한국컴퓨터그래픽스학회
Korea Computer Graphics Society Vol. 32, No. 3, P. 47~57

데이터셋 분포 기반 보상을 활용한 강화학습 기반 3차원 실내 가구 배치

조윤식[O,1] 김진모[1,2,*]

한성대학교 일반대학원 정보컴퓨터공학과[1], 한성대학교 컴퓨터공학부[2]
{yunsik.cho, jinmo.kim}@hansung.ac.kr

Reinforcement Learning-Based 3D Indoor Furniture Arrangement Using Dataset Distribution-Based Rewards

Yunsik Cho[O,1] Jinmo Kim[1,2*]

Department of Information and Computer Engineering, Graduate School, Hansung University[1]

Division of Computer Engineering, Hansung University[2]

요 약

본 논문은 방 평면도와 실내 장면 데이터셋의 가구 배치 분포 및 관계로부터 각 가구의 위치와 방향을 결정하는 3차원 실내 가구 배치 문제를 다룬다. 기존 강화학습 기반 실내 가구 배치 연구는 보상 함수의 기준과 제약이 수작업 규칙이나 참조 장면에서 추출된 명시적 정보에 의존한다는 한계를 가진다. 본 연구는 대규모 3차원 장면 데이터셋인 3D-FRONT를 바탕으로 카테고리별 후면과 측면 벽 사이의 거리 분포를 추출하고, 이를 다중 에이전트 강화학습 정책의 관측 및 보상 신호로 활용하는 데이터셋 분포 기반 보상 방법을 제안한다. 일반 에이전트는 카테고리별 벽 거리 분포로부터 유도된 기준 거리를 활용해 가구 배치 정책을 학습하며, 자식 에이전트는 부모 가구를 기준으로 정의된 상대 위치와 방향 목표에 수렴하도록 학습된다. 침실, 거실, 식당 장면에 대하여 가구 OBB의 정사영(orthographic) 탑 뷰로 렌더링한 이미지를 대상으로 FID 실험을 진행한 결과, 제안하는 방법은 ATISS 및 DiffuScene과 유사한 또는 향상된 분포 매칭 성능을 보였다. 이는 데이터셋 통계를 강화학습 보상으로 직접 변환하는 방식이 실내 가구 배치 정책 학습에 효과적으로 활용될 수 있음을 보인다.

Abstract

This paper addresses the problem of 3D indoor furniture arrangement, where the position and orientation of each furniture object are determined from a given room floor plan and the furniture arrangement distributions and relationships observed in indoor scene datasets. Existing reinforcement learning-based indoor furniture arrangement methods are limited in that their reward criteria and constraints rely on hand-crafted rules or explicit information extracted from reference scenes. We propose a dataset distribution-based reward method that extracts category-specific rear and side wall-distance distributions from 3D-FRONT, a large-scale 3D scene dataset, and uses them as observation and reward signals for a multi-agent reinforcement learning policy. General agents learn furniture arrangement policies using reference distances derived from category-specific wall-distance distributions, while child agents are trained to converge to relative position and orientation targets defined with respect to their parent furniture objects. We evaluate bedroom, living room, and dining room scenes using FID on orthographic top-view renderings of furniture OBBs. The results show that the proposed method achieves distribution-matching performance comparable to or better than ATISS and DiffuScene. This demonstrates that directly converting dataset statistics into reinforcement learning rewards can be effective for learning indoor furniture arrangement policies.

키워드: 실내 가구 배치, 강화학습, 데이터셋 분포 기반, 3차원 장면 합성, 다중 에이전트 학습

Keywords: Indoor Furniture Arrangement, Reinforcement Learning, Dataset Distribution-Based Reward, 3D Scene Synthesis, Multi-Agent Learning

*corresponding author: Jinmo Kim / Department of Information and Computer Engineering, Graduate School & Division of Computer Engineering, Hansung University (jinmo.kim@hansung.ac.kr)

Received : 2026.05.27./ Review completed : 1st 2026.06.07. 2nd 2026.06.25./ Accepted : 2026.06.29.
DOI : 10.15701/kcgs.2026.32.3.47
ISSN : 1975-7883(Print)/2383-529X(Online)

1. 서론

3차원 실내 장면 합성은 방 구조, 가구 구성, 객체의 위치·회전·크기 등을 결정하여 현실적인 실내 장면을 생성하는 문제이다. 이 중 실내 가구 배치는 침실이나 거실과 같이 주어진 방의 형태와 배치할 가구 집합으로부터 각 가구의 위치와 방향을 결정하는 하위 문제이다. 실내 장면 합성 및 가구 배치 문제는 가상 환경 구축, 인테리어 디자인 보조, 로봇 시뮬레이션 데이터 생성 등 여러 응용 분야에서 핵심 기반 기술로 다루어져 왔으며, 가상 현실 콘텐츠 제작과 체화 인공지능(Embodied AI) 학습 등으로 응용 범위가 지속적으로 확장되고 있다[1, 2, 3, 4]. 본 연구는 이러한 실내 장면 합성 연구 중에서도, 방 평면도와 가구 배치 분포와 관계에 관한 정보가 주어진 조건에서 각 가구의 평면 위치와 방향을 결정하는 조건부 실내 가구 배치 문제를 다룬다.

초기 실내 장면 합성 연구는 사람이 정의한 규칙과 휴리스틱에 기반하여 가구를 배치하였다[5, 6, 7, 8]. 이러한 규칙 기반 방법은 해석 가능성이 높은 장점이 있으나 새로운 규칙을 수작업으로 추가해야 하는 응용 및 확장성에 한계를 가진다[9]. 이후 3D-FRONT[10], SUNCG[12], Structured3D[13], ScanNet[14]과 같은 대규모 실내 장면 데이터셋이 공개되면서, 데이터셋에 내재된 배치 패턴을 학습하는 장면 합성 연구가 활발히 진행되고 있다[1, 15, 16, 17, 18, 19]. 대규모 장면 데이터셋 기반 장면 합성 연구는 크게 자기회귀 모델, 확산 모델, 언어 모델 기반 방법으로 구분된다. 자기회귀 모델은 트랜스포머 구조를 활용하여 객체를 순차적으로 생성함으로써 객체 간 의존 관계를 모델링하며[1, 17], 확산 모델 기반 접근법은 객체 배치를 노이즈 제거 과정을 응용하여 병렬적으로 생성함으로써 다양성과 사실성을 동시에 확보하였다[15, 16]. 최근에는 대형 언어 모델(LLM, Large Language Model)을 활용하여 자연어 입력으로부터 장면을 합성하는 연구도 활발히 제안되고 있다[18, 2]. 이러한 데이터 기반 장면 합성 연구는 3D-FRONT[10]와 같은 대규모 데이터셋에 내재된 가구 배치 분포를 학습함으로써, 수작업으로 설계된 규칙에 의존하지 않고 사실적이고 다양한 실내 장면을 생성할 수 있다는 점에서 의의가 있다.

이러한 생성 모델 기반 접근과 달리 장면 합성을 순차적 의사결정 문제로 정식화하고 강화학습을 적용하는 연구도 제안되고 있다. SceneMover[20]는 가구 재배치를 위한 충돌 없는 이동 계획 문제에 몬테카를로 트리 탐색(MCTS, Monte Carlo Tree Search)과 심층 강화학습을 결합한 초기 접근을 제시하였다. 이후 가구 배치를 마르코프 결정 과정으로 정식화하고 시뮬레이션 환경에서 강화학습 에이전트를 학습하는 연구가 이어졌으며[21, 22], HAISOR[23]는 단일 에이전트 DQN(Deep Q-Network)과 MCTS를 결합하여 인간 활동을 고려한 가구 배치를 순차적으로 결정하는 방식을 제시하였다. Cho & Kim[24]은 가구를 개별 에이전트로 정의하는 객체 수준 강화학습 프레임워크를 제안하고, 단일 참조 장면에서 추출한 카테고리별 기준값에 기반한 보상 설계를 통해 다양한 방 유형과 가구 조합으로 일반화되는 정책을 학습하였다. 이러한 접근법은 명시적인 의사결정 과정을 모델링한다는 점에서 의의가 있으나, 보상 함수의 기준값과 제약이 인간 디자이너의 수작업 또는 소수 참조 장면에 의존한다는 한계를 가진다. 이는 새로운 가구 카테고리나 방 유형으로 확장할 때마다 보상 함수를 재설계해야 함을 의미하며, 대규모 데이터셋에 내재된 풍부한 배치 정보를 학습 신호로 직접 활용하지 못한다.

본 연구는 이러한 한계를 해결하기 위해, 3차원 장면 데이터셋의 가구 배치 통계 자체가 카테고리별로 서로 다른 배치 경향을 담고 있다는 관찰에서 출발한다. 예를 들어, 침대 카테고리는 벽에 인접한 위치에, 책상 카테고리는 벽에서 떨어진 위치에 빈번하게 배치되는 분포를 보인다. 이러한 통계를 강화학습 정책의 보상 신호로 활용하면, 인간 디자이너가 카테고리별 제약을 수작업으로 명시할 필요 없이 데이터셋 자체가 각 카테고리의 가구가 어디에 배치되어야 하는지를 정책에게 직접 알려주는 구조를 구성할 수 있다. 본 연구는 이러한 직관에 기반하여 3D-FRONT[10]의 ATISS[1] 전처리 결과로부터 가구 카테고리별 벽 거리 분포를 자동으로 추출하고, 이를 강화학습 정책의 관측 및 보상 신호로 변환하는 데이터셋 분포 기반 보상 설계 방법을 제안한다. 구체적으로, 선행 연구[24]의 객체 수준 강화학습 프레임워크를 바탕으로 기준값을 단일 참조 장면이 아닌 데이터셋 전체 분포의 정점으로 정의하는 방식으로 보상 설계를 확장한다. 또한 개별 가구의 분포 적합도뿐 아니라 장면 전체의 일관성을 유도하기 위해, 모든 가구가 카테고리별 분포의 대표 영역에 도달했을 때 부여되는 그룹 보상을 결합한다. 실험적으로는 3D-FRONT[10]의 침실, 거실, 식당 세 방 유형을 대상으로 OBB(Oriented Bounding Box) 기반 탑 뷰 FID(Fréchet Inception Distance) 평가를 수행하고, 제안하는 방법이 ATISS[1] 및 DiffuScene[15]과 유사한 수준의 분포 매칭 성능 달성을 확인한다.

2. 관련연구

2.1 강화학습 기반 실내 가구 배치

강화학습 기반 접근은 실내 가구 배치를 정적인 예측 문제가 아니라, 에이전트가 환경 안에서 행동을 선택하며 장면 배치를 점진적으로 개선해 나가는 순차적 의사결정 문제로 모델링한다. 이러한 방식은 충돌 회피, 접근 가능성, 인간 중심 동선, 부품 가동성 등 명시적인 배치 조건을 보상 함수나 환경 피드백으로 반영할 수 있다는 장점을 가진다. SceneMover[20]는 초기 배치와 목표 배치가 모두 주어진 상황에서, 현재 장면을 목표 장면으로 변환

하기 위한 가구 이동 순서를 계획한다. Di & Yu[21]는 가구 배치를 마르코프 결정 과정(MDP, Markov Decision Process)으로 모델링하고 심층 강화학습을 통해 가구가 배치되지 않은 방에서 가구의 위치와 크기를 직접 결정하는 접근을 제시하였으며, 이후 계층적 강화학습을 도입하여 대규모 방에서의 배치 품질을 개선하는 후속 연구도 진행되었다[22]. 이러한 연구들은 가구가 배치되지 않은 방에서 처음부터 가구 배치 정책을 학습하는 설정을 다룬다는 점에서 SceneMover[20]와 차이가 있으나, 보상 함수가 사전에 정의된 기하학적 제약(충돌, 경계, 인접 관계 등)의 가중합으로 설계된다는 점은 공통적이다. HAISOR[23]는 초기 장면을 입력받아 Dueling Double DQN과 MCTS를 결합한 에이전트를 통해 인간 활동 가용성(Human Affordance)을 고려한 행동 시퀀스를 출력하여 장면 가구 배치를 최적화한다. 보상 함수는 충돌 회피, 자유 공간 확보, 사용자 상호작용성과 같은 인간 중심 기준들의 가중합으로 설계된다. Cho & Kim[24]는 가구를 개별 에이전트로 정의하는 객체 수준 강화학습 프레임워크를 제안하고, 다양한 방 유형과 가구 조합으로 일반화되는 단일 정책을 학습하였다. 보상은 단일 참조 장면에서 추출한 카테고리별 기준값과 가구 현재 상태의 차이로 정의된다. 하지만, 이 방식은 데이터셋 전체에 내재된 카테고리별 배치 분포를 학습 신호로 직접 반영하지는 않는다.

이처럼 강화학습 기반 연구들은 보상설계와 환경 상호작용을 통해 장면 가구 배치 정책을 학습할 수 있음을 보여주었다. 다만 기존 연구들은 수작업으로 정의한 제약이나 소수 참조 장면에서 추출한 기준값으로 보상 함수를 구성해 왔으며, 카테고리별 배치 경향이나 방 유형별 분포 차이와 같은 데이터셋 전체의 통계적 규칙성을 보상 신호로 직접 활용하지는 않았다.

2.2 생성 모델 기반 실내 장면 합성

생성 모델 기반 실내 장면 합성 연구는 대규모 장면 데이터셋으로부터 가구의 카테고리, 위치, 크기, 방향, 형상 사이의 통계적 관계를 학습하는 방향으로 발전해 왔다. 이 계열의 연구는 한 번의 추론 또는 순차적 샘플링을 통해 다양한 실내 장면을 생성할 수 있으며, 3D-FRONT[10]와 같은 대규모 3차원 장면 데이터셋의 등장 이후 정량 평가가 가능한 표준적 실험 환경을 형성해 왔다[1, 15].

2.2.1 자기회귀 및 확산 모델

초기 딥러닝 기반 연구[25]는 합성곱 생성 모델(Convolutional Generative Models)을 이용해 실내 장면을 순차적으로 생성하였다. 이후 ATISS[1]는 트랜스포머 기반 자기회귀 모델을 도입하여, 바닥 평면도와 방 유형을 조건으로 순차적인 가구배치를 바탕으로 장면 합성을 수행하였다. ATISS[1]는 3D-FRONT[10] 기반 실내 장면 합성에서 FID, KL 발산, 분류 정확도 등의 지표를 사용하여 생성 장면의 품질을 평가하였으며, 이후 연구들의 주요 비교 기준으로 활용되고 있다.

확산 모델 계열은 데이터에 노이즈를 점진적으로 추가한 다음 역과정으로 제거(denoising)하며 새로운 데이터를 생성하는 방식을 장면 합성에 응용하는 확산 모델 기반 연구들이 진행되었다. DiffuScene[15]은 정렬되지 않는 객체 속성 집합의 잡음을 제거하여 실내 장면을 합성하는 확산 모델을 설계하였다. MiDiffusion[16]은 객체 속성을 이산-연속 혼합 변수로 다루는 확산 과정을 통해 장면 합성 품질을 추가로 개선하였다. PhyScene[26]은 충돌 회피, 평면도 제약, 도달 가능성과 같은 물리적 조건을 추론 과정에 가이던스 형태로 추가하여 체화 인공지능 환경에서의 상호작용 가능성을 강화하였다. FreeScene[27], CommonScenes[28], LEGO-Net[19]은 텍스트, 장면 그래프, 정돈된 재배치와 같은 조건을 결합하여 장면 생성의 표현 범위를 확장하였다.

이러한 자기회귀 및 확산 모델은 학습 데이터에 나타나는 배치 분포를 효과적으로 모델링하고, 시각적으로 자연스러운 장면을 생성하는 데 강점을 갖지만 충돌 회피, 방 구조 준수, 도달 가능성, 특정 기준 가구에 대한 종속 배치와 같은 명시적 제약은 기본적인 분포 학습 목적만으로 보장되기 어렵기 때문에, 별도의 정규화 항, 추론 시점의 가이던스, 조건부 생성 또는 후처리 과정으로 이를 보완해 왔다[15, 26].

2.2.2 언어 및 시각-언어 모델 기반 접근

최근에는 LLM 또는 시각-언어 모델을 활용하여 실내 장면 생성을 고수준 지시와 연결하려는 연구가 활발히 진행되고 있다. LayoutGPT[18]는 LLM을 시각적 플래너로 활용하여, CSS 스타일 시트 형식의 예시를 바탕으로 텍스트 조건에 맞는 각 가구의 위치, 크기, 방향을 예측한다. 이를 통해 3차원 실내 장면 합성에서 지도학습 기반 방법과 비교 가능한 성능을 달성하였다. Holodeck[2]는 자연어 입력으로부터 방 구조, 객체 목록, 재질, 객체 간 공간 관계를 구성하고, 이를 바탕으로 체화 인공지능을 위한 3차원 장면을 생성한다. InstructScene[29]은 객체 간 의미적 관계를 표현하는 장면 그래프 사전(scene graph prior)과 배치 디코더(layout decoder)를 결합하여 자연어 지시 기반 3차원 실내 장면 합성을 수행하며, LLM 기반 의미 추론과 확산 모델 기반 배치 생성을 통합한 혼합 접근을 제시하였다. 이러한 접근은 사용자의 자연어 지시를 장면 제약으

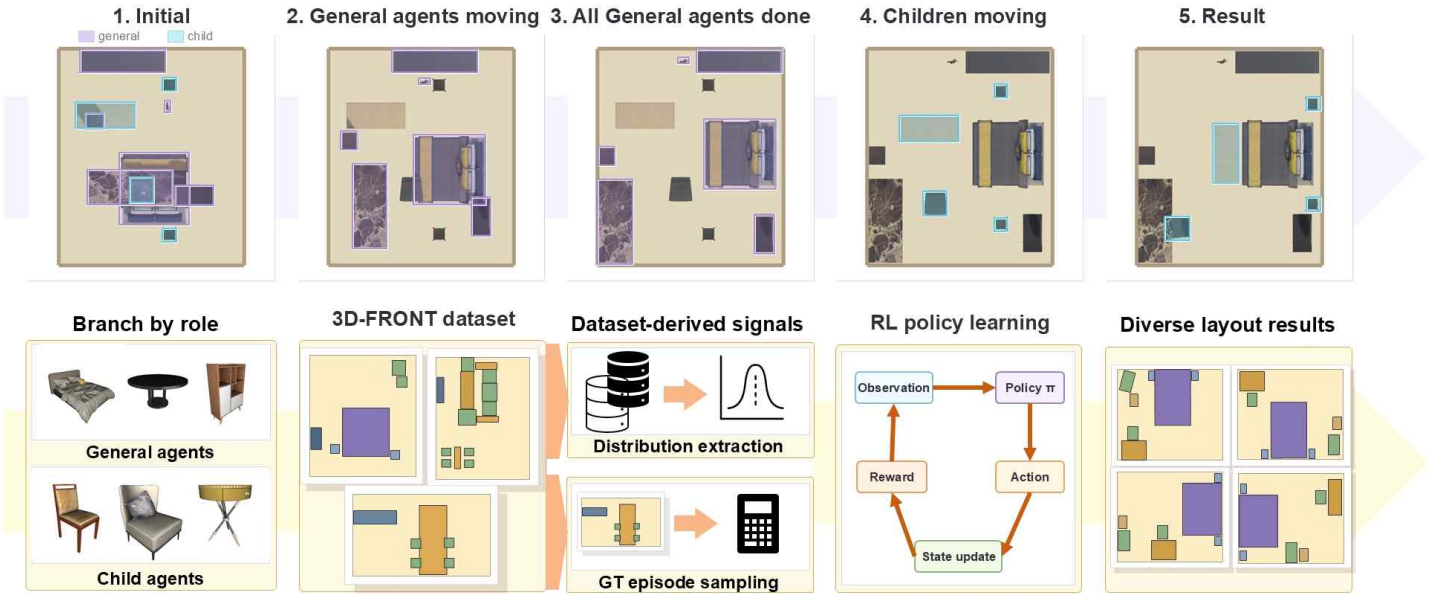

Figure 1. Overview of the proposed two-phase multi-agent furniture arrangement framework. The upper row illustrates the sequential arrangement process, where general agents first move and converge to place major furniture objects, and child agents subsequently move after their corresponding parent/general agents are fixed. The lower row summarizes the dataset-derived learning signals used in the framework. From the 3D-FRONT[10] dataset, category-wise rear and side wall-distance distributions are extracted as dataset-level reward signals for general agents, while sampled GT episodes provide episode-level parent-child relations and relative targets for child agents. These agent-specific signals are incorporated into the two-phase furniture arrangement process to produce diverse indoor layout results.

로 변환할 수 있다는 점에서 유연하지만, 배치 과정이 사전학습된 언어 모델의 추론 능력에 의존하며 장면 배치 정책 자체가 환경과의 상호작용을 통해 독립적인 실행 정책으로 학습되는 것은 아니다.

이처럼 기존 연구들은 데이터 분포 학습 또는 명시적 보상 설계 중 한쪽에 초점을 맞추어 발전해 왔다. 본 연구는 이러한 두 흐름의 한계에 주목하여, 선행 연구[24]를 바탕으로, 가구 배치 제약을 인간 디자이너의 수작업으로 명시하는 대신 3D-FRONT[10] 데이터셋의 통계로부터 보상 신호를 자동으로 구성하는 강화학습 기반 장면 합성 방법을 제안한다. 이는 3차원 실내 장면 합성 연구 현황을 비교 분석한 조사[30]에서 향후 연구 방향으로 제시된 강화학습 기반 접근을 구체화하는 시도에 해당한다. 본 연구를 통해 카테고리별 배치 경향이 데이터셋 통계로부터 직접 학습 신호로 반영되며, 새로운 가구 카테고리나 방 유형으로의 확장이 보상 재설계 없이 가능하다.

3. 데이터셋 분포 기반 보상 설계를 통한 실내 가구 배치

Figure 1은 제안하는 데이터셋 분포 기반 실내 가구 배치 프레임워크의 개요를 보여준다. 상단은 일반 에이전트가 먼저 가구를 배치하고 고정된 상태에서, 자식 에이전트가 배치되는 순차적 과정을 나타낸다. 하단은 3D-FRONT[10] 데이터셋의 원본 장면 배치, 즉 객체의 실제 위치·크기·방향 정보로 구성된 GT(Ground Truth) 가구 배치로부터 카테고리별 벽 거리 분포와 부모-자식 상대 관계를 추출하고, 이를 강화학습 보상 신호로 변환하여 정책 학습에 사용하는 흐름을 요약한다.

3.1 실내 가구 배치 문제 정의

본 연구의 실내 가구 배치 문제는 방의 바닥 평면도와 배치할 가구 집합이 주어졌을 때, 각 가구의 평면 위치와 방향을 결정하는 문제로 정의된다.

문제 설정. 주어진 입력은 방의 바닥 평면도 F (다각형으로 표현)와 배치할 가구의 집합 $O = \{o_1, o_2, \ldots, o_N\}$이며, 각 가구 o_i는 카테고리 c_i와 OBB의 크기(half-extent) $e_i \in R^2$를 속성으로 갖는다. 본 연구의 출력은 각 가구 i에 대한 평면 배치 속성 (x_i, z_i, θ_i)로, OBB 중심 위치 (x_i, z_i)와 방향각(yaw) θ_i를 의미한다. 모든 거리·간격 계산은 탑 뷰 2차원 평면에서 수행하며, 가구 크기는 OBB의 크기로 표기한다.

다중에이전트 의사결정 모델링. 본 연구는 각 가구를 하나의 에이전트로 정의하고, 가구 배치 문제를 다중 에이전트 마르코프 결정 과정으로 모델링한다. 시점 t에서 환경 상태는 방 평면도와 모든 가구의 배치 속성으로 구성되며, 각 에이전트는 자신의 부분 관측으로부터 5분기 평면 이동 (정지, $\pm x$, $\pm z$)과 3분기 방향 회전 (정지, $\pm 90^\circ$)의 이산 행동을 선택한다. 전반적인 모델링 구조

는 선행 연구[24]를 따른다.

부모-자식 역할 분리. 일부 가구는 다른 가구에 종속적으로 배치되는 경향이 있어(예: 식탁-의자, 침대-협탁), 본 연구는 선행 연구[24]의 에이전트 역할 분류 기법을 따라 가구를 일반 에이전트와 자식 에이전트로 분리하여 학습한다. 본 연구에서 활용하는 21개 대표 카테고리는 일반 에이전트 14개와 자식 에이전트 7개로 구성되며, 일반 에이전트는 데이터셋 분포 기반 보상으로, 자식 에이전트는 부모 가구의 위치와 방향을 기준으로 한 상대적 위치 및 자세 관계를 학습한다.

3.2 일반 에이전트의 분포 기반 보상 설계

3D-FRONT[10]는 전문가가 설계한 대규모 합성 실내 장면 데이터셋이다. 본 연구는 ATISS[1]가 제공하는 전처리 결과를 활용한다. ATISS[1] 전처리는 3D-FRONT[10] 원본에서 침실, 거실, 식당 세 방 유형의 장면을 추출하고, 각 장면을 방의 바닥 평면도 다각형과 가구별 카테고리 라벨·OBB 중심 위치·크기·방향의 표준화된 형식으로 변환한다. 본 연구는 이 표준 형식으로부터 일반 에이전트의 학습 신호로 사용할 두 가지 분포-후면 벽 거리 분포 $\hat{P}(d_{rear}\mid c)$와 측면 벽 거리 분포 $\hat{P}(d_{side}\mid c)$-를 추출한다. Figure 2는 가구의 후면 및 측면 벽 거리를 계산하는 방식을 보인다.

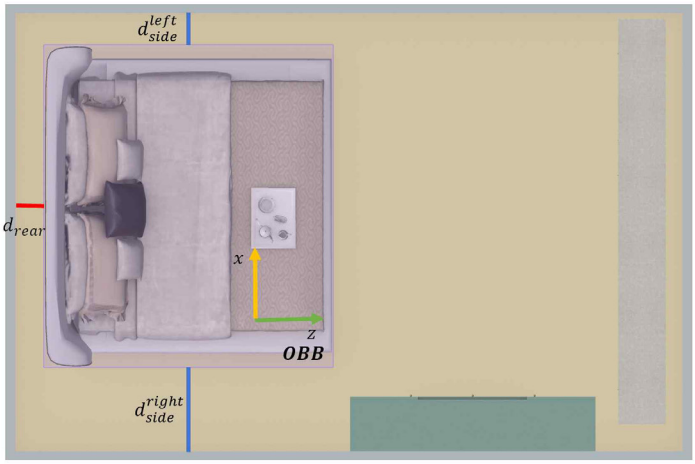

Figure 2. Computation of rear and side wall distances for general agents. Based on the yaw of each furniture object, the $-z$ edge of the OBB is defined as the rear side, and the $\pm x$ edges are defined as the side directions. The rear wall distance d_{rear} is measured from the rear side to the wall, while the side wall distance d_{side} is defined as the smaller distance from the two side(left, right) directions to the wall.

벽 거리 정의. 가구 i의 방향각 θ_i가 주어졌을 때, OBB의 4개 모서리는 가구 로컬 좌표계의 $\pm x$, $\pm z$ 방향과 일치하며, $-z$ 방향 모서리를 후면(rear), $\pm x$ 방향 두 모서리를 측면(side)으로 정의한다. 따라서 방향이 회전하면 후면·측면 모서리가 가리키는 방향도 함께 회전한다. 후면 벽 거리 d_{rear}^{i}은 가구 i의 후면 모서리에서 후면이 향하는 방향으로 측정한 벽까지의 평면 거리로, 측면 벽 거리 d_{side}^{i}는 두 측면 모서리 각각에서 각 측면이 향하는 방향으로 측정한 벽까지의 거리 중 더 작은 값으로 정의한다(Figure 2).

히스토그램 기반 분포추정. 각 카테고리 c에 대해 $K=6$개의 비균등 빈(bin)으로 구성된 1차원 히스토그램을 구축한다. 거릿값을 빈 단위로 이산화하는 이유는 매 step마다 수행되는 보상 조회 비용을 줄이고, 표본 수가 적은 카테고리에서도 분포를 안정적으로 추정하기 위함이다. 빈 경계는 짧은 거리 영역을 조밀하게 두는 휴리스틱에 따라 후면 벽 거리에 대해 $d_{rear}=[0,\ 0.1,\ 0.4,\ 0.8,\ 1.5,\ 2.5,\ 4.0]$, 측면 벽 거리에 대해 $d_{side}=[0,\ 0.2,\ 0.5,\ 1.0,\ 2.0,\ 3.5,\ 5.0]$로 설정한다. 카테고리 c의 조건부 거리 분포는 표본 수 n_c에 대한 빈별 정규화로 추정한다:

$$\hat{P}_{rear}(k\mid c)=h_{real}^{c}[k]/n_c,\ \hat{P}_{side}(k\mid c)=h_{side}^{c}[k]/n_c$$

여기서 $h_{rear}^{c}[k]$와 $h_{side}^{c}[k]$는 각각 카테고리 c의 가구 중 후면 벽 거리와 측면 벽 거리가 k번째 빈에 속하는 표본 수이며, n_c는 카테고리 c의 전체 표본 수이다.

카테고리 통합 및 적용 범위. 분포 추출은 ATISS[1] 전처리 데이터의 *future_category* 라벨 30개를 기준으로 수행한다. 이 중 천장에 고정되어 평면 배치 대상에서 제외되는 Pendant Lamp와 Ceiling Lamp를 제거하고, Single Bed와 Kids Bed를 Bed로 통합하는 등 의미적으로 유사한 카테고리를 대표 카테고리로 통합한다. 이를 통해 최종 21개 대표 카테고리(일반 카테고리 14개, 자식 카테고리 7개)를 도출한다. 분포 기반 보상은 일반 에이전트 14개 카테고리에 대해 적용하며, 각 카테고리 c마다 후면 벽 거리 분포 $\hat{P}(d_{rear}\mid c)$와 측면 벽 거리 분포 $\hat{P}(d_{side}\mid c)$를 구축한다. 따라서 측면 배치 경향 역시 카테고리별 수작업 규칙 없이 데이터 분포로부터 자동으로 반영된다.

관측 벡터. 일반 에이전트의 관측 벡터는 자신의 상태와 분포 기반 차이 정보로 구성된다. 자신의 상태와 RaycastSensor 기반 외부 센서 구성은 선행 연구[24]를 따른다. 본 연구의 차이는 분포 기반 차이 정보에 있으며, 후면 벽 거리 차이 $\Delta d_{rear}=|d_{peak}(c_i)-d_{rear}|/h$와 측면 벽 거리 차이 $\Delta d_{side}=|d_{peak}(c_i)-d_{side}|/h$를 포함한다 ($h$는 방 평면도의 정규화 크기). $d_{peak}(c_i)$는 카테고리 c_i의 분포에서 가장 빈도가 높은 빈의 중간값이며, 기준값을 단일 참조 장면이 아니라 데이터셋 분포의 정점으로 정의하는 점에서 선행 연구[24]와 차이를 갖는다. Figure 3은 추출된 카테고리별 벽 거리 분포의 예시를 보여준다. 침대와 소파처럼

벽에 인접하는 경향이 강한 카테고리는 짧은 후면 벽 거리 구간에 높은 확률이 집중되는 반면, 식탁과 같이 방 중앙에 배치되는 경우가 많은 카테고리는 상대적으로 큰 측면 벽 거리 구간에서도 높은 확률을 보인다. 본 연구는 이러한 카테고리별 분포 차이를 보상 조회 테이블로 사용하여, 에이전트가 데이터셋의 배치 경향을 따르도록 학습한다.

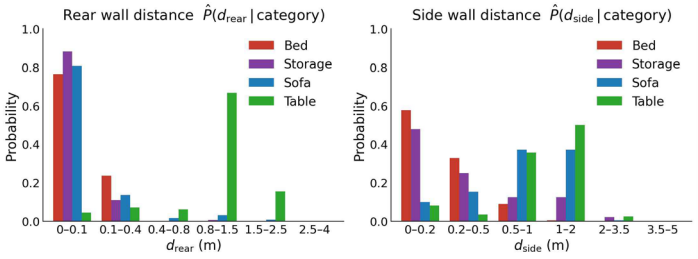

Figure 3. Examples of category-wise rear (left) and side (right) wall-distance distributions extracted from 3D-FRONT[10]. Each category exhibits a distinct wall-proximity tendency, and this study uses these distributions as distribution-based reward signals for the reinforcement learning policy.

보상 함수. 일반 에이전트에 대한 보상은 선행 연구[24]를 따라 각 step에서 벽 거리 보상, 충돌 페널티, 수렴 보너스, step cost를 합산하여 정의하였다. 본 연구에서는 벽 거리 보상의 기준값을 수작업으로 설정하지 않고, 3D-FRONT GT layout에서 추출한 카테고리별 후면 및 측면 벽 거리 분포의 최빈 bin으로부터 자동으로 계산하였다. 다음 식은 일반 에이전트 i의 step t에서의 보상을 나타내며, 편의상 수식에서는 시간 인덱스 t를 생략하였다:

$$r_i = w_{rear}\Delta d_{rear} + w_{side}\Delta d_{side} - w_{col}overlap - stepcost$$

여기서 w는 각 보상 항에 대한 가중치를 의미하고, $overlap$은 일반 에이전트 i의 OBB가 다른 일반 에이전트의 OBB와 겹치는 경우 1, 그렇지 않은 경우 0의 값을 갖는 이진 플래그이다. 또한 $stepcost$은 에이전트가 가능한 한 빠르게 목표 상태에 도달하도록 유도하기 위한 hurry-up 페널티를 의미한다.

3.3 자식 에이전트의 보상 함수 설계

자식 에이전트는 부모 가구의 배치 속성에 상대적으로 학습된다. 본 연구의 자식 에이전트 학습 방식은 선행 연구[24]와 마찬가지로 부모 가구와의 상대적 거리·방향 관계를 학습 신호로 삼되, 목표에 가까울수록 학습 신호가 강해지는 보상 형태에서 본 연구가 차이를 갖는다.

부모-자식 매핑. 3D-FRONT[10] 장면 데이터는 객체 간 부모-자식 관계를 명시하지 않으므로, 본 연구는 에피소드(episode) 초기화 시 부모-자식 매핑을 GT 위치에 기반하여 추론한다. 각 자식 카테고리는 3D-FRONT[10]에서 통계적으로 함께 출현하는 상위 카테고리 집합을 부모 후보군으로 가지며 (예: Nightstand-Bed), 이는 본 연구의 분포 기반 보상 설계와 일관된 데이터 기반 매핑이다. 모든 가능한 (부모-자식 후보) 쌍을 GT 평면 거리 순으로 정렬하여 가까운 쌍부터 자식을 부모에 할당하며, 매핑은 에피소드 동안 고정된다. 부모 수렴 후 자식의 기준 배치 상태는 부모의 현재 배치 상태에 GT 장면에서 추출한 부모-자식 상대 offset을 적용하여 결정된다(Figure 4).

관측 벡터. 자식 에이전트의 관측 벡터는 자신의 상태와 부모에 상대적인 기준 정보로 구성된다. 부모의 상대 정보는 자신의 좌표계에서 본 기준 위치 벡터(v_x, v_z)와 기준 방향 오차($\cos\theta$, $\sin\theta$)를 포함한다. 자식 에이전트는 부모와의 상대 정보를 활용하므로 RaycastSensor를 사용하지 않는다.

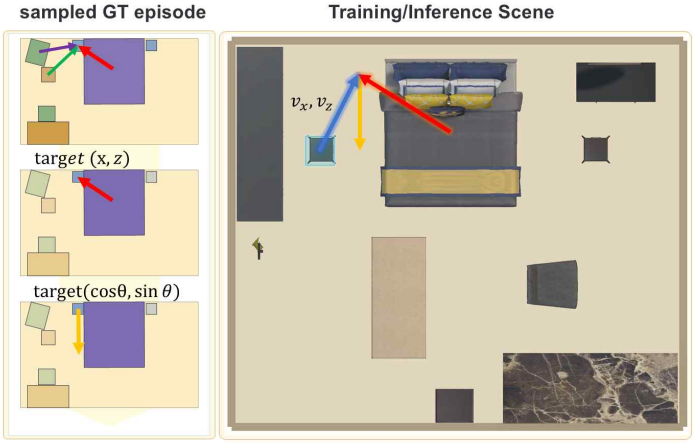

Figure 4. GT-position-based parent-child mapping and dynamic child reference placement. Since 3D-FRONT [10] lacks explicit parent-child annotations, each child object is assigned to the nearest parent candidate based on GT planar distance. The red arrow denotes the GT parent-child relative offset. After the parent converges, this offset and the GT relative yaw are applied to the current parent placement to compute the child reference position and target orientation, indicated by the blue and yellow arrows, respectively.

보상 함수. 선행 연구[24]가 부모-자식 거리·방향 오차의 절댓값에 비례한 페널티를 사용하는 반면, 본 연구는 기준에 가까울수록 보상이 급격히 커지는 3차 감쇠 함수 형태를 채택한다. 따라서 자식 에이전트의 보상은 부모에 대한 기준 위치 및 방향 오차의 3차 감쇠 함수 형태로 정의된다:

$$r_{child} = \frac{\lambda_{gt}}{2}[(\max(0, 1 - d_p/\tau_p))^3 + (\max(0, 1 - |\epsilon_\theta|/\tau_\theta))^3]$$

여기서 $d_p = \|v_x, v_z\|$는 자식의 현재 위치와 기준 위치 사이의 평면 거리, ϵ_θ는 방향 오차이며, τ_p와 τ_θ는 각각 위치·방향 오차의 유효 범위이다. 이는 자식이 기준 근방에 도달했을 때 학습 신호를 강하게 부여하여 학습 목표를 명확히 하기 위함이다. 자식 에이전트는 기준 위치 및 목표 방향에 가까워질수록 큐빅 형태의 보상을 받으며, 기준 거리와 방향 오차가 별도로 설정된 수렴 임계값 이내에 들고 부모 에

이전트가 수렴한 경우 수렴 보너스를 부여받는다.

3.4 학습 절차

본 연구의 정책 학습은 MA-POCA(Multi-Agent POsthumous Credit Assignment)[31] 학습 알고리즘을 사용한다. 학습은 두 단계로 순차 진행한다. MA-POCA는 다중 에이전트 협동 학습을 위한 강화학습 알고리즘이다. 학습 단계에서 팀 전체 상태와 다른 에이전트의 정보를 활용하는 중앙집중식 크리틱(critic)으로 각 행동의 기여도를 평가하고 실행 단계에서는 각 에이전트가 자신의 관찰만으로 독립적인 행동을 수행한다. 또한 셀프 어텐션(self-attention) 기반 구조를 통해 학습 도중 제거되거나 비활성화되는 에이전트의 과거 행동 기여도까지 추정할 수 있다. 먼저, 1단계에서는 자식 카테고리를 학습 환경에서 제외하고 일반 에이전트의 정책만 학습한다. 여기서 하나의 에피소드는 하나의 방 가구 배치를 완성하기까지의 배치 과정을 의미하며, step은 에이전트들이 한 번씩 행동을 선택하고 환경이 이를 반영하여 상태를 갱신하는 단위이다. 매 에피소드는 3D-FRONT[10] 데이터셋에서 무작위로 선택된 방의 평면도와 가구 집합으로 초기화되며, 일반 에이전트들은 방 안의 임의 위치에 배치된다. 매 step 모든 에이전트는 자신의 관측 벡터를 기반으로 행동을 동시에 선택하며, 환경은 방 경계 제약을 적용해 다음 상태로 전이한다. 분포의 정점 영역에 도달한 에이전트는 수렴 상태로 전환되어 고정되며, 모든 에이전트가 수렴 상태에 도달하면 그룹 수렴 보너스와 함께 에피소드가 종료된다. 그 전에 에피소드 최대 step 수에 도달하면 시간 종료로 마무리된다.

2단계에서는 1단계에서 학습된 일반 에이전트의 정책 가중치를 고정한 상태로 자식 에이전트의 정책을 새롭게 학습한다. 자식 카테고리가 학습 환경에 추가되어 부모-자식 매핑(3.3절)에 따라 부모 가구에 할당되며, 일반 에이전트는 학습된 정책으로 추론만 수행한다. 자식 에이전트는 매핑된 부모가 수렴 상태로 고정된 이후 행동을 시작하여 기준 배치 상태로 수렴해 간다. 이와 같이 학습을 두 단계로 나누어 진행하는 이유는 자식 에이전트의 행동이 부모의 수렴 상태에 의존하기 때문이다. 일반 에이전트의 정책을 먼저 안정화한 뒤 자식 정책을 학습하는 분리 학습 방식이 두 정책을 동시에 학습하는 방식보다 학습 안정성에 유리하다.

4. 실험

4.1 강화학습 환경 구축

본 연구의 모든 학습, 추론 및 실험은 Intel Core i9-13900 CPU, NVIDIA GeForce RTX 4070 Ti GPU, 32GB RAM 하드웨어 환경에서 수행되었다. 강화학습은 Unity ML-Agents Toolkit[32] 기반의 Unity 시뮬레이션 환경에서 수행하였다. 정책 네트워크는 일반과 자식 모두 3-layer MLP (hidden units 256) 구조이며, 두 정책은 학습률 3×10^{-4}, 할인율 $\gamma=0.99$, Generalized Advantage Estimation(GAE) $\Lambda=0.95$, clip $\epsilon=0.2$, batch size 1024, buffer size 20480의 동일한 MA-POCA[31] 하이퍼파라미터를 공유한다. 1단계 일반 에이전트 학습은 1.5×10^{6} step, 2단계 자식 에이전트 학습은 2.0×10^{6} step 동안 진행하였다.

데이터셋 구성. 3D-FRONT[10] 데이터셋에서 침실, 거실, 식당 세 카테고리의 방 평면도와 가구 구성을 사용한다. 분포 기반 보상의 히스토그램은 침실 4,041, 거실 987, 식당 900개의 총 5,928방에서 카테고리별로 추출된다. 학습 시 강화학습 환경은 같은 데이터셋에서 임의로 선택된 방으로 에피소드를 초기화하며, 한 에피소드의 최대 step은 1,500으로 설정하였다.

학습 가중치. 일반 에이전트 보상에서 후면 벽 기준 거리 오차, 측면 벽 기준 거리 오차, 충돌 페널티, step cost에 대한 가중치는 각각 $w_{rear}=1.0$, $w_{side}=1.5$, $w_{col}=3.0$, $stepcost=0.1$로 설정하였고, 개별 수렴 보너스와 그룹 수렴 보너스는 각각 2000, 10으로 설정하였다. 충돌 페널티는 일반 에이전트 간 OBB 중첩에 대해서만 적용하며, 자식 에이전트의 보상에는 객체 간 충돌 페널티를 포함하지 않는다. 자식 에이전트 보상에서는 목표 위치와 방향에 가까워질수록 보상이 커지도록 $\lambda_{gt}=1.0$, 위치 보상 유효 범위 $\tau_p=1.5$, 방향 보상 유효 범위 $\tau_\theta=45^\circ$로 설정하였다. 즉, 자식 가구가 기준 위치로부터 1.5 m 이내이고 기준 방향으로부터 45° 이내일 때만 위치 및 방향에 대한 보상이 부여된다. 모든 가구는 학습과 평가에서 OBB로 근사하며, 일반 에이전트의 충돌 판정과 벽 거리 측정은 이 OBB를 기준으로 수행하였다.

4.2 FID 비교 실험

본 연구는 주어진 방 평면도와 가구 집합을 입력으로 생성된 배치 분포가 실제 데이터셋 분포에 얼마나 가까운지를 측정하기 위해 baseline과 FID 비교 실험을 진행하였다. 공정한 비교를 위해 본 연구와 모든 baseline의 출력 장면은 동일한 OBB 기반 탑 뷰 렌더링 결과로 변환한 뒤 FID를 계산하였다. 구체적으로, 생성된 장면은 OBB 기반 카테고리별 색상 큐브로 256×256 해상도의 탑 뷰 이미지로 렌더링된다. 비교 baseline으로는 자기회귀 트랜스포머 기반의 ATISS[1]와 확산 모델 기반의 DiffuScene[15]을 사용하였다. ATISS[1]의 저자들은 학습된 가중치 모델을 제공하지 않으므로, 논문에서 보고된 설정에 따라 각 방 유형별로 200 epoch, batch size 500 조건에서 총 100,000 iteration 동안 직접 학습하였고, DiffuScene[15]은 저자가 제공하는 학습된 가중치 모델을 사용하여 결과를 생성하였다. Figure

5는 동일한 방 데이터에 대해 GT, ATISS[1], DiffuScene[15], 본 연구가 생성한 결과를 FID 평가에 사용한 이미지를 보인다. 본 연구의 결과는 GT의 가구 배치 패턴과 시각적으로 유사한 가구 배치를 보이며, 정량적인 분포 매칭 정확도는 FID 결과로 평가한다. Table 1은 침실, 거실, 식당 세 카테고리에 대한 FID 결과를 보여준다. 이는 본 연구의 강화학습 기반 방식이 명시적인 지도학습이 아닌 분포 기반 보상 신호만으로 학습되었음에도, 학습 분포를 직접 활용하는 생성 모델 기반 baseline과 비교 가능한 결과 도출 가능성을 보인다.

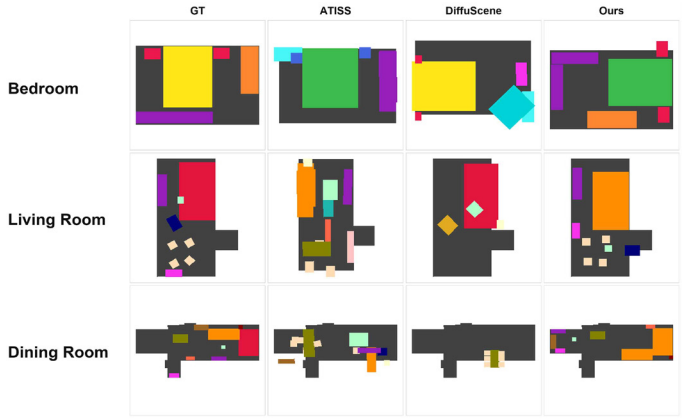

Figure 5. Comparison of OBB-based top-down renderings used for FID evaluation. For one bedroom, one living room, and one dining room scene, the results generated by GT, ATISS[1], DiffuScene[15], and the proposed method are shown.

Table 1. Category-wise FID comparison measured on the official ATISS[1] evaluation dataset (531 rooms). Lower values indicate closer agreement with the GT distribution.

Category	rooms	Avg. Objects	ATISS[1]	DiffuScene[15]	**Ours**
Bedroom	162	4.14	53.43	**16.52**	18.22
Living room	192	9.92	39.40	29.92	**27.92**
Dining room	177	9.48	40.18	34.19	**32.41**

Figure 6. Rendered indoor scenes generated by the proposed method. For bedroom, living room, and dining room scenes, two rooms are shown for each category with both top-down and oblique views.

4.3 정성 결과

OBB 렌더링은 FID 평가용 추상화이므로, 본 절은 본 연구의 정책이 생성한 장면을 3D-FUTURE[11]의 가구 모델로 렌더링하여 시각적 품질을 검토한다. Figure 6은 침실, 거실, 식당 세 카테고리에서 각 2개씩 탑 뷰와 측면 뷰 두 시점으로 본 연구의 결과를 보이며, 각 카테고리에 적합한 가구 배치를 형성한 결과를 정성적으로 확인할 수 있다.

특히, 학습 효율성 측면에서, DiffuScene[15]의 정량 비교에는 저자가 제공하는 학습 완료 가중치를 사용하였으나, 학습 비용을 참고하기 위해 동일 실험 환경에서 직접 초기 학습 구간을 실행하고 전체 학습 시간을 추정하였다. 그 결과 GPU 사용률이 평균 약 70% 수준으로 유지되는 조건에서 100,000 epoch까지 학습할 경우 약 28일이 소요될 것으로 추정되었다. 반면, 본 연구는 Unity ML-Agents[32] 환경의 20배 시뮬레이션 가속과 총 3.5×10^6 step의 학습(약 40분)으로 비교 가능한 FID를 달성하였다. 이는 본 실험 환경에서 제안하는 분포 기반 보상 설계가 생성 모델 학습에 비해 훨씬 적은 학습 시간으로 유사한 분포 매칭 성능에 도달할 수 있음을 보인다.

5. 한계 및 향후 연구

본 연구의 결과는 분포 기반 보상 설계가 강화학습 기반 실내 가구 배치에서 가능성을 보였으나, 다음과 같은 한계가 향후 연구 과제로 남는다.

문제 정의의 한계. 본 연구는 주어진 가구 카테고리 구성에서 각 가구의 평면 배치(위치와 방향)를 학습하므로, 어떤 카테고리의 가구를 생성할지를 추론하는 카테고리 추론은 본 연구의 문제 정의에 포함되지 않는다. 이는 ATISS[1], DiffuScene[15] 등 카테고리 추론을 포함하는 생성 모델 기반 baseline과의 비교 측면에서 본 연구가 다루는 문제 범위가 좁다는 점을 보인다. 카테고리 추론을 통합하는 분포 기반 보상 설계는 향후 연구 방향이다.

환경 제약 및 역할별 충돌 처리의 한계. 본 연구는 방 경계 위반을 환경 전이 단계에서 제한하며, 일반 에이전트에 대해서는 객체 간 OBB 중첩에 대한 충돌 페널티를 보상 함수에 포함한다. 따라서 일반 에이전트는 카테고리별 벽 거리 분포를 따르면서도 다른 일반 가구와의 겹침을 줄이는 방향으로 학습된다. 그러나 자식 에이전트의 보상 함수에는 객체 간 충돌 페널티가 포함되지 않으며, 부모 가구를 기준으로 한 상대 위치와 방향 목표에 수렴하는 것에 초점을 둔다. 이로 인해 일부 사례에서는 자식 가구가 부모 또는 주

변 가구와 겹치거나, 부모 기준 목표 위치를 따라가는 과정에서 방 경계를 다소 벗어난 위치에 수렴할 수 있다. 향후 연구에서는 자식 에이전트에도 충돌 회피와 경계 준수를 학습 신호로 통합하여, 부모-자식 상대 배치와 물리적 타당성을 함께 만족하는 보상 설계로 확장할 필요가 있다.

그룹 보상 Ablation 실험의 부재. 본 연구는 학습 과정에서 MA-POCA[31] 알고리즘을 활용하여 그룹 수렴 보너스의 항목을 보상설계에 포함하였다. 그룹 보상이 학습 결과에 미치는 영향을 확인하기 위해 그룹 보상을 활용하지 않는 PPO 알고리즘과의 비교실험을 진행하였지만, 유의미한 차이를 확인하지 못하였다. 이는 강화학습의 실험 환경에서 명시적 제약이 강하게 설정되어 있고, 사용자 이동 동선 확보, 다양한 카테고리의 서로 다른 에이전트 간 배치 구조학습 등 협동을 수행하여 받게 되는 그룹 보상설계를 활용하지 않았기 때문으로 예상된다. 향후 연구에서는 장면 배치 상태를 평가하는 시스템을 도입하고 이를 그룹 보상으로 설계하여 에이전트 간 협업을 통한 장면 합성 결과를 만들어 낼 수 있을 것으로 기대된다.

6. 결론

본 연구는 강화학습 기반 실내 가구 배치를 위한 데이터셋 분포 기반 보상 설계 방법을 제안하였다. 기존 연구가 수작업으로 정의한 배치 규칙이나 소수 참조 장면에서 추출한 기준값에 의존한 것과 달리, 본 연구는 3D-FRONT[10] 데이터셋의 ATISS[1] 전처리 결과로부터 카테고리별 후면 및 측면 벽 거리 분포를 추출하고, 이를 강화학습 정책의 보상 신호로 활용하였다. 이를 통해 일반 에이전트는 카테고리별 벽 인접 경향을 학습하고, 자식 에이전트는 부모 가구를 기준으로 한 상대 위치 및 방향 관계를 학습하도록 구성하였다. 침실, 거실, 식당 장면을 대상으로 한 OBB 기반 탑 뷰 FID 평가에서 제안 방법은 주어진 가구 집합 조건하에 ATISS[1] 및 DiffuScene[15]과 비교 가능한 분포 매칭 성능을 보였다. 이는 명시적인 지도학습 기반 생성 모델이 아니더라도, 데이터셋에 내재된 배치 통계를 보상 신호로 변환함으로써 실내 가구 배치 정책을 효과적으로 학습할 수 있음을 보여준다. 다만 본 연구는 가구 카테고리와 객체 개수가 주어진 조건부 배치 문제를 다루므로, 카테고리 및 객체 수를 함께 생성하는 전체 장면 합성 문제와는 범위가 다르다. 또한 자식 에이전트의 충돌 처리와 일부 경계 위반 사례는 향후 개선이 필요한 부분이다. 향후 연구에서는 카테고리 추론을 포함하는 통합 정책, 자식 에이전트의 물리적 제약 학습, 그리고 장면 전체의 기능적 일관성을 반영하는 협력 보상 설계로 확장할 수 있다.

감사의 글

본 연구는 한성대학교 학술연구비 지원과제임 (김진모, Jinmo Kim).

References

[1] D. Paschalidou, A. Kar, M. Shugrina, K. Kreis, A. Geiger, and S. Fidler, “ATISS: Autoregressive transformers for indoor scene synthesis,” *Advances in Neural Information Processing Systems*, vol. 34, pp. 12013-12026, 2021.

[2] Y. Yang, F.-Y. Sun, L. Weihs, E. VanderBilt, A. Herrasti, W. Han, J. Wu, N. Haber, R. Krishna, L. Liu, C. Callison-Burch, M. Yatskar, A. Kembhavi, and C. Clark, “Holodeck: Language guided generation of 3D embodied AI environments,” in Proc. *IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 16227-16237, 2024.

[3] M. Savva, A. Kadian, O. Maksymets, Y. Zhao, E. Wijmans, B. Jain, J. Straub, J. Liu, V. Koltun, J. Malik, D. Parikh, and D. Batra, “Habitat: A platform for embodied AI research,” in Proc. *IEEE/CVF Int. Conf. Computer Vision*, pp. 9339-9347, 2019.

[4] E. Kolve, R. Mottaghi, W. Han, E. VanderBilt, L. Weihs, A. Herrasti, M. Deitke, K. Ehsani, D. Gordon, Y. Zhu, A. Kembhavi, A. Gupta, and A. Farhadi, “AI2-THOR: An interactive 3D environment for visual AI,” *arXiv preprint* arXiv:1712.05474, 2017.

[5] L.-F. Yu, S.-K. Yeung, C.-K. Tang, D. Terzopoulos, T. F. Chan, and S. Osher, “Make it home: Automatic optimization of furniture arrangement,” *ACM Transactions on Graphics*, vol. 30, no. 4, article 86, 2011.

[6] T. Weiss, A. Litteneker, N. Duncan, M. Nakada, C. Jiang, L.-F. Yu, and D. Terzopoulos, “Fast and scalable position-based layout synthesis,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 12, pp. 3231-3243, 2019.

[7] P. Kán and H. Kaufmann, “Automated interior design using a genetic algorithm,” in Proc. *23rd ACM Symposium on Virtual Reality Software and Technology*, article 25, pp. 1-10, 2017.

[8] Y. Li, X. Wang, Z. Wu, G. Li, S. Liu, and M. Zhou, “Flexible indoor scene synthesis based on multi-object particle swarm intelligence optimization and user intentions with 3D gesture,” *Computers & Graphics*, vol. 93, pp. 1-12, 2020.

[9] Q. Sun, H. Zhou, W. Zhou, L. Li, and H. Li, “Forest2Seq: Revitalizing order prior for sequential indoor scene synthesis,” *in Computer Vision - ECCV 2024*, Lecture Notes in Computer Science, vol. 15083, pp. 251-268, Springer, 2024.

[10] H. Fu, B. Cai, L. Gao, L.-X. Zhang, J. Wang, C. Li, Q.

Zeng, C. Sun, R. Jia, B. Zhao, and H. Zhang, "3D-FRONT: 3D furnished rooms with layouts and semantics," in Proc. *IEEE/CVF Int. Conf. Computer Vision*, pp. 10933-10942, 2021.

[11] H. Fu, R. Jia, L. Gao, M. Gong, B. Zhao, S. Maybank, and D. Tao, "3D-FUTURE: 3D furniture shape with texture," *International Journal of Computer Vision*, vol. 129, no. 12, pp. 3313-3337, 2021.

[12] S. Song, F. Yu, A. Zeng, A. X. Chang, M. Savva, and T. Funkhouser, "Semantic scene completion from a single depth image," in Proc. *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1746-1754, 2017.

[13] J. Zheng, J. Zhang, J. Li, R. Tang, S. Gao, and Z. Zhou, "Structured3D: A large photo-realistic dataset for structured 3D modeling," in Proc. *European Conf. Computer Vision*, pp. 519-535, 2020.

[14] A. Dai, A. X. Chang, M. Savva, M. Halber, T. Funkhouser, and M. Nießner, "ScanNet: Richly-annotated 3D reconstructions of indoor scenes," in Proc. *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 5828-5839, 2017.

[15] J. Tang, Y. Nie, L. Markhasin, A. Dai, J. Thies, and M. Nießner, "DiffuScene: Denoising diffusion models for generative indoor scene synthesis," in Proc. *IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 20507-20518, 2024.

[16] S. Hu, D. M. Arroyo, S. Debats, F. Manhardt, L. Carlone, and F. Tombari, "Mixed diffusion for 3D indoor scene synthesis," in Proc. *IEEE/CVF Winter Conf. Applications of Computer Vision*, pp. 1262-1272, 2026.

[17] X. Wang, C. Yeshwanth, and M. Nießner, "SceneFormer: Indoor scene generation with transformers," in Proc. *Int. Conf. 3D Vision*, pp. 106-115, 2021.

[18] W. Feng, W. Zhu, T.-J. Fu, V. Jampani, A. Akula, X. He, S. Basu, X. E. Wang, and W. Y. Wang, "LayoutGPT: Compositional visual planning and generation with large language models," *Advances in Neural Information Processing Systems*, vol. 36, pp. 18225-18250, 2023.

[19] Q. A. Wei, S. Ding, J. J. Park, R. Sajnani, A. Poulenard, S. Sridhar, and L. Guibas, "LEGO-Net: Learning regular rearrangements of objects in rooms," in Proc. *IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 19037-19047, 2023.

[20] H. Wang, W. Liang, and L.-F. Yu, "Scene mover: Automatic move planning for scene arrangement by deep reinforcement learning," *ACM Transactions on Graphics*, vol. 39, no. 6, article 233, 2020.

[21] X. Di and P. Yu, "Multi-agent reinforcement learning of 3D furniture layout simulation in indoor graphics scenes," *arXiv preprint* arXiv:2102.09137, 2021.

[22] X. Di and P. Yu, "Hierarchical reinforcement learning for furniture layout in virtual indoor scenes," *arXiv preprint* arXiv:2210.10431, 2022.

[23] J.-M. Sun, J. Yang, K. Mo, Y.-K. Lai, L. J. Guibas, and L. Gao, "Haisor: Human-aware indoor scene optimization via deep reinforcement learning," *ACM Transactions on Graphics*, vol. 43, no. 2, article 15, 2024.

[24] Y. Cho and J. Kim, "From objects to agents: Object-level reinforcement learning for 3D indoor scene synthesis," *under review*, 2026.

[25] D. Ritchie, K. Wang, and Y.-A. Lin, "Fast and flexible indoor scene synthesis via deep convolutional generative models," in Proc. *IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 6182-6190, 2019.

[26] Y. Yang, B. Jia, P. Zhi, and S. Huang, "PhyScene: Physically interactable 3D scene synthesis for embodied AI," in Proc. *IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 16262-16272, 2024.

[27] T. Bai, W. Bai, D. Chen, T. Wu, M. Li, and R. Ma, "FreeScene: Mixed graph diffusion for 3D scene synthesis from free prompts," in Proc. *IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 5893-5903, 2025.

[28] G. Zhai, E. P. Örnek, S.-C. Wu, Y. Di, F. Tombari, N. Navab, and B. Busam, "CommonScenes: Generating commonsense 3D indoor scenes with scene graph diffusion," *Advances in Neural Information Processing Systems*, vol. 36, 2023.

[29] C. Lin and Y. Mu, "InstructScene: Instruction-driven 3D indoor scene synthesis with semantic graph prior," in Proc. *Int. Conf. Learning Representations*, 2024.

[30] Y. Cho, J. Lee, S.-H. Cho, and J. Kim, "Research status for 3D indoor scene synthesis applications: A comparative analysis of interactive, optimization, and deep learning methods," *Journal of the Korea Computer Graphics Society*, vol. 31, no. 2, pp. 13-25, 2025.

[31] A. Cohen, E. Teng, V.-P. Berges, R.-P. Dong, H. Henry, M. Mattar, A. Zook, and S. Ganguly, "On the use and misuse of absorbing states in multi-agent reinforcement learning," in *AAAI-22 Workshop on Reinforcement Learning in Games*, 2022.

[32] A. Juliani, V.-P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar, and D. Lange, "Unity: A general platform for intelligent agents," *arXiv preprint* arXiv:1809.02627, 2020.

〈저자소개〉

조 윤 식

- 2021년 한성대학교 컴퓨터공학부 학사
- 2022년 한성대학교 일반대학원 컴퓨터공학과 석사
- 2022년~현재 한성대학교 일반대학원 정보컴퓨터공학과 박사과정
- 관심분야: 가상현실, 증강현실, 컴퓨터그래픽스, HCI
- https://orcid.org/0000-0003-2118-0904

김 진 모

- 2006년 동국대학교 멀티미디어학과 학사
- 2008년 동국대학교 영상대학원 멀티미디어학과 석사
- 2012년 동국대학교 영상대학원 멀티미디어학과 박사
- 2012년~2014년 동국대학교 영상문화콘텐츠연구원 전임연구원
- 2014년~2019년 부산가톨릭대학교 소프트웨어학과 조교수
- 2019년~현재 한성대학교 컴퓨터공학부 부교수
- 관심분야: 컴퓨터그래픽스, 가상현실, 증강현실, 메타버스, 게임 공학 등
- https://orcid.org/0000-0002-1663-9306