Review completed : 1st 2026.06.07. 2nd 2026.06.25./ Accepted : 2026.06.29.
DOI : 10.15701/kcgs.2026.32.3.85
ISSN : 1975-7883(Print)/2383-529X(Online)

시뮬레이션하는 것은 기술적 제약으로 인해 일반 사용자에게 쉽지 않은 과제였다.

기존의 상용 인테리어 시뮬레이션 앱들은 IKEA Place [1], Wayfair, 오늘의 집 등 자체적으로 구축한 3D 가구 카탈로그를 기반으로 동작한다. 이러한 방식은 앱이 제공하는 카탈로그 내 제품만 배치 가능하여, 사용자가 현재 보유한 가구를 활용하지 못한다는 근본적 한계를 갖는다.

이러한 제약의 기술적 원인은 세 가지이다. 첫째, 임의의 가구 사진으로부터 고품질 3D 모델을 자동 생성하는 기술에 최근까지 한계가 있었다. 둘째, 실내 사진에서 개별 가구를 인스턴스 수준으로 정확히 분리하는 분할 정확도가 충분하지 않았다. 셋째, 가구 간 가림(occlusion)으로 인해 불완전하게 촬영된 가구를 자동으로 복원하는 것이 어려웠다.

그러나 최근 3D Gaussian Splatting [6], SAM3D [7], SAM2 [3], Stable Diffusion Inpainting [5] 등 핵심 기반 기술들이 빠르게 발전하면서 이러한 한계를 극복할 수 있는 환경이 마련되었다. 본 논문은 이러한 최신 기술들을 사용하여, 단일 실내 사진으로부터 가구를 자동으로 분리·복원·3D 변환하고 인터랙티브하게 재배치할 수 있는 end-to-end 시스템을 연구하였다.

본 논문의 주요 기여점은 다음과 같다.

· **LaMa + SD(Stable Diffusion) ControlNet(Canny) 2 단계 빈 방 재구성 파이프라인 제안.** 기존 SD 단독 인페인팅의 환각 현상(Hallucination) 문제를 해결하며, 3D 메쉬나 도메인 미세조정(Fine-tuning) 없이 단일 2D 사진만으로 동작한다.

· **SAM3D 를 실내 가구 복원에 적용하기 위한 입력 최적화 전략 제안.** 배경 요구사항 상충 문제는 컨텍스트 패딩 비율 분석으로, 랜덤 시드 의존성 문제는 카테고리별 다중 시드 선택 기준으로 해결하였다.

· **SAM3D 의 실내 공간 전체 적용 탐색 및 한계 분석.** 단일 객체 shape prior 가 오목한 실내 공간에서 갖는 구조적 한계를 규명하고, Three.js 기반 절차적 방 구성으로 대체한 설계 판단을 제시한다.

# 2. 관련 연구

## 2.1 가구 인식 및 분할(Segmentation)

실내 장면에서 개별 가구를 정확히 인식하고 분리하는 분할 기술은 인테리어 재구성 시스템의 핵심 전제 조건이다. Kirillov et al. [2]의 SAM(Segment Anything Model)은 대규모 데이터로 학습된 범용 분할 모델로, 사용자의 포인트·박스·텍스트 프롬프트를 기반으로 임의의 객체를 분리할 수 있다. 그러나 SAM 은 단일 이미지에만 적용 가능하다는 한계가 있다. 이를 개선한 Ravi et al. [3]의 SAM2 는 이미지와 비디오 모두에서 인스턴스 수준의 분할을 지원하며, 다중 포인트 프롬프트를 통해 복잡한 형상의 객체에서도 높은 정확도를 달성한다. 본 논문은 SAM2 를 활용하여 실내 사진에서 개별 가구를 인스턴스 단위로 분리한다.

## 2.2 이미지 인페인팅 및 빈 방 재구성

Suvorov et al. [4]의 LaMa 는 푸리에 컨볼루션을 활용하여 대규모 마스크 영역의 구조적 복원에 탁월한 성능을 보인다. Rombach et al. [5]의 Stable Diffusion 기반 인페인팅은 잠재 확산 모델을 활용하여 의미론적으로 일관된 고해상도 복원 결과를 생성한다.

그러나 Stable Diffusion 단독 적용 시 가구 제거 맥락에서 그림자와 마스크 경계의 잔여물을 설명하려는 환각 현상 문제가 발생함이 보고되었다 [11].

Dolhasz et al. [12]은 3D 메쉬에서 추출한 Canny edge 를 ControlNet [8] 조건으로 사용하여 구조적으로 일관된 파노라마 인페인팅을 달성하였다. 그러나 이 방법은 파노라마 카메라와 3D 메쉬 스캔 장비가 필요하며, 실내 전용 도메인 미세조정을 요구한다는 점에서 일반 사용자 환경에 적용하기 어렵다. 본 논문은 LaMa 의 구조 복원 결과에서 직접 Canny edge 를 추출하여 SD ControlNet 조건으로 활용함으로써, 추가 장비나 미세조정 없이 단일 2D 사진만으로 동일한 효과를 달성한다.

## 2.3 단일 이미지 기반 3D 복원

단일 이미지로부터 3D 객체를 복원하는 연구는 다양한 접근법을 취한다. pixelNeRF [10]은 소수 이미지에서 NeRF 기반 3D 표현을 학습한다. TripoSR [15]은 대규모 3D 데이터로 사전 학습된 트랜스포머 기반 모델로, 단일 이미지에서 빠른 3D 복원이 가능하다. 그러나 이러한 방법들은 가구 일부가 가려진 occlusion 환경에서 완전한 형상 복원에 한계를 보인다. SAM3D [7]는 대규모 3D 오브젝트 데이터셋으로 학습된 3D shape prior 를 활용하여 단일 이미지에서 가려진 영역의 3D 형상을 자동으로 추정한다. Baradad et al. [13]은 배경 컨텍스트가 단안 깊이 추정 성능에 중요한 영향을 미침을 보였다. 본 논문은

SAM3D 를 실내 가구 복원에 적용하는 과정에서 깊이 추정과 형상 복원의 배경 요구사항이 상충함을 발견하고 이를 해결하는 입력 분리 전략을 제안한다.

### 2.4 인테리어 시뮬레이션 시스템

기존 2D 인테리어 시뮬레이션 시스템은 평면 이미지 위에 가구 이미지를 합성하는 방식으로 동작하며, 원근감 표현 불가, 충돌 감지 불가, 다시점 확인 불가, 가구 간 가림(occlusion)으로 인해 불완전하게 촬영된 가구를 복원하지 못하는 등 구조적 한계를 가진다. IKEA Place [1] 등 AR 기반 시스템은 3D 배치가 가능하나 자체 카탈로그에만 의존한다. 본 논문의 시스템은 단일 사진으로부터 가구를 직접 3D 로 변환하여 물리 기반 충돌 감지와 다시점 확인이 가능한 완전한 3D 배치 환경을 제공한다는 점에서 기존 방법들과 근본적으로 차별화된다.

## 3. 제안 시스템

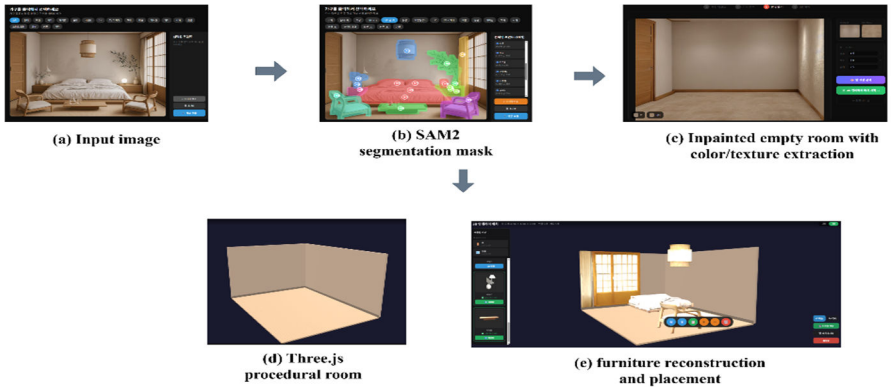



Figure 1. Overview of the proposed system: (a) input indoor image, (b) SAM2 segmentation mask, (c) color/texture extraction and inpainted empty room, (d) Three.js procedural 3D room generation, (e) SAM3D-based furniture reconstruction and interactive placement.

### 3.1 개요

제안 시스템은 Figure 1 과 같이 세 개의 파이프라인으로 구성되며, 각 파이프라인의 결과물이 Three.js 기반 인터랙티브 3D 씬에 통합된다. 첫 번째는 2D 빈 방 재구성 파이프라인(3.2 절)으로, 단일 실내 사진을 입력으로 받아 SAM2 기반 가구 분할과 LaMa + SD ControlNet 2 단계 인페인팅을 거쳐 가구가 제거된 빈 방 이미지를 출력한다. 두 번째는 Three.js 기반 인터랙티브 3D 방 구성 파이프라인(3.3 절)으로, 빈 방 이미지에서 색상·텍스처를 추출하여 절차적으로 3D 방을 구성한다. 세 번째는 가구 3D 복원 파이프라인(3.4 절)으로, 분할된 가구 크롭 이미지를 입력으로 받아 SAM3D 기반 다중 시드 최적화와 버텍스 컬러 적용을 통해 3D 가구 메쉬를 출력한다. 빈 방 재구성과 가구 복원 파이프라인은 병렬로 수행 가능하며, 최종 결과물은 Three.js 기반 인터랙티브 배치 시스템(3.5 절)에서 통합된다. 이 설계는 각 파이프라인의 독립 교체·개선이 가능하고, 빈 방 재구성과 가구 복원을 병렬로 수행하여 전체 처리 시간을 단축하는 장점을 갖는다.

### 3.2 2D 빈 방 재구성 파이프라인

#### 3.2.1 SAM2 기반 인스턴스 분할

빈 방 재구성의 첫 단계는 원본 실내 사진에서 가구 영역을 분리하는 것이다. 본 논문은 SAM2 [3]를 활용하여 가구 마스크를 추출한다. 사용자는 Figure 2 와 같은 UI 에서 가구 종류를 선택한 뒤 해당 가구 영역을 클릭하는 포인트 프롬프트 방식으로 분할을 수행한다. 가구 1 개당 평균 1~3 회 클릭으로 정확한 마스크를 획득할 수 있으며, 다중 포인트 입력을 통해 복잡한 형상의 가구에서도 정확도를 높일 수 있다. 동일 레이블의 포인트는 그룹화되어 가구별 독립적인 마스크로 생성된다.

추출된 마스크에는 마스크 경계를 정제하여 인페인팅 단계에서의 아티팩트를 최소화한다.

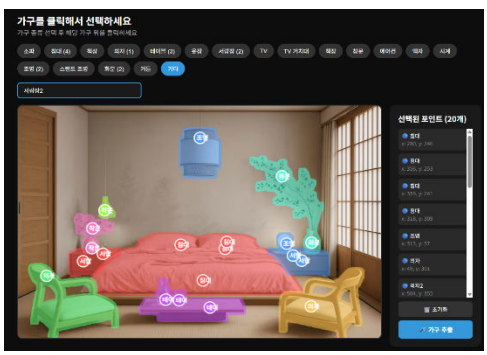


Figure 2. SAM2-based instance segmentation interface. An instance-level mask is generated when the user selects a furniture category and clicks on the target furniture.

### 3.2.2 LaMa + SD ControlNet 2 단계 인페인팅

가구 마스크를 이용하여 원본 사진에서 가구를 제거하고 빈 방을 재구성한다. 본 논문은 Stable Diffusion Inpainting 단독 적용 시 발생하는 환각현상 문제를 해결하기 위해 2 단계 순차 파이프라인을 제안한다.

Stable Diffusion 은 가구를 제거한 영역에서 그림자, 마스크 경계의 잔여물 등 시각적 단서를 설명하려는 경향으로 인해 새로운 가구나 구조물을 생성하는 환각현상이 빈번히 발생한다 [11]. 실제로 Canny edge 조건 없이 Stable Diffusion Inpainting 을 단독 적용할 경우, 모델은 마스크 영역의 기하학적 구조를 참조할 수 없어 원본 장면에 존재하지 않던 가구나 패턴을 임의로 생성하는 현상이 나타난다. 이는 빈 방 재구성이라는 목표와 상충하며, 실내 공간의 구조적 일관성을 저해한다. 기존 연구에서 Canny edge 기반 ControlNet 이 이를 완화함이 보고되었으나 [12], 해당 방법은 3D 메쉬와 도메인 미세조정을 필요로 한다.

본 논문은 이러한 문제를 위하여 Figure 3 과 같이 단일 2D 사진만으로 해결하는 방법을 제안한다. 먼저 SAM2 로 생성된 통합 가구 마스크를 LaMa [4]에 입력하여 바닥, 벽 등 배경 구조를 1 차 복원한다. LaMa 는 푸리에 컨볼루션 기반으로 대규모 영역의 구조적 일관성 복원에 강점을 보이나, 의미론적 완성도가 부족하다. 이어서 LaMa 1 차 복원 결과에서 추출하되 마스크 영역 내부 edge 를 제거한 Canny edge 를 ControlNet 조건으로 사용하여 Stable Diffusion Inpainting [5]을 적용한다. Canny edge 는 의미론적 사전 지식 없이 기하학적 구조를 포착하여 다양한 실내 환경에 적응적으로 동작하며, 모델이 마스크 영역 외부의 구조적 맥락을 참조함으로써 환각현상 발생을 억제한다. 이 2 단계 순차 적용 방식은 구조적 일관성과 시각적 사실감을 동시에 확보하며, 기존 방법 [12]과 달리 3D 정보나 도메인 미세조정 없이 단일 사진만으로 동작한다.

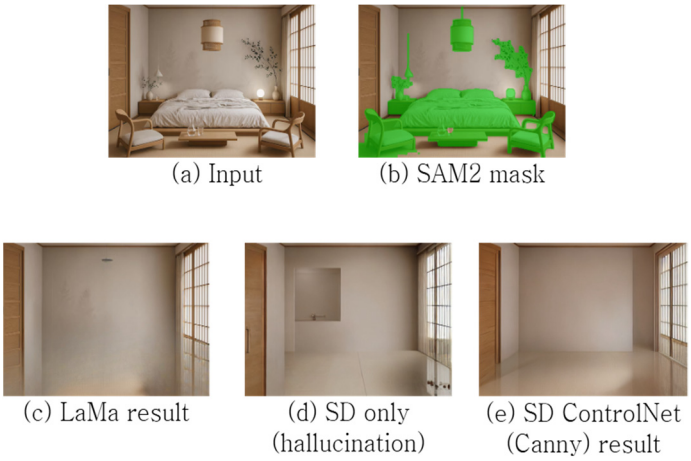



Figure 3. Empty room reconstruction pipeline. (a) Input image with furniture, (b) SAM2 segmentation mask, (c) LaMa inpainting result, (d) SD inpainting only: hallucination occurs as a ceiling fan not present in the original scene is generated in the masked region, (e) SD ControlNet (Canny): hallucination is suppressed through Canny edge conditioning, producing a structurally consistent empty room.

## 3.3 3D 방 메쉬 생성 파이프라인

### 3.3.1 SAM3D 기반 방 전체 복원 탐색 및 한계 분석

본 절은 최종 시스템에 채택되지 않은 탐색적 실험으로, SAM3D 를 실내 공간 전체에 적용했을 때의 가능성과 한계를 분석한다. 이 분석을 바탕으로 3.3.2 절의 절차적 방 구성 방식을 채택하게 되었다. 객체용 모델이 오목한 실내 공간 구조를 어떻게 해석하는지 분석하고, 그 결과와 한계를 서술한다.

2D 빈 방 재구성 결과를 기반으로, 실제 방의 3D 형상을 SAM3D [7]를 활용하는 파이프라인을 구성하였다. 사용자가 UI 에서 방 내부 영역을 클릭하면 SAM2 가 방 내부 전체 마스크를 생성한다. 이 마스크를 기반으로 bounding box 영역을 추출하고 최소 512px, 최대 1024px 로 조정한다. 마스크 외 영역은 회색(RGB 128)으로 채워 SAM3D 가 방 내부 구조에만 집중하도록 한다. 색상 부여를 위해 SAM2 방 마스크를 Y 축 기준으로 분할한다. 마스크 상위 50% 픽셀의 평균 RGB 를 벽 색상으로, 이미지 하단 15% 중앙 80% 영역의 평균 RGB 를 바닥 색상으로 추출한다. SAM3D vertex color 를 그대로 사용하되, Y 좌표 하위 20% vertex 에 대해서만 바닥 색상을 덮어씌운다. 메쉬 품질 향상을 위해 trimesh 기반 후처리를 적용한다. 면적 $1\times10^{-8}$ 이하인 degenerate face 를 제거하고, 이음새 구멍을 메운 뒤, Laplacian smoothing(1 회, $\lambda$=0.2)으로 거친 부분을 완화한다. 이후 벽·바닥 vertex 를 각 평면 방향으로 snap 하는 _flatten_room_mesh 를 적용하여 울퉁불퉁함을 교정한다.

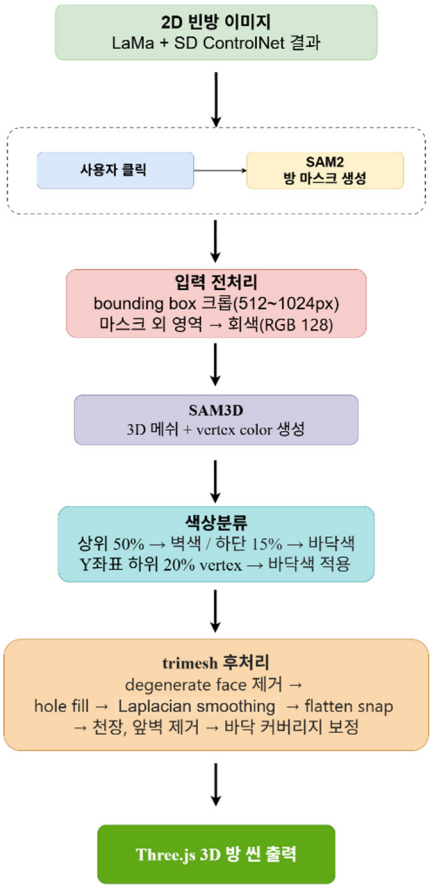
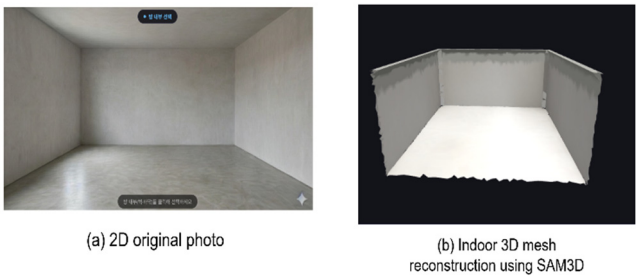



Figure 4. SAM3D-based 3D room mesh generation pipeline.

그러나 생성된 SAM3D 메쉬를 Three.js 로 시각화한 결과 한계가 확인되었다. SAM3D 는 볼록한 단일 객체를 가정한 shape prior 를 학습하였기 때문에, 오목한 구조인 실내 공간에 적용 시 바닥 면 미생성, 벽면 질감 왜곡, 천장 오류 등의 한계가 발생하였다(Figure 5). 이는 단일 객체 도메인과 실내 씬 도메인 간의 근본적인 구조적 차이에서 기인한다. 이러한 한계로 인해 SAM3D 메쉬를 실제 서비스에 적용하는 것은 적합하지 않다고 판단하였으며, 이후 절에서는 Three.js 기반의 절차적 방 구성 방식으로 전환한 내용을 서술한다.



(a) 2D original photo

(b) Indoor 3D mesh reconstruction using SAM3D

Figure 5. 3D mesh generation results for indoor space using SAM3D. Wall and floor geometry is reconstructed relatively accurately in simple rooms, but limitations such as texture distortion and missing floor surfaces occur in spaces with complex elements (furniture, windows, doors, etc.).

### 3.3.2 최종 채택: Three.js 기반 절차적 3D 방 구성

SAM3D 기반 방 재구성의 한계를 바탕으로, 본 연구는 Three.js 기반 절차적 3D 방 구성 방식을 최종 채택하였다. 실제 방 이미지의 상단 50% 영역 평균 RGB 를 벽 색상으로, 하단 15% 중앙 80% 영역 평균 RGB 를 바닥 색상으로 추출하여 기하학적으로 구성된 공간에 적용함으로써, 메쉬 품질에 의존하지 않으면서도 입력 이미지와 시각적으로 일관된 실내 공간을 안정적으로 생성할 수 있다. Three.js 의 BoxGeometry 로 방 구조를 생성하고 각 면에 추출된 색상을 MeshStandardMaterial 로 적용하며, 각 면에는 비가시 충돌 평면(invisible collision plane)을 배치하여 가구 배치 좌표계와 시각적 방 형상의 일치를 보장한다. 이를 통해 사용자는 드래그·드롭 방식으로 가구를 배치하고 실시간으로 인테리어를 시뮬레이션할 수 있다.

### 3.3.3 인터랙티브 가구 배치 충돌 감지 시스템

Three.js 기반 인테리어 시뮬레이터에서 가구를 현실감 있게 배치하기 위해, 드래그 이동 중 발생하는 가구 간 충돌을 실시간으로 감지하고 물리적으로 차단하는 시스템을 구현하였다. 충돌 감지에는 AABB(Axis-Aligned Bounding Box) 기법을 적용하였다. 단, Three.js 의 Box3.setFromObject()는 렌더 루프 실행 전까지 월드 행렬이 갱신되지 않아 부정확한 좌표를 반환하므로, userData 에 저장된 각 축 반경과 group.position 으로 AABB 를 직접 산출한다.

사용자가 가구를 드래그할 때마다 이동 전 좌표를 저장한 뒤 새 위치로 이동시키고, 배치된 모든 바닥 아이템에 대해 X·Z 평면 및 Y 축 높이 범위 교차 여부를 동시에 검사한다. 세 축 조건이 모두 성립할 때만 충돌로 판정하며, 충돌 시 이동 전 좌표로 즉시 복원한다. 벽 부착 아이템(창문, 액자 등) 및 천장 아이템(조명 등)은 충돌 검사에서 제외한다.

또한 테이블 위에 화분을 놓는 등 가구 위 적재 시나리오를 지원하기 위해, 드래그 중 raycaster 가 다른 가구의 상단 면(법선 벡터의 y 성분이 0.5 이상인 면)에 교차하는지 우선 검사한다. 교차가 감지되면 해당 Y 좌표에 아이템 반높이를 더한 위치로 자동 스냅하며 충돌로 처리하지 않고, 교차가 없을 경우 바닥 평면 기준으로 수평 충돌 검사를 수행한다.

Figure 6. Collision detection system — placing a plant on top of a table surface.

## 3.4 가구 3D 복원 파이프라인

### 3.4.1 SAM3D 실내 적용을 위한 입력 최적화

SAM3D 는 단일 클린 객체 이미지를 전제로 설계되어 실내 사진 환경에 직접 적용 시 품질 저하가 발생한다. 본 절에서는 이를 해결하기 위해 수행한 입력 최적화 실험과 그 결과를 기술한다.

**(A) 카테고리별 마스크 전처리(Morphological CLOSE)**

가구 마스크를 SAM3D 에 입력하기 전, 마스크 품질을 높이기 위해 카테고리별 Morphological CLOSE 연산을 적용한다. 소파·침대·테이블 등 일반 가구에는 25×25 커널을 적용한다. 이는 큰 물체일수록 마스크 내부에 큰 구멍이 생기기 쉬워, 넓은 커널로 구멍을 채워야 깔끔한 실루엣을 얻을 수 있기 때문이다.

반면 조명·화분·시계·액자 등 얇거나 작은 물체에는 5×5 커널을 적용한다. 이는 작은 물체에 큰 커널을 적용하면 물체 자체가 뭉개지거나 인접 물체와 마스크가 합쳐지는 문제가 발생하므로, 최소한의 보정으로 원래 형태를 보존하기 위함이다.

**(B) 컨텍스트 패딩 비율이 깊이 추정과 형상 복원에 미치는 트레이드오프 분석**

SAM3D 는 단일 클린 객체 이미지를 전제로 설계되었으므로, 실내 사진에서 가구 영역만 크롭하여 입력하는 것이 자연스러운 방식이다. 그러나 본 연구는 이 과정에서 컨텍스트 패딩 비율이 깊이 추정과 형상 복원에 상반된 영향을 미친다는 사실을 발견하였다.

패딩이 부족하면 MoGe-v2[16]가 공간 구조 정보 없이 깊이를 추정하게 되어 scale calibration 이 불안정해진다. 반면 패딩이 과도하면 주변 가구·구조물의 형태가 SAM3D 의 학습된 3D shape prior 에 혼입되어 형상이 왜곡된다. 예를 들어 화분 복원 시 50% 패딩을 적용하면 주변 배경 요소(창문, 직물 등)의 형태가 SAM3D 의 shape prior 에 혼입되어 복원 메쉬가 오염되는 오류가 발생하였다 (Figure 7 참조).

가구 바운딩 박스 기준 0%, 15%, 50% 패딩을 비교한 실험에서 15%가 두 문제 사이의 균형점임을 확인하였다. 이 트레이드오프는 단일 객체 환경에서 설계된 모델을 실내 멀티 오브젝트 환경에 적용할 때 발생하는 도메인 갭에서 기인하며, SAM3D 원 논문에서 분석된 바 없는 새로운 설계 고려사항이다.

한편 최종 시스템에서는 Mesh Postprocessing 활성화를 통해 메쉬 품질을 보완함으로써, 컨텍스트 패딩 없이 gray background 방식만으로도 안정적인 복원이 가능함을 확인하였다. 따라서 최종 시스템은 처리 단순성과 안정성을 위해 gray background 방식을 채택하였다.

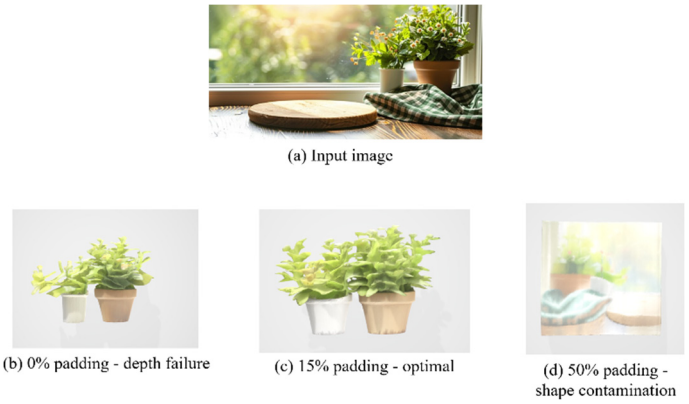



Figure 7. Analysis of context padding ratios and background color for SAM3D input preprocessing. (a) Input image, (b) 0% padding — depth estimation failure due to insufficient spatial information, (c) 15% padding — optimal balance between depth estimation and shape reconstruction, (d) 50% padding — shape contamination due to background interference (plant incorrectly reconstructed as a table).

**(C) 랜덤 시드 의존성 해결을 위한 카테고리별 다중 시드 선택**

SAM3D 는 생성 과정의 확률적 특성으로 인해 동일한 입력에 대해 시드에 따라 생성되는 형태가 크게 달라지는 일관성 문제를 보인다. 생성 기반 3D 복원 모델의 랜덤 시드 의존성 문제는 기존 연구에서 인식된 바 있으며 [14], Best-of-N 샘플링이 일반적인 해결책으로 제안된다. 그러나 기존 방법은 범용 품질 지표를 사용하거나 선택 기준이 없는 경우가 대부분이다.

본 논문은 실내 가구 도메인에서 형상 일관성을 높이기 위한 카테고리별 형상 기반 선택 기준을 제안한다. Stage 1 을 5 회 반복 실행하여 최적 시드를 선택하며, 일반 가구(소파, 의자, 테이블 등)는 생성된 메쉬의 세 축(x, y, z) 중 최소 범위를 최대화하는 기준을 적용해 입체적이고 일관된 형태를

유도한다. 얇거나 작은 물체(조명, 화분, 시계, 액자 등)는 복셀 개수를 최대화하여 구조적 완전성을 우선한다. 이 카테고리별 기준은 범용 품질 평가 모델 없이 기하학적 특성만으로 최적 시드를 자동 선택한다.

시계 객체를 대상으로 단일 시드와 5 회 다중 시드 전략을 각 10 회 실험한 결과, 입체성 지표의 표준편차가 0.162 에서 0.032 로 감소하여 실행 간 형상 일관성이 크게 향상되었다. 다만 Stage 1 을 5 회 반복함에 따라 평균 처리시간이 약 10 초에서 41 초로 증가하는 트레이드오프가 존재한다.

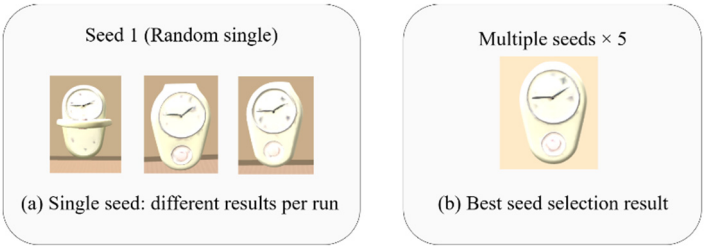



Figure 8. Effect of multi-seed selection. (a) Single-seed results across 3 runs — output varies across runs. (b) Results after selecting the optimal seed from 5 candidates — stably reconstructed mesh.

### 3.4.2 SAM3D 기반 3D 복원 및 자동 형상 완성

가구 3D 복원의 핵심으로 SAM3D [7]를 활용한다. SAM3D 는 대규모 3D 오브젝트 데이터셋으로부터 학습된 3D shape prior 를 활용하여, 단일 이미지만으로 가구의 완전한 3D 형상을 복원한다. 소파의 절반만 보이더라도 모델은 소파의 일반적인 3D 구조 패턴을 통해 가려진 나머지 절반의 팔걸이·등받이·다리 형상을 자동으로 생성한다.

본 시스템의 SAM3D 구현은 MoGe-v2 [16]기반 깊이 추정과 Diffusion Transformer 를 결합하며, 버텍스 컬러 방식(use_vertex_color=True)으로 색상을 표현한다. 입력 해상도는 최대 1024px 이며, 소형 가구의 경우 최소 512px 로 업스케일하여 처리한다. 생성된 메쉬의 면수가 80,000 faces 를 초과하면 Quadric Decimation 으로 경량화하여 렌더링 성능을 유지한다.

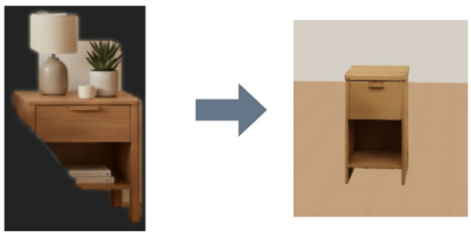



Figure 9. SAM3D occlusion handling: (a) input image of partially occluded furniture, (b) complete 3D reconstruction with occluded geometry restored.

### 3.4.3 버텍스 컬러 기반 가구 색상 재현

본 시스템은 단일 이미지로부터 생성된 3D 메쉬의 색상 재현 방식으로 UV 텍스처 베이킹과 버텍스 컬러를 비교 실험하였으며, 처리 시간 안정성, 렌더링 품질, 아키텍처 적합성을 종합적으로 검토하여 최종 방식을 결정하였다.

**UV 텍스처 베이킹의 한계**

UV 텍스처 베이킹은 메쉬 추출 이후 다시점 렌더링을 통해 색상 정보를 UV 공간에 투영하는 후처리 방식이다.

시스템 구현상 이 과정은 with_texture_baking=True, use_vertex_color=False 옵션으로 활성화되며, 생성된 메쉬의 glb.visual.uv 및 glb.visual.material. baseColorTexture 에서 UV 맵을 추출하는 독립된 렌더링 파이프라인을 통해 수행된다. 이 방식은 두 가지 근본적 한계를 가진다.

첫째, 처리 시간의 불안정성이다. 4 개 객체에 대한 실험 결과, 테이블(복셀 6,383 개, 53 초)과 문(복셀 10,322 개, 58 초)은 유사한 처리 시간을 보였으나, 의자(복셀 26,910 개, 2,550 초)와 침대(복셀 21,826 개, 5,243 초)에서는 처리 시간이 수천 초로 급증하였다. 특히 복셀 수가 더 적은 침대가 의자보다 2 배 이상 긴 처리 시간을 나타냈는데, 이는 UV 베이킹 처리 시간이 복셀 수뿐만 아니라 메쉬 형상 복잡도(UV 아일랜드 구성 방식)에도 크게 좌우되기 때문이다. 곡면과 복수의 구조적 구성 요소를 가진 형상일수록 UV 언래핑 과정에서 생성되는 UV 아일랜드 수가 급증하여 처리 시간이 예측 불가능하게 증가한다.

둘째, 단일 이미지 입력 조건에서의 색상 오염이다. UV 베이킹은 다수의 렌더링 뷰를 평균화하여 색상을 결정하는데, 단일 이미지 입력 환경에서는 미관측 표면(뒷면, 측면)에 대한 실제 색상 정보가 존재하지 않는다. 이로 인해 인접 표면의 색상이 UV 아일랜드 경계를 넘어 번지는 color bleeding 현상, 그리고 기하학적으로 불일치한 가상 뷰들의 평균화로 인한 색상 오염이 발생하였다. (Figure 10(c)).

**버텍스 컬러의 구조적 우위**

버텍스 컬러는 UV 베이킹과 달리 SAM3D 의 학습된 메쉬 디코더(SLatMeshDecoder)가 색상을 직접 회귀하는 방식으로 생성된다.

코드 분석 결과, SparseFeatures2Mesh 의 _calc_layout() 메서드는 네트워크 출력 피처를 SDF(8 채널),

deform(24 채널), color(48 채널, RGB+노멀맵 각 8 개 복셀), weights(21 채널)로 구성하며, 이 중 color 피처가 FlexiCubes 메쉬 추출기에 voxelgrid_colors 로 직접 전달되어 버텍스 색상을 결정한다.

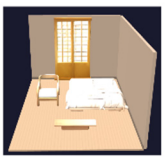
즉 버텍스 색상은 렌더링 투영 알고리즘이 아니라 end-to-end 학습된 신경망의 직접 출력이다. 모델은 대규모 3D 데이터 학습을 통해 단일 이미지에서 관측되지 않는 표면의 색상을 공간적으로 일관되게 추론하는 prior 를 내재화하고 있으며, 이 과정은 좌표 기반 sparse latent space 내에서 수행되므로 뷰 간 불일치 문제가 발생하지 않는다. 또한 최종 출력 단계에서 밝기 보정(×1.15) 및 감마 보정(γ=0.85)을 적용하여 3D 뷰어 환경에서의 색상 재현성을 추가로 개선하였다. 버텍스 컬러 방식은 객체 형상 복잡도와 무관하게 32~40 초의 일정한 처리 시간을 보여, 복잡한 형상에서 최대 134 배의 처리 시간 단축 효과를 나타냈다.

Table 1. Comparison of UV texture baking and vertex color processing time per object (RTX 4090).

| 객체 | 복셀 수 | UV 텍스처 베이킹 | 버텍스 컬러 | 단축 배율 |
|---|---|---|---|---|
| 테이블 | 6,383 | 53 초 | 32 초 | 1.7 배 |
| 문 | 10,322 | 58 초 | 32 초 | 1.8 배 |
| 의자 | 26,910 | 2,550 초 | 40 초 | 64 배 |
| 침대 | 21,826 | 5,243 초 | 39 초 | 134 배 |

처리 시간 안정성, 단일 이미지 입력 조건에서의 렌더링 품질, SAM3D 아키텍처 설계 의도와의 정합성을 종합적으로 고려하여 버텍스 컬러를 최종 색상 재현 방식으로 채택하였다.

(a) Input image



(b) Vertex color result

(c) UV texture baking result

Figure 10. Color rendering comparison: (a) input image, (b) vertex color result — stable and clean color reproduction, (c) texture baking result — color artifacts occur under single-image input conditions.

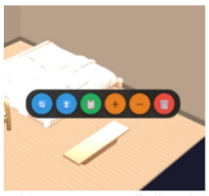
### 3.5 인터랙티브 배치 시스템

복원된 3D 가구 메쉬가 Three.js [9] 기반 인터랙티브 3D 씬에 통합된다. 사용자는 왼쪽 패널에서 생성된 가구 목록을 확인하고, 드래그 앤 드롭 방식으로 3D 방에 가구를 배치할 수 있다. 시스템은 3D 원근 뷰와 2D 조감도 뷰 간 전환을 지원하여 사용자가 다양한 시점에서 배치 결과를 확인할 수 있다.

배치된 가구에 대해 자유 이동, 90° 단위 회전, 크기 조절, 복사, 삭제 조작이 가능하며, 창문·액자 등 벽면 부착 카테고리는 벽면을 자동 인식하여 벽에 부착되는 방식으로 배치된다. 가구 크기는 표준 가구 치수 데이터베이스를 기반으로 자동 스케일링되어 실제 공간 비율에 근사한 배치를 지원한다. 작업 이력 관리를 위해 최대 30 단계의 Undo/Redo 기능을 제공하며, 완성된 인테리어 디자인은 IndexedDB 를 활용하여 저장·불러오기가 가능하다.

(a) 3D perspective view with placed furniture

(b) 2D top-down view



(c) context menu displayed upon furniture selection

Figure 11. Interactive furniture placement interface: (a) 3D perspective view with placed furniture, (b) 2D top-down view, (c) context menu displayed when furniture is clicked.

## 4. 실험 및 결과

### 4.1 실험 환경

실험은 거실, 침실, 서재 등 3 가지 실내 사진 93 장을 대상으로 시스템 전체 파이프라인의 동작을 검증하였으며, 사진별 가구 수는 1 개에서 10 개까지 다양하고 가구 간 occlusion 이 있는 케이스를 포함한다. 시스템은 NVIDIA GeForce RTX 4090 (24GB VRAM) GPU, CUDA 12.1 환경에서 구동되었으며, 백엔드는 Python 3.11 및 FastAPI 로, 프론트엔드는 React 와 Three.js 로 구현하였다.

Table 2. Average processing time per pipeline stage (seconds).

| 단계 | 처리 내용 | 평균 시간(초) |
|---|---|---|
| SAM2 분할 | 인스턴스 마스크 추출 | 0.3 초/클릭 |
| LaMa 인페인팅 | 구조적 배경 복원 | 0.33 초 |
| SD ControlNet | 의미론적 빈 방 완성 | 3.09 초 |
| SAM3D 복원 (×5) | 가구 3D 메쉬 생성 | ~35 초/가구 |

## 4.2 Ablation Study

제안 시스템의 각 구성요소 기여도를 검증하기 위해 실내 이미지 93 장을 대상으로 Table 4 와 같이 비교 실험을 수행하였다. 빈 방 재구성 측면에서 SD 단독 적용은 다수의 사례에서 환각 현상이 관찰되었고, LaMa 단독 적용은 구조적 일관성은 높으나 의미론적 완성도가 낮은 결과를 보였다. 반면 제안하는 LaMa + SD ControlNet(Canny) 2 단계 파이프라인은 구조적 일관성과 시각적 사실감을 모두 고려한 결과를 생성하였다.

이를 정량적으로 검증하기 위해 실제 빈 방 이미지를 GT 로 확보 가능한 10 장을 선정하여 마스크 영역 기준 MSE, PSNR, SSIM, LPIPS 지표를 측정하였다. 실험 결과(Table 3), 제안 방법은 MSE 0.029, PSNR 15.344, SSIM 0.76 으로 세 지표에서 가장 우수한 성능을 보였으며, LPIPS 는 0.232 로 LaMa 단독(0.228) 대비 근소한 차이를 보이나, 나머지 세 지표에서의 전반적 우위를 고려할 때 제안 방법이 구조 보존과 시각적 사실감을 균형 있게 달성함을 확인하였다.

다만 본 연구에서는 비교 대상이 되는 기존 인페인팅 기반 가구 제거 시스템의 공개 소스가 존재하지 않아 타 방법과의 직접적인 정량 비교는 수행하지 못하였으며, 향후 유사 목적 시스템의 소스가 공개되면 이를 활용한 정량적 비교 실험을 진행할 계획이다.

Table 3. Quantitative comparison of inpainting quality on masked regions

| | LaMa | LaMa + SD | LaMa + SD ControlNet(Canny) |
|---|---|---|---|
| MSE ↓ | 0.031 | 0.033 | 0.029 |
| PSNR ↑ | 15.085 | 14.836 | 15.344 |
| SSIM ↑ | 0.757 | 0.742 | 0.76 |
| LPIPS ↓ | 0.228 | 0.231 | 0.232 |

메쉬 품질 측면에서는 다중 시드 전략이 시계 객체를 대상으로 10 회 반복 실험한 결과, 형상 일관성 지표의 표준편차를 0.162 에서 0.032 로 감소시켜 복원 안정성을 향상시켰다(Figure 8).

버텍스 컬러 방식은 여러 종류의 가구에 대해 반복적으로 비교 검증한 결과, 복잡한 형상에서 UV 텍스처 베이킹 대비 최대 134 배의 처리 시간 단축 효과를 달성하였다(Table 1). 또한 가구 3D 복원 측면에서 배경 포함 크롭(0% 패딩)은 깊이 추정 실패가 빈번하였고, 50% 패딩은 형상 오염이 발생하였다. 15% 패딩 + 회색 배경 조합이 가장 안정적인 복원 결과를 보였다.

Table 4. Ablation study results. ("-" indicates not applicable to this evaluation axis; ✓ indicates the proposed configuration.)

| 설정 | 빈 방 품질 | 메쉬 형상 | 색상 품질 | 근거 |
|---|---|---|---|---|
| SD 단독 | 환각 발생 | - | - | Figure 3 (d) |
| LaMa 단독 | 구조 흐림 | - | - | Figure 3 (c) |
| LaMa + SD ControlNet (제안) | ✓ 환각 없음 | - | - | Figure 3 (e) |
| 0% 패딩 (크롭만) | - | 깊이 추정 실패 | - | Figure 7 (b) |
| 50% 패딩 | - | 형상 오염 | - | Figure 7 (d) |
| 15% 패딩 + 회색 배경 (제안) | - | ✓ 최적 균형 | - | Figure 7 (c) |
| 시드 1 회 | - | 형상 일관성 저하 | - | Figure 8 (a) |
| 시드 5 회 카테고리 선택 (제안) | - | ✓ 일관성 향상 | - | Figure 8 (b) |
| UV 텍스처 베이킹 | - | - | 색상오염 | Figure 10 (c) |
| 버텍스 컬러 (제안) | - | - | ✓ 안정적 재현 | Figure 10 (b), Table 1 |

## 4.3 기존 방법과의 정성적 비교

제안 시스템의 가구 3D 복원 품질을 TripoSR [15]와 정성적으로 비교하였다. 동일한 가구 이미지를 각 방법에 입력하여 생성된 3D 모델을 비교한 결과, occlusion 이 있는 케이스에서 제안 방법이 더 완전한 가구 형상을 복원함을

확인하였다. TripoSR 은 가려진 영역의 형상을 추정하지 못하고 가시 영역의 표면만 재구성하는 한계를 보인 반면, 제안한 방법은 학습된 3D shape prior 를 활용하여 occlusion 으로 인해 보이지 않는 부분의 형상을 자동으로 완성한다.

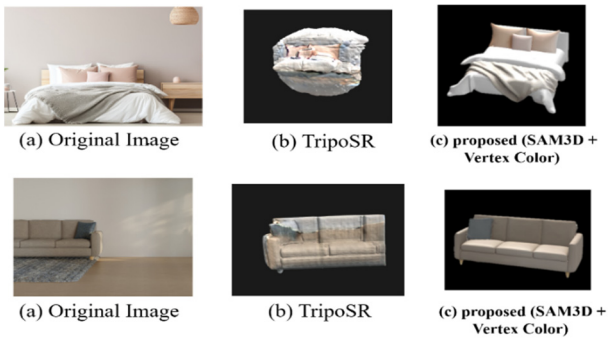



Figure 12. Qualitative comparison: (a) input image, (b) TripoSR, (c) proposed method (SAM3D + vertex color).

## 4.4 기존 상용 앱과의 기능 비교

Table 5 는 제안 시스템과 기존 상용 인테리어 앱들의 기능을 비교한 것이다. 제안 시스템은 실제 보유 가구 활용, 자동 가구 분리, Occlusion 복원 등 기존 앱이 제공하지 못하는 기능을 지원하는 반면, 자체 카탈로그, 실시간 처리, AR 지원 측면에서는 한계가 있다.

Table 5. Feature comparison between the proposed system and existing commercial interior design applications.

| 기능 | IKEA Place | Wayfair | 오늘의집 | 제안 시스템 |
|---|---|---|---|---|
| 실제 보유 가구 활용 | X | X | X | O |
| 자동 가구 분리 | X | X | X | O |
| Occlusion 자동 복원 | X | X | X | O |
| 자동 3D 변환 | X | X | X | O |
| 빈 방 자동 재구성 | X | X | X | O |
| 자체 카탈로그 제공 | O | O | O | X |
| 실시간 처리 | O | O | O | X |
| AR 지원 | O | O | O | X |

## 4.5 사용자 평가

시스템의 실용성을 객관적으로 검증하기 위해 인테리어 비전공자인 20 대 대학생 20 명을 대상으로 사용자 평가를 진행하였다. 참가자들은 사전 설명 없이 시스템을 처음 접한 상태에서 실내 사진 1 장을 입력하여 ① 가구 마스킹, ② 빈 방 재구성 확인, ③ 3D 가구 복원 후 재배치의 3 단계 과제를 순서대로 수행하였다. 과제 완료 후 7 점 리커트 척도 기반 10 개 문항에 응답하였으며, 평가는 개별적으로 진행되어 참가자 간 상호 영향을 배제하였다.

Table 6. User results (7-point Likert scale, 20 participants).

| *카테고리* | 문항 | M | SD |
|---|---|---|---|
| *가구 분할 정확도* | Q1. 마스크 정확도 | 6.10 | 1.25 |
| | Q2. 복잡한 가구 안정적 분리 | 6.35 | 0.93 |
| ***3D 복원 및 색상 품질*** | Q3. Occlusion 3D 자동완성 | 6.45 | 0.60 |
| | Q4. 색상·질감 일관성 | 6.25 | 1.21 |
| | Q5. 빈 방 재구성 자연스러움 | 6.25 | 0.85 |
| *인터페이스 편의성* | Q6. 드래그앤드롭 조작 | 6.60 | 0.75 |
| | Q7. 충돌감지·스냅 기능 | 6.45 | 1.00 |
| | Q8. 뷰 전환·Undo/Redo | 6.50 | 0.61 |
| *전반적 만족도* | Q9. 기존 앱 대비 차별성 | 6.10 | 1.41 |
| | Q10. 전반적 만족·활용 의향 | 6.20 | 1.28 |
| *전체 평균* | - | **6.33** | **1.01** |

## 4.6 사용자 평가 분석

전체 10 개 문항의 평균은 6.33 점(7 점 만점, 90.4%)으로 높은 수준의 사용자 만족도를 보였다. 카테고리별로는 인터페이스 및 기능 편의성(6.52)이 가장 높은 평가를 받았으며, 특히 드래그앤드롭을 통한 3D 가구 배치 조작(Q6, 6.60)이 전 문항 중 가장 높은 점수를 기록하여 웹 기반 인터랙티브 환경의 직관성이 검증되었다.

기술적 핵심 기능인 Occlusion 자동완성(Q3, 6.45)과 뷰 전환·Undo/Redo(Q8, 6.50)는 각각 표준편차 0.60, 0.61 로 전 문항 중 가장 낮은 편차를 보였는데, 이는 가려진 가구를 자동으로 3D 복원하는 기능과 인터페이스 조작 완성도가 평가자 간 이견 없이 일관되게 긍정적으로 인정받았음을 의미한다. 또한 기존 상용 앱 대비 차별성 항목(Q9, 6.10)에서도 높은 점수를 보여, 사용자 보유 가구를 직접

3D 로 변환하여 배치한다는 본 시스템의 핵심 가치가 실용적으로 유효함을 확인하였다.

주관식 응답에서는 처리 속도 지연 및 세그멘테이션 오인식이 개선점으로 지적되었으며, 이는 결론에서 기술한 한계와 일치한다. 반면 "상용화를 기원합니다", "잘 만든 것 같아요" 등의 의견이 수집되어 시스템의 실용적 가치에 대한 높은 기대를 확인하였다.

## 5. 결론 및 한계

본 논문은 단일 실내 사진으로부터 가구를 자동 분리·복원·3D 변환하고 인터랙티브하게 재배치할 수 있는 end-to-end 인테리어 재구성 시스템을 제안하였다.

LaMa 와 SD ControlNet(Canny)의 2 단계 파이프라인으로 환각현상 없는 빈 방 재구성을 달성하였으며, SAM3D 를 실내 가구 복원에 적용하는 과정에서 발견한 배경 요구사항 상충 문제를 컨텍스트 패딩 비율 분석으로 해결하였다. 카테고리별 다중 시드 선택 기준으로 랜덤 시드 의존성 문제를 완화하였으며, UV 텍스처 베이킹과 버텍스 컬러의 비교 실험을 통해 단일 이미지 조건에서 버텍스 컬러가 처리 효율과 색상 일관성 측면에서 적합함을 확인하였다. SAM3D 를 방 전체 메쉬 생성에 탐색적으로 적용하였으나 단일 객체 도메인과 실내 씬 도메인 간의 구조적 차이로 인한 한계를 확인하였으며, 최종적으로 Three.js 기반 절차적 3D 방 구성 방식을 채택하여 안정적인 3D 방 환경을 구현하였다.

본 시스템은 다음과 같은 한계가 있고 이들을 해결하기 위한 향후 연구를 계획 중이다. 첫째, 단일 이미지 기반 복원의 특성상 후면 형상 추정에 부정확성이 존재할 수 있다. 이는 단일 이미지에서 관측되지 않는 표면에 대한 정보가 근본적으로 부재하기 때문이며, 향후 연구에서는 실내 가구 도메인에 특화된 3D shape prior 를 학습한 모델이나 단일 이미지 기반 3D 복원 성능이 향상된 후속 모델로의 교체를 통해 후면 형상 추정의 정확도를 점진적으로 개선해 나갈 계획이다.

둘째, 가구 1 개당 약 35 초의 처리 시간이 소요되어 다수의 가구를 연속으로 생성할 경우 사용자 경험이 저하될 수 있다. 처리 시간 단축을 위한 향후 전략으로는 SAM3D 경량화 모델 적용, 다중 가구 병렬 처리, 복원 결과 캐싱을 통한 동일 가구 재사용 등을 고려할 수 있다.

셋째, 투명 재질이나 고세밀도 형상(예: 등받이 격자 구조)의 가구에서 복원 품질이 저하될 수 있다. 넷째, 방의 실제 스케일 추정이 구현되지 않아 가구 크기 비율이 실제와 다를 수 있다. 다섯째, SAM3D 는 단일 객체 복원에 최적화된 모델이므로 실내 공간 전체에 적용 시 오목한 공간 구조를 볼록한 객체로 잘못 해석하여, 벽·바닥 이음새의 구멍(hole) 아티팩트, 마스크 경계부의 geometry 찢어짐, 복잡한 공간에서의 바닥 면 미생성 등의 문제가 발생하였다. trimesh 후처리 및 평면 snap 으로 일부 완화하였으나 완전한 해결은 어려웠으며, 본 시스템은 이를 Three.js 기반 절차적 3D 방 구성 방식으로 대체하였다. 이는 실내 공간 구조에 특화된 3D 복원 모델의 필요성을 시사한다.

### 감사의 글

본 연구는 2024 년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업 지원을 받아 수행되었음(2024-0-00062)

## References

[1] IKEA, "IKEA Place App," https://www.ikea.com, 2024.
[2] A. Kirillov et al., "Segment Anything," in Proc. ICCV, 2023.
[3] N. Ravi et al., "SAM 2: Segment Anything in Images and Videos," arXiv:2408.00714, 2024.
[4] R. Suvorov et al., "Resolution-robust Large Mask Inpainting with Fourier Convolutions," in Proc. WACV, 2022.
[5] R. Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models," in Proc. CVPR, 2022.
[6] B. Kerbl et al., "3D Gaussian Splatting for Real-Time Radiance Field Rendering," ACM Trans. Graph., vol. 42, no. 4, 2023.
[7] SAM 3D Team et al., "SAM 3D: 3Dfy Anything in Images," arXiv:2511.16624, 2025.
[8] L. Zhang and F. Agrawala, "Adding Conditional Control to Text-to-Image Diffusion Models," in Proc. ICCV, 2023.
[9] R. Cabello et al., "Three.js - JavaScript 3D Library," https://threejs.org, 2024.
[10] A. Yu, V. Ye, M. Tancik, and A. Kanazawa, "pixelNeRF: Neural Radiance Fields from One or Few Images," in Proc. CVPR, 2021.
[11] M. Slavcheva et al., "An Empty Room is All We Want," in Proc. CVPRW, 2024.
[12] A. Dolhasz et al., "Defurnishing with X-Ray Vision," in Proc. CVPRW, 2025.
[13] M. Baradad et al., "Background Prompting for Improved Object Depth," arXiv:2306.05428, 2023.
[14] K. Xu, L. Zhang, and J. Shi, "Good Seed Makes a Good Crop: Discovering Secret Seeds in Text-to-Image Diffusion Models," arXiv:2405.14828, 2024.
[15] T. Wohlgemuth et al., "TripoSR: Fast 3D Object Reconstruction from a Single Image," arXiv, 2024.
[16] R. Wang, S. Xu, Y. Dong, Y. Deng, J. Xiang, Z. Lv, G. Sun, X. Tong, and J. Yang, "MoGe-2: Accurate Monocular Geometry with Metric Scale and Sharp Details," arXiv:2507.02546, 2025.

## 〈저자소개〉

**심 우 현**
- 2022년~현재 조선대학교 컴퓨터공학과 학부생
- 관심분야: 인공지능(AI), 그래픽스, 컴퓨터비전
- https://orcid.org/0009-0001-7126-8915

**위 영 림**
- 2022년~현재 조선대학교 컴퓨터공학과 학부생
- 관심분야: 인공지능(AI), 그래픽스, 컴퓨터비전
- https://orcid.org/0009-0003-6588-5946

**손 예 진**
- 2022년~현재 조선대학교 컴퓨터공학과 학부생
- 관심분야: 인공지능(AI), 그래픽스, 컴퓨터비전
- https://orcid.org/0009-0007-5633-3882

**김 성 기**
- 2009년 서울대학교 박사
- 2009년~2014년 삼성전자
- 2024년~현재 조선대학교 컴퓨터공학과 부교수
- 관심분야: 그래픽스/게임 알고리즘, VR/AR, 인공지능, GPU를 통한 알고리즘 최적화, CPU/GPU를 통한 고성능 컴퓨팅
- https://orcid.org/0000-0002-2664-3632