

보행자와 장소 간 관계를 고려한 고밀도 군중 흐름 시뮬레이션

장예정[°]

권태수^{*}

한양대학교

willlido32@hanyang.ac.kr, taesoo@hanyang.ac.kr

Dense Crowd Flow Simulation Considering Pedestrian-Venue Relationships

Yejeong Jang[°]

Taesoo Kwon^{*}

Hanyang University

요약

본 연구는 군중이 밀집된 환경에서 장소의 구조가 보행자 흐름에 미치는 영향을 반영한 고밀도 군중 시뮬레이션 프레임워크를 제안한다. 기존의 군중 시뮬레이션 및 다중 에이전트 궤적 예측 연구는 주로 보행자 간 상호작용에 초점을 맞추어 왔다. 그러나 군중 밀도가 높아질수록 장소 고유의 구조와 특성이 군중의 이동 방향과 밀집도 변화에 큰 영향을 미친다. 본 연구에서는 이러한 특징적인 공간 요소를 장소 노드로 정의하고, 각 노드에 의미 정보를 부여하여 보행자 간 관계뿐 아니라 보행자와 장소, 장소 간 관계까지 함께 모델링하였다. 또한 그래프 신경망(graph neural network)을 활용하여 관계의 종류와 특성에 따른 상호작용 가중치를 반영함으로써, 장소 구조에 따른 군중 흐름 변화를 효과적으로 학습하였다. 제안 방법은 벡터화된 궤적 표현을 기반으로 보행자 궤적을 예측하며, 위치 예측 오차뿐 아니라 충돌 발생률과 장소 노드 주변의 군중 밀집도 등 시뮬레이션 관점의 지표도 함께 분석하였다. 이를 통해 본 연구는 장소의 의미와 구조를 반영한 고밀도 군중 흐름 모델링의 가능성을 제시하고자 한다.

Abstract

We propose a dense crowd simulation framework that reflects the influence of venue structures on pedestrian flows in crowded environments. Existing studies on crowd simulation and multi-agent trajectory prediction have mainly focused on interactions among pedestrians. However, as crowd density increases, the inherent structure and characteristics of each venue have a significant impact on pedestrian movement directions and changes in crowd density. In this study, such distinctive spatial elements are defined as venue nodes, and semantic information is assigned to each node to model not only pedestrian-pedestrian relationships but also pedestrian-venue and venue-venue relationships. In addition, a graph neural network is used to reflect interaction weights according to the types and characteristics of relationships, enabling the model to effectively learn changes in crowd flow caused by venue structures. The proposed method predicts pedestrian trajectories based on vectorized trajectory representations and analyzes not only position prediction errors but also simulation-oriented metrics such as collision rate and crowd density around venue nodes. Through this approach, this study aims to demonstrate the potential of dense crowd flow modeling that incorporates the semantic meaning and structure of venues.

키워드: 군중 시뮬레이션, 컴퓨터 애니메이션, 다중 에이전트 궤적 예측, 그래프 신경망

Keywords: Crowd Simulation, Computer Animation, Multi-Agent Trajectory Prediction, Graph Neural Networks

*corresponding author: Taesoo Kwon / Hanyang University (taesoo@hanyang.ac.kr)

1 서론

군중 시뮬레이션은 다수의 보행자 또는 캐릭터가 하나의 공간 내에서 이동하며 상호작용하는 과정을 생성하는 컴퓨터 애니메이션의 핵심 분야 중 하나이다. 게임, 영화, 가상 현실(VR), 도시 환경 분석, 로봇 내비게이션, 안전 시뮬레이션 등 다양한 분야에 활용될 수 있어 중요하다. 특히 군중이 밀집된 환경에서는 보행자 간 거리가 매우 짧고, 개별 보행자의 움직임이 주변 사람 및 공간 구조, 방향성 등 다양한 요인의 영향을 받게 되는데, 이를 고려하여 수많은 군중을 자연스럽게 생성하기 위해서는 단순히 목표 지점을 향한 이동이나 충돌 회피뿐만 아니라, 장소의 의미론적 구조와 군중의 흐름 간 관계를 함께 고려할 필요가 있다.

군중 속 충돌 회피를 고려하기 위해 다양한 방법이 제안되었는데, Social Force Model [1], ORCA [2] 등 맥락 기반 힘을 고려하거나 속도 기반 충돌 회피를 고려하는 방법을 통해 다수의 보행자의 움직임을 효율적으로 생성할 수 있게 되었다. 이러한 방법들은 계산 효율이 높고 안정적이지만, 에이전트들의 행동이 사전에 정의된 규칙에 크게 의존하여 실제 군중 속에서 나타나는 다양한 장소 기반의 흐름 변화를 반영하기 어렵고, 고밀도 군중 시뮬레이션의 경우처럼 수많은 보행자가 등장할 경우 계산량이 크게 늘어난다는 단점이 있다. 최근에는 보다 복잡한 군중 움직임을 학습하기 위해 GREIL-Crowds [3]과 같이 강화학습 그리고 데이터 기반 접근법이 많이 활용되고 있으며, 3차원 환경에서의 밀집된 군중 시뮬레이션에서 발생하는 캐릭터 간 충돌을 보정하는 연구 [4] 또한 제안되었다. 이러한 연구 흐름은 고밀도 군중 시뮬레이션이 단순한 보행자의 경로 예측 문제가 아니라, 전체적인 군중의 행동 및 공간 구조, 충돌 안정성, 그리고 자연스러운 움직임이 함께 고려되어야 한다는 것을 보여준다.

한편, 다중 에이전트 궤적 예측 분야에서는 보행자 간 상호작용을 학습하기 위한 그래프 기반 모델 [5], [6]과 Transformer 기반 모델 [7]이 활발하게 연구되어 왔다. 특히 VectorNet [8]과 HiVT [9]는 에이전트의 궤적과 장면 정보를 벡터 또는 polyline 단위로 표현하고, 지역적 맥락과 전역적인 상호작용을 계층적으로 모델링함으로써 복잡한 다중 에이전트 장면을 효율적으로 처리할 수 있음을 보였다. 이러한 벡터화 표현은 고밀도 군중 시뮬레이션의 복잡한 보행자의 이동과 주변 관계를 구조적으로 표현하기 위한 유용한 접근법이 될 수 있다. 그러나 기존의 궤적 예측 연구를 고밀도 군중 시뮬레이션에 그대로 적용하기에는 여전히 한계가 있다. 대부분의 모델은 보행자 또는 차량과 같은 동적 에이전트를 중심으로 상호작용을 학습하며, 장소 자체가 가진 의미론적 역할을 명시적으로 모델링하지 않는다. 예를 들어 출입구는 군중의 유입과 유출을 유도하고, 병목 구간은 정체와 압축 현상이 발생하며, 통로에서는 방향성 있는 흐름이 형성된다. 이처럼 장소는 군중 행동을 결정하는 중요한 요소이지만, 기존 모델에서는 이러한 장소의 기능이 독립적인 학습 단위로 충분히 표현되지 않았다.

본 연구는 이러한 문제를 해결하기 위해 장소 노드 주석 표기

와 그래프 뉴럴 네트워크(Graph Neural Network)를 결합한 고밀도 군중 시뮬레이션 프레임워크를 제안한다. 제안 방법은 보행자만을 그래프의 노드로 보는 것이 아니라, 군중 흐름에 영향을 주는 각각의 장소 요소를 노드로 정의하고, 각 장소 노드에 의미론적 주석을 기입한다. 출입구, 병목 구간, 통로 등 모든 장소는 각각 다른 기능을 가진 장소 노드로 표현되며, 모델은 이를 통해 보행자-보행자 관계뿐만 아니라, 보행자-장소 관계 및 장소-장소 관계를 함께 학습할 수 있다.

또한 본 연구는 모든 간선을 동일하게 처리하지 않고, 간선의 종류와 의미에 따라 서로 다른 상호작용 가중치를 반영한다. 군중이 밀집된 상황에서 가까운 보행자와의 관계, 같은 흐름을 공유하는 보행자와의 관계, 병목 구간으로 접근하는 보행자와 장소 간의 관계, 출구를 향해 이동하는 보행자와 출구 지점의 관계는 서로 다른 의미를 가진다. 따라서 본 연구에서는 그래프 뉴럴 네트워크를 통해 관계별 가중치를 다르게 학습하고, 이를 통해 장소 구조가 군중 흐름에 미치는 영향을 보다 명확하게 표현하고자 한다.

본 연구의 주요 기여 항목은 다음과 같다:

- 고밀도 군중 시뮬레이션에서 장소를 단순한 배경이 아니라 군중의 행동에 영향을 주는 의미론적 그래프 노드로 모델링하는 장소 노드 주석 기입 구조를 제안한다.
- 출입구, 병목 구간, 통로 등 장소의 기능을 주석으로 정의하고, 보행자 궤적과 결합하여 장소의 의미와 군중의 흐름 사이의 관계를 학습할 수 있도록 하였다. 또한 보행자-보행자, 보행자-장소, 장소-장소 관계를 그래프 구조로 표현하고, 간선 유형 또는 특징에 따라 서로 다른 상호작용 가중치를 반영하는 그래프 뉴럴 네트워크 구조를 설계하였다.
- 기존 벡터 기반 궤적 표현 연구와 지역적/전역적 상호작용 모델링의 장점을 장소 노드 주석 및 그래프 구조와 결합하였다. 이를 통해 본 연구는 보행자 간 상호작용뿐만 아니라 보행자-장소, 장소-장소 관계를 함께 고려하는, 장소를 반영한 고밀도 군중 시뮬레이션 프레임워크를 구성하였다.
- 실제 보행자 궤적 데이터셋을 활용하여, 위치 예측 오차뿐만 아니라 충돌 발생률, 장소별 밀집도 변화, 장소 노드 주변의 흐름 변화 등 시뮬레이션 관점의 평가를 수행할 수 있는 실험 파이프라인을 구성하였다.

2 관련 연구

2.1 군중 시뮬레이션

군중 시뮬레이션은 다수의 보행자가 공간 내에서 이동하며 상호작용하는 과정을 생성하는 문제로, 전통적으로는 사람이 정의한 규칙이나 수식 모델을 기반으로 발전해 왔다. Social Force Model [1]은 보행자가 목표 방향으로 이동하려는 힘과 주변 보행자 및

장애물로부터 멀어지려는 반발력을 사용하여 군중의 움직임을 표현하였다. 이러한 접근법은 병목 구간, 집단 흐름과 같은 군중과 관련된 현상을 직관적으로 설명할 수 있다는 장점이 있다. 한편 대표적인 속도 기반 모델인 ORCA [2]는 이와 달리, 에이전트가 선택할 수 있는 속도 공간에서 충돌을 회피할 수 있는 속도 조건을 계산한다. ORCA는 다중 에이전트가 충돌 회피 책임을 나누어 가진다는 가정 아래, 각 에이전트가 충돌 없는 속도를 선택하도록 하여 실시간 시뮬레이션 환경에서도 안정적이다. 하지만 앞선 방법들은 에이전트들의 행동이 사전에 정의된 규칙에 크게 의존하거나 군중 시뮬레이션처럼 수많은 에이전트가 등장할 경우 계산량이 크게 늘어난다는 단점이 있다.

최근에는 실제 군중 데이터나 예시 행동을 활용하여 보행자 행동을 학습하는 데이터 기반 군중 시뮬레이션 연구도 활발히 진행되고 있다. GREIL-Crowds [3]는 참조 군중 데이터에 기반한 강화학습 프레임워크를 통해 목표 추종, 그룹 형성, 배회 행동 등 다양한 군중 행동을 학습하고자 하였다. 이는 군중 행동을 사전에 모두 규칙으로 정의하지 않고도 예시 데이터로부터 복잡한 행동 패턴을 학습할 수 있음을 보여준다. 또한 밀집 군중 애니메이션에서는 궤적 생성 이후의 캐릭터 간 충돌과 3차원 애니메이션 품질도 중요한 문제로 다루어지는데, 고밀도 3차원 군중 환경에서 캐릭터 간 충돌을 보정하여, 입력 군중 움직임의 전체 흐름을 유지하면서도 시각적으로 자연스러운 3차원 군중 애니메이션을 생성하는 방법 또한 제안되었다 [4]. 이러한 연구는 고밀도 군중 시뮬레이션이 단순한 경로 생성 문제가 아니라, 충돌 안정성, 군중 흐름, 애니메이션 품질을 함께 고려해야 하는 문제임을 보여준다. 한편 다수 캐릭터 간의 밀접한 상호작용을 생성하기 위한 상호작용 합성 연구도 제안되었으며, 캐릭터 간 상호작용의 자연스러움에 초점을 두고 있다 [16].

이러한 기존 군중 시뮬레이션 연구는 주로 보행자 간 충돌 회피 또는 개별 에이전트의 행동 생성에 초점을 두는 반면, 실제 밀집 군중 환경에서는 출입구, 통로, 병목 구간, 정체 가능 구역과 같은 장소 구조가 군중 흐름에 직접적인 영향을 미친다. 본 연구는 이러한 장소 요소를 단순한 배경으로 처리하지 않고, 장소 노드 주석 기입을 통해 그래프 안의 의미론적 노드로 표현한다는 점에서 기존 연구와 차별화된다.

2.2 다중 에이전트 궤적 예측

보행자 궤적 예측 연구는 과거 위치 관측을 바탕으로 미래 이동을 예측하는 문제를 다루며, 군중 시뮬레이션과 밀접하게 연결된다. 초기 딥러닝 기반 연구인 Social LSTM [14]은 각 보행자의 이동을 LSTM으로 인코딩하고, 주변 보행자의 은닉 상태를 사회적 풀링(pooling)으로 결합하여 보행자 간 상호작용을 모델링하였다. 이후 Social GAN [15]은 보행자의 미래 이동이 하나로 결정되지 않는다는 점에 주목하여, 생성적 적대 신경망을 통해 사회적으로 타당한 다중 미래 궤적을 생성하고자 하였다.

그래프 기반 접근 방법 또한 널리 사용되었다. Social-STGCNN

[6]은 보행자를 시공간 그래프의 노드로 구성하고, 그래프 합성곱 신경망을 통해 보행자 간 상호작용을 학습하였다. Trajectron++ [5]는 에이전트 간 관계를 그래프 구조로 표현하고, 다양한 에이전트의 동역학적 특성과 장면 정보를 함께 고려하여 미래 궤적을 예측하였다. 이러한 연구들은 보행자 궤적 예측에서 개별 보행자의 과거 이동뿐만 아니라, 주변 에이전트와의 관계를 함께 학습하는 것이 중요함을 보여준다.

최근에는 Transformer 기반 구조를 활용하여 시간적 관계와 에이전트 간 관계를 동시에 모델링하려는 연구가 증가하였다. AgentFormer [7]은 시간 축과 에이전트 축의 관계를 함께 고려하는, 에이전트가 반영된 Transformer 구조를 제안하여 다중 에이전트 예측에서 attention 기반 모델의 가능성을 보였다. QCNet [10]은 질의 중심 방식으로 궤적 예측 문제를 재구성하여, 반복적인 좌표 정규화와 중복 계산을 줄이고 효율적인 다중 에이전트 예측을 수행하고자 하였다.

이러한 궤적 예측 연구들은 보행자 또는 차량과 같은 동적 에이전트 간 상호작용을 효과적으로 모델링한다. 그러나 대부분의 방법은 출입구, 병목 구간, 통로와 같은 장소 요소가 군중 흐름에 미치는 영향을 명시적인 관계로 표현하지 않는다. 고밀도 군중 시뮬레이션에서는 특정 장소가 보행자의 유입, 정체, 분산, 회피 행동을 유도할 수 있으므로, 본 연구에서는 보행자-보행자 관계뿐만 아니라 보행자-장소 관계를 함께 고려하는 그래프 구조를 설계한다.

2.3 벡터화 및 그래프 기반 장면 표현

최근 다중 에이전트 궤적 예측 연구에서는 장면을 이미지 형태로 변환하는 래스터화(reastered) 표현 방법 대신, 지도 요소와 에이전트 궤적을 벡터 또는 폴리선(polyline) 단위로 표현하는 vectorized representation(이하 벡터화 표현)이 주목받고 있다. VectorNet [8]은 도로 지도와 에이전트 움직임을 폴리선 형태의 벡터로 표현하고, 이를 계층적 그래프 구조로 처리하여 장면의 구조적 정보를 보존하고자 하였다. 이러한 표현은 이미지 기반 입력보다 불필요한 공간 낭비를 줄이고, 궤적과 장면 요소의 기하학적 관계를 직접적으로 다룰 수 있다는 장점이 있다.

HiVT [9]는 다중 에이전트 모션 예측 문제를 지역적인 맥락 추출과 전역적인 상호작용 모델링으로 분리하여 처리하였다. 각 에이전트의 지역적인 이동 문맥을 먼저 인코딩한 뒤, 에이전트 간 전역적인 상호작용을 모델링함으로써 많은 수의 에이전트가 존재하는 장면에서도 효율적인 예측이 가능하도록 하였다. 이러한 분리 구조는 각 보행자의 개별 이동 패턴과 주변 보행자와의 상호작용 모두 중요하기 때문에 밀집된 군중 환경에서도 유용하다.

HiVT는 다중 에이전트 궤적 예측에서 지역적 맥락과 전역적인 상호작용을 효율적으로 분리하여 모델링한다는 장점이 있다. 그러나 HiVT의 주요 목적은 에이전트의 과거 궤적과 주변 상호작용을 기반으로 미래 위치를 예측하는 것이며, 출입구, 병목 구간, 통로와 같은 장소 요소를 독립적인 의미 노드로 정의하거나,

모델은 관측 구간 T_{obs} 동안의 보행자 궤적을 입력으로 받아, 이후 T_{pred} 동안의 미래 궤적을 예측한다.

$$\hat{Y}_1^{1:T_{pred}} = f_{\theta}(X_1^{1:T_{obs}}, G_V), \quad (2)$$

여기서 $X_1^{1:T_{obs}}$ 는 보행자들의 과거 궤적, $\hat{Y}_1^{1:T_{pred}}$ 는 예측된 미래 궤적, G_V 는 장소 노드와 간선 정보를 포함하는 그래프를 의미한다.

본 연구에서는 보행자의 절대 위치뿐만 아니라 시간에 따른 이동 벡터를 함께 사용한다.

$$v_i^t = p_i^{t+1} - p_i^t, \quad (3)$$

이동 벡터는 보행자의 이동 방향과 속도 변화 및 보행자 궤적을 벡터로 표현하여 구성하는 데 사용된다. 이러한 표현 방식은 기존 VectorNet [8] 및 HiVT [9]에서 사용된 방식과 유사하게, 복잡한 다중 에이전트 장면을 구조적으로 표현하는 데 유리하다.

3.2 장소 노드에 대한 주석 기입

본 연구의 핵심은 장소를 단순한 배경 정보가 아니라 군중 흐름에 영향을 주는 의미적 노드로 모델링하는 것이다. 이를 위해 장면 내 주요 장소 요소를 장소 노드로 정의하고, 각 노드에 의미론적 주석을 기입한다. 각 장소 노드 v_k 는 다음과 같이 표현된다:

$$v_k = (c_k, l_k). \quad (4)$$

여기서 c_k 는 장소 노드의 공간적 위치 또는 중심 좌표를 의미하며, l_k 는 해당 장소의 의미적 레이블을 의미한다. 예를 들어 l_k 는 입구, 출구, 병목 구간, 대기 구역과 같은 장소 유형으로 정의될 수 있다.

장소 노드 주석 기입 과정은 모델이 특정 장소 주변에서 반복적으로 나타나는 군중 행동을 학습할 수 있도록 한다. 예를 들어 출입구 주변에서는 유입과 유출 흐름이 강하게 나타날 수 있으며, 병목 구간 주변에서는 보행자의 속도 감소와 밀집도 증가가 발생할 수 있다. 본 연구는 이러한 장소별 기능을 그래프 노드의 속성으로 포함하여, 보행자 움직임과 장소 의미 사이의 관계를 학습한다.

3.3 장소가 반영된 그래프 구성

전처리된 보행자 궤적과 장소 노드 주석 기입을 바탕으로, 본 연구는 보행자 노드와 장소 노드를 함께 포함하는 장소가 반영된 그래프를 구성한다.

$$G_V = (N_A, N_V, E), \quad (5)$$

여기서 N_A 는 보행자 노드 집합, N_V 는 장소 노드 집합, E 는 노드 간 간선 집합을 의미한다. 본 연구의 그래프는 보행자-보행자

관계뿐만 아니라 보행자-장소 관계와 장소-장소 관계를 함께 포함한다.

보행자-보행자 간선은 두 보행자 사이의 상대 위치, 거리, 이동 방향 차이를 나타낸다. 이는 기존 다중 에이전트 궤적 예측 모델에서 다루어온 객체 간 상호작용과 관련이 있다 [5], [6], [7]. 보행자-장소 간선은 보행자와 주변 장소 노드 사이의 관계를 나타낸다. 예를 들어 보행자가 출입구에 가까워지는 경우, 해당 간선은 유입 또는 유출 흐름과 관련된 정보를 가진다. 장소-장소 간선은 장소 노드 사이의 연결 관계를 나타내며, 통로와 출입구, 통로와 병목 구간처럼 장면 내 흐름 구조를 표현하는 데 사용된다.

이러한 그래프 구성은 기존 보행자 중심 그래프와 달리, 장소 구조를 군중 흐름의 일부로 직접 포함한다는 점에서 차별성을 가진다.

3.4 지역 인코더

지역 인코더는 각 보행자의 관측 궤적에서 개별적인 이동 패턴을 추출한다. 보행자 a_i 의 입력 특징은 관측 구간 동안의 위치 변화와 이동 벡터로 구성되며, 지역 인코더는 이를 통해 보행자의 이동 방향, 속도 변화, 정지 또는 회전 경향과 같은 지역적 움직임 특징들을 생성한다.

$$h_i^{local} = Enc_{local}(X_i, V_i), \quad (6)$$

여기서 X_i 는 보행자 a_i 의 관측 위치 시퀀스, V_i 는 이동 벡터 시퀀스, h_i^{local} 은 보행자별 지역적 특징을 의미한다.

지역 인코더의 역할은 주변 보행자나 장소와의 관계를 고려하기 전에, 각 보행자의 기본적인 이동 의도를 안정적으로 표현하는 것이다. 예를 들어 일정한 방향으로 이동하는 보행자, 속도가 감소하는 보행자, 방향을 전환하는 보행자는 서로 다른 지역적 특징을 갖게 된다. 이러한 정보는 이후 전역 인코더와 상호작용 모듈에서 주변 관계와 결합된다.

3.5 전역 인코더

전역 인코더는 장면 전체의 상호작용 정보를 학습한다. 기존 HiVT [9]는 지역적 맥락과 전역적인 상호작용을 분리하여 다중 에이전트 예측을 효율적으로 수행하였다. 본 연구 역시 이러한 지역적/전역적 분리 구조를 활용하되, 전역적인 상호작용의 대상을 보행자에 한정하지 않고 장소 노드까지 확장한다.

전역 인코더는 보행자 노드와 장소 노드의 특징을 함께 입력으로 받아, 장면 단위의 상호작용 표현을 생성한다.

$$H^{global} = Enc_{global}(H_A^{local}, H_V), \quad (7)$$

여기서 H_A^{local} 은 지역 인코더에서 얻은 보행자 특징 집합이고, H_V 는 장소 노드 주석 기입 과정으로부터 얻은 장소 특징 집합이다. 이를 통해 모델은 보행자 간 상호작용뿐만 아니라, 보행자와

장소 사이의 관계를 함께 고려할 수 있다.

예를 들어 병목 구간 근처에 여러 보행자가 접근하는 경우, 전역 인코더는 단순히 보행자 사이의 거리만 보는 것이 아니라 해당 보행자들이 병목 장소 노드와 어떤 관계를 갖는지 함께 학습한다. 따라서 장면의 장소 구조가 군중 흐름에 미치는 영향을 모델 내부 표현에 반영할 수 있다.

3.6 간선이 반영된 상호작용 모듈

군중이 밀집된 환경에서는 모든 간선이 동일한 의미를 갖지 않는다. 즉, 보행자-보행자 간선, 보행자-장소 간선, 장소-장소 간선은 서로 다른 의미를 가진다. 또한 같은 보행자-장소 간선이라 하더라도, 출입구와의 관계인지 병목 구간과의 관계인지에 따라 군중 흐름에 미치는 영향이 달라진다.

이를 반영하기 위해 본 연구는 간선이 반영된 상호작용 모듈을 사용한다. 각 간선 e_{ij} 는 다음과 같은 특징을 가진다:

$$e_{ij} = (type_{ij}, d_{ij}, r_{ij}), \quad (8)$$

여기서 $type_{ij}$ 는 간선의 종류, d_{ij} 는 두 노드 사이의 거리, r_{ij} 는 상대 위치 또는 방향 정보를 의미한다. 간선이 반영된 모듈은 간선 특징을 이용하여 메시지 전달 과정에서 관계별 가중치를 다르게 계산한다.

$$h'_i = \phi \left(h_i, \sum_{j \in \mathcal{N}(i)} \alpha_{ij} \cdot \psi(h_j, e_{ij}) \right), \quad (9)$$

여기서 h_i 는 노드 i 의 특징, $\mathcal{N}(i)$ 는 이웃 노드 집합, ψ 는 이웃 노드 특징과 간선 특징을 결합하는 함수, α_{ij} 는 간선별 상호작용 가중치, ϕ 는 노드 업데이트 함수이다.

이 구조를 통해 모델은 관계의 종류에 따라 서로 다른 영향력을 반영할 수 있다. 예를 들어 가까운 보행자와의 간선은 충돌 회피와 관련된 정보를 강하게 전달할 수 있고, 출입구 장소 노드와의 간선은 이동 방향과 유출 가능성에 영향을 줄 수 있다. 병목 장소 노드와의 간선은 속도 감소나 밀집도 증가와 관련된 정보를 전달할 수 있다.

3.7 예측 디코더

Edge-aware 상호작용을 통해 업데이트된 보행자 노드 특징은 미래 궤적 예측에 사용된다. 각 보행자 노드의 최종 특징을 h_i^* 라 할 때, 예측 디코더는 이를 기반으로 미래 이동량을 예측한다.

$$\Delta \hat{p}_i^t = Dec(h_i^*, t), \quad (10)$$

예측된 이동량은 마지막 관측 위치를 기준으로 누적되어 미래 위치로 변환된다.

$$p_i^{T_{obs}} + \sum_{\tau=1}^t \Delta \hat{p}_i^\tau, \quad (11)$$

이처럼 미래 위치를 직접 예측하는 대신 이동량을 예측하면, 장면 좌표계에 대한 의존성을 줄이고 보행자의 이동 변화 자체를 학습하기 쉽다. 또한 장소가 반영된 그래프에서 얻은 장소 기반 상호작용 정보가 디코더에 반영되므로, 특정 장소 구조에 따른 흐름 변화를 예측할 수 있다.

3.8 학습 목적 함수

모델은 예측 궤적과 정답 궤적 사이의 위치 오차를 최소화하도록 학습된다. 기본 궤적 손실 함수는 다음과 같이 정의한다:

$$\mathcal{L}_{traj} = \frac{1}{NT_{pred}} \sum_{i=1}^N \sum_{t=1}^{T_{pred}} \left\| \hat{p}_i^{T_{obs}+t} - p_i^{T_{obs}+t} \right\|_2^2. \quad (12)$$

또한 고밀도 군중 시뮬레이션에서는 보행자 간 비정상적인 충돌이나 중첩을 줄이는 것이 중요하므로, 필요에 따라 충돌 감지 손실 함수를 추가할 수 있다.

$$\mathcal{L}_{col} = \frac{1}{T_{pred}} \sum_{t=1}^{T_{pred}} \sum_{i=1}^N \sum_{j=i+1}^N \max(0, \delta - d_{ij}^t), \quad (13)$$

여기서 d_{ij}^t 는 시간 t 에서 두 보행자 사이의 예측 거리이며, δ 는 충돌로 판단하는 거리 임계값이다.

최종 손실 함수는 다음과 같이 정의된다:

$$\mathcal{L}_{total} = \mathcal{L}_{traj} + \lambda_{col} \mathcal{L}_{col}. \quad (14)$$

여기서 λ_{col} 은 궤적 예측 정확도와 충돌 회피 사이의 균형을 조절하는 가중치이다.

4 실험 및 분석

4.1 실험 데이터셋

본 연구에서는 보행자 궤적과 장면 정보를 포함하는 2차원 이미지 또는 영상 평면 데이터셋을 사용하여 제안 방법을 평가한다. 사용 데이터셋은 SDD [11], JRDB [12], SiT [13]으로 구성된다. 각 데이터셋은 서로 다른 관측 환경과 장면 특성을 가지므로, 제안하는 데이터 전처리 및 장소가 반영된 그래프 구성이 다양한 환경에서도 적용 가능한지 확인하는 데 적합하다.

SDD [11]는 드론 기반 조감 시점에서 수집된 캠퍼스 환경 데이터셋으로, 보행자뿐만 아니라 자전거, 차량 등 다양한 이동 객체를 포함한다. 본 연구에서는 이 중 보행자 클래스를 사용하였으

4.3 구현 세부사항

제안 모델은 보행자 궤적을 벡터화 표현으로 변환한 뒤, 보행자 노드와 장소 노드를 함께 포함하는 장소가 반영된 그래프를 구성한다. 이후 지역 인코더는 각 보행자의 개별 이동 패턴을 인코딩하고, 전역 인코더 및 그래프 신경망은 보행자-보행자, 보행자-장소, 장소-장소 관계를 함께 학습한다. 최종 예측 디코더는 업데이트된 보행자 노드 특징을 기반으로 미래 궤적을 예측한다.

모든 실험에서는 동일한 관측 길이와 예측 길이를 사용하였으며, 각 데이터셋은 동일한 전처리 파이프라인을 거쳐 학습에 사용되었다. 기존 연구와의 공정한 비교를 위해 가능한 경우 동일한 양의 학습/테스트 분할 및 동일한 입력 궤적 길이를 사용하였다. 모든 실험은 Nvidia GeForce RTX 3080 Ti 환경에서 수행하였다.

4.4 평가 지표

본 연구에서는 예측 정확도와 시뮬레이션 관점의 품질을 함께 평가하였다. 먼저 위치 기반 지표로 ADE와 FDE를 사용하였다. ADE는 전체 예측 구간에서 예측 위치와 정답 위치 사이의 평균 오차를 측정하며, FDE는 마지막 예측 시점에서의 위치 오차를 측정하는 지표이다.

$$ADE = \frac{1}{NT_{pred}} \sum_{i=1}^N \sum_{t=1}^{T_{pred}} |\hat{p}_i^t - p_i^t|_2, \quad (17)$$

$$FDE = \frac{1}{N} \sum_{i=1}^N |\hat{p}_i^{T_{pred}} - p_i^{T_{pred}}|_2, \quad (18)$$

고밀도 군중 시뮬레이션에서는 보행자 간 충돌 가능성을 평가하기 위해 충돌 발생률을 사용한다. 두 보행자 사이의 거리가 임계 값 δ 보다 작을 경우 충돌 또는 비정상적 근접으로 간주하였다.

$$C_{ij}^t = \begin{cases} 1, & \text{if } |\hat{p}_i^t - \hat{p}_j^t|_2 < \delta, \\ 0, & \text{otherwise} \end{cases}, \quad (19)$$

또한 본 연구에서는 장소 노드 주변의 흐름 변화를 분석하기 위해 장소 수준의 산술 지표를 추가로 사용하였다. 특정 장소 노드 v_k 주변 반경 R 안에 존재하는 보행자 수를 기반으로 장소의 밀도를 계산할 수 있다.

$$D_k^t = \frac{1}{\pi R_k^2} \sum_{i=1}^{N_t} 1[|p_i^t - c_k|_2 < R_k], \quad (20)$$

또한 특정 장소 노드 주변을 통과하는 보행자의 수를 흐름 계산 지표로 정의할 수 있다. 이는 출입구, 통로, 병목 구간과 같이

흐름 변화가 중요한 장소에서 모델이 실제 군중 흐름을 얼마나 잘 반영하는지 평가하는 데 사용된다.

4.5 실험 결과 분석

Table 1은 SDD 데이터셋을 본 연구에서 제안한 모델에 넣었을 때, 위치 예측 성능 및 충돌 발생률을 비교한 결과이다. 같은 SDD 데이터셋의 bookstore 속 video 0 장면에서 Case 1은 과거 궤적을 8 프레임만큼, Case 2는 16 프레임만큼 참고한 경우이다. 그리고 Case 3은 SDD 데이터셋의 deathCircle 속 video 0 장면에서 과거 궤적을 16 프레임만큼 참고한 결과이다.

실험	ADE↓	FDE↓	충돌 발생률↓	속도 MAE↓	경로 길이 MAE↓
Case 1	0.652	1.238	0.00375	0.251	0.841
Case 2	0.533	0.988	0.00436	0.240	0.647
Case 3	0.882	1.664	0.00226	0.401	1.021

Table 1: SDD 데이터 장면별 궤적 예측 결과 분석

같은 장소에서 과거의 궤적을 더 많이 참고할 수록 좋은 성능을 보였으며, 같은 길이의 과거 궤적을 참고한 경우 보다 복잡한 장면 구조일수록 전반적인 궤적 예측 지표는 좋지 않았지만, 충돌 발생률이 낮아지는 결과를 보여주었다.

Table 2는 SDD 데이터셋의 각 장면을 대상으로 장소 노드 주석을 기입하지 않았을 때와 기입했을 때, ADE와 FDE의 차이를 비교한 결과이다. 낮을수록 결과 예측의 정확도가 높다고 할 수 있다. 아래 비교에 사용된 장면은 순서대로 각각 bookstore, death-Circle, gates, hyang, nexus의 video 0이다.

Scene	ADE w/o	ADE w/	Δ ADE	FDE w/o	FDE w/	Δ FDE
Scene 1	0.6811	0.6870	+0.0059	1.3115	1.3080	-0.0036
Scene 2	1.1023	1.0875	-0.0148	2.0494	2.0345	-0.0149
Scene 3	1.2205	1.2378	+0.0173	2.3103	2.3475	+0.0372
Scene 4	1.1422	1.1474	+0.0052	2.1344	2.1691	+0.0348
Scene 5	0.8376	0.8512	+0.0137	1.5523	1.5924	+0.0401
Mean	0.9967	1.0022	+0.0054	1.8716	1.8903	+0.0187

Table 2: 장소 노드 주석 기입 유무에 따른 궤적 예측 정확도 분석

Table 3은 SDD 데이터셋의 각 장면을 대상으로 장소 노드 주석을 기입하지 않았을 때와 기입했을 때, 속도 MAE와 경로 MAE의 차이를 비교한 결과이다. 낮을수록 좋은 모션 일관성을 보인다고 할 수 있다.

Scene	속도 w/o	속도 w/	Δ 속도	경로 w/o	경로 w/	Δ 경로
Scene 1	0.2830	0.2700	-0.0130	0.9959	0.9235	-0.0723
Scene 2	0.4322	0.4249	-0.0073	1.3784	1.2801	-0.0983
Scene 3	0.4904	0.4919	+0.0015	1.4726	1.5274	+0.0549
Scene 4	0.5089	0.5291	+0.0202	1.7406	1.8559	+0.1154
Scene 5	0.3713	0.3616	-0.0097	1.0696	1.0881	+0.0185
Mean	0.4172	0.4155	-0.0017	1.3314	1.3350	+0.0036

Table 3: 장소 노드 주석 기입 유무에 따른 모션 일관성 분석

장소 노드별 주석 기입은 평균 ADE/FDE 개선으로 이어지지 않았으나, 원형 교차나 명확한 이동 구조가 있는 장면에서는

움직임에 대한 일관성과 일부 위치 지표를 개선하였다. 특히, 원형 도로 장면의 경우 장소 노드 주석을 기입했을 때 전반적으로 모든 지표가 개선되었다.

5 결론 및 향후 연구

본 연구에서는 군중이 밀집된 환경의 시뮬레이션을 위해 장소의 의미적 구조와 보행자 간 상호작용을 함께 고려하는 장소가 반영된 그래프 기반 프레임워크를 제안하였다. 기존의 군중 시뮬레이션 및 다중 에이전트 궤적 예측 연구는 주로 보행자 간 충돌 회피나 동적 에이전트 사이의 상호작용을 중심으로 발전해 왔다. 그러나 실제 밀집 군중 환경에서는 출입구, 통로, 병목 구간, 정체 가능 구역과 같은 장소 요소가 군중의 이동 방향, 밀집도 변화, 흐름 분산에 직접적인 영향을 미친다. 본 연구는 이러한 장소 요소를 장소 노드로 정의하고, 각 노드에 의미론적 주석을 기입함으로써 장소 구조와 군중 흐름 사이의 관계를 모델링하고자 하였다.

제안 방법의 주요 기여는 다음과 같다. 첫째, 장소를 단순한 배경 정보가 아니라 군중 행동에 영향을 주는 의미적 그래프 노드로 표현하는 장소 노드 기입 구조를 제안하였다. 둘째, 보행자-보행자 관계뿐만 아니라 보행자-장소, 장소-장소 관계를 함께 포함하는 장소를 반영한 그래프를 구성하였다. 셋째, 간선 유형 및 특징에 따라 서로 다른 상호작용 가중치를 반영하는 그래프 신경망 구조를 통해 장소별 군중 흐름의 차이를 학습할 수 있도록 하였다. 마지막으로, 기존 벡터 기반 궤적 표현의 장점을 활용하면서도 이를 장소 의미 기반 그래프 구조로 확장하여 고밀도 군중 시뮬레이션에 적합한 궤적 예측 및 흐름 분석 파이프라인을 구성하였다. 실험 결과, 장소 노드 주석이 모든 장면에서 일관된 ADE/FDE 개선으로 이어지지는 않았으나, 원형 교차 구조와 같이 장소의 흐름 방향성이 명확한 장면에서는 위치 예측 오차와 모션 일관성 지표가 함께 개선되는 경향을 확인하였다. 이는 제안한 장소를 반영한 그래프가 단순 위치 예측 정확도를 향상 향상시키기보다는, 장소 구조가 군중 흐름에 강하게 작용하는 장면에서 더 효과적으로 작동할 가능성을 보여준다.

그러나 본 연구에는 몇 가지 한계가 존재한다. 첫째, 장소 노드 주석의 품질이 모델 성능에 영향을 줄 수 있다. 장소 노드를 수동으로 정의하는 경우 주석의 기준이 주관적일 수 있으며, 자동으로 추출하는 경우 장면의 실제 의미 구조를 충분히 반영하지 못할 수 있다. 둘째, 본 연구는 주로 2차원 보행자 궤적을 기반으로 하므로 보행자의 자세, 시선, 몸 방향, 보행 주기와 같은 세부 행동 정보는 직접적으로 고려하지 않는다. 셋째, 서로 다른 데이터셋은 관측 시점, 좌표계, 주석 방식, 장면 밀도에서 차이가 있기 때문에, 공통 전처리 과정을 적용하더라도 데이터셋 간 차이가 완전히 제거되지는 않는다.

향후 연구에서는 장소 노드 주석 과정을 보다 자동화하고 정교화하는 방향으로 확장할 수 있다. 예를 들어 궤적 밀도, 이동 방향 분포, 정체 발생 빈도 등의 통계적 특성을 활용하여 출입구, 병목

구간, 통로와 같은 장소 노드를 자동으로 추출할 수 있다. 또한 2차원 궤적 정보뿐 아니라 보행자의 이동 방향, 속도 변화, 자세 정보 등을 함께 활용한다면 보다 풍부한 군중 상호작용을 모델링할 수 있을 것이다. 나아가 예측된 궤적을 3차원 캐릭터 애니메이션, 모션 매칭(motion matching), 물리 기반 제어와 연계한다면, 실제 게임, 가상현실, 도시 안전 분석 등 다양한 응용 분야에서 활용 가능한 고밀도 군중 시뮬레이션 시스템으로 확장할 수 있을 것이다.

6 감사의 글

본 연구는 G-deep XR(2024.04) - RS-2024-00399136 [중앙행정기관 - 문화체육관광부/전문기관 - 한국콘텐츠진흥원]과 중견 연구(2024.05) - RS-2024-00354549 [중앙행정기관 - 과학기술정보통신부/전문기관 - 한국연구재단]을 통해 진행되었다.

References

- [1] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Physical Review E*, vol. 51, no. 5, pp. 4282–4286, 1995.
- [2] J. van den Berg, S. J. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," *Robotics Research, Springer Tracts in Advanced Robotics*, vol. 70, pp. 3–19, 2011.
- [3] P. Charalambous, J. Pettré, V. Vassiliades, Y. Chrysanthou, and N. Pelechano, "Greil-crowds: Crowd simulation with deep reinforcement learning and examples," *ACM Transactions on Graphics*, vol. 42, no. 4, pp. 1–15, 2023.
- [4] G. Gomez-Nogales, M. Prieto-Martin, C. Romero, M. Comino-Trinidad, P. Ramon-Prieto, A.-H. Olivier, L. Hoyet, M. Otaduy, J. Pettre, and D. Casas, "Resolving collisions in dense 3d crowd animations," *ACM Transactions on Graphics (TOG)* 43, 5, Article 163, pp. 1–14, 2024.
- [5] T. Salzmann, B. Ivanovic, P. Chakravarty, and M. Pavone, "Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data," *Proceedings of the European Conference on Computer Vision (ECCV)*, p. 683–700, 2020.
- [6] A. Mohamed, K. Qian, M. Elhoseiny, and C. Claudel, "Social-stgcnn: A social spatio-temporal graph convolutional neural network for human trajectory prediction," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 14424–14432, 2020.

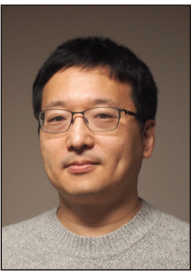
- [7] Y. Yuan, X. Weng, Y. Ou, and K. M. Kitani, “Agentformer: Agent-aware transformers for socio-temporal multi-agent forecasting,” *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, p. 9813–9823, 2021.
- [8] J. Gao, C. Sun, H. Zhao, Y. Shen, D. Anguelov, C. Li, and C. Schmid, “Vectornet: Encoding hd maps and agent dynamics from vectorized representation,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 11522–11530, 2020.
- [9] Z. Zhou, L. Ye, J. Wang, K. Wu, and K. Lu, “Hivt: Hierarchical vector transformer for multi-agent motion prediction,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8823–8833, 2022.
- [10] Z. Zhou, J. Wang, Y.-H. Li, and Y.-K. Huang, “Query-centric trajectory prediction,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 17863–17873, 2023.
- [11] A. Robicquet, A. Sadeghian, A. Alahi, and S. Savarese, “Learning social etiquette: Human trajectory understanding in crowded scenes,” *Proceedings of the European Conference on Computer Vision (ECCV)*, p. 549–565, 2016.
- [12] R. Martín-Martín, M. Patel, H. Rezatofighi, A. Sheno, J. Gwak, E. Frankel, A. Sadeghian, and S. Savarese, “Jrdb: A dataset and benchmark of egocentric robot visual perception of humans in built environments,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, p. 6748–6765, 2023.
- [13] J. W. Bae, J. Kim, J. Yun, C. Kang, J. Choi, C. Kim, J. Lee, J. Choi, and J. W. Choi, “Sit dataset: Socially interactive pedestrian trajectory dataset for social navigation robots,” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- [14] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, “Social lstm: Human trajectory prediction in crowded spaces,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 961–971, 2016.
- [15] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, “Social gan: Socially acceptable trajectories with generative adversarial networks,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2255–2264, 2018.
- [16] Z. Chang, H. Wang, G. Koulouris, and H. P.H., “Large-scale multi-character interaction synthesis,” *ACM Transactions on Graphics (TOG)*, Article 114, pp. 1–10, 2025.

〈 저 자 소 개 〉



장 예 정

- 2020-2025: 가천대학교 컴퓨터공학과 학사
- 2025-현재: 한양대학교 컴퓨터소프트웨어학과 석사
- <https://orcid.org/0009-0008-9736-0884>



권 태 수

- 1996-2000: 서울대학교 전기컴퓨터공학부 학사
- 2000-2002: 서울대학교 전기컴퓨터공학부 석사
- 2002-2007: 한국과학기술원 전산학전공 박사
- 현재: 한양대학교 컴퓨터·소프트웨어학과 교수 재직 중.
- 관심분야: 컴퓨터 애니메이션, 물리 기반 캐릭터 애니메이션, 기계 학습.
- <https://orcid.org/0000-0002-9253-2156>