

# Grid-World 경로탐색을 위한 Diffusion Heatmap 재구성: Conditional DDPM 기반 아모타이즈드 플래닝

장현준, 유재천

성균관대학교 전자전기컴퓨터공학과

10hour@skku.edu, yoojc@skku.edu

## Diffusion Heatmap Reconstruction for Grid-World Path Planning: Conditional DDPM-Based Amortized Planning

Jang Hyun Jun, Yoo Jae Chern

Dept. of Electrical and Computer Engineering, Sungkyunkwan University

### 요 약

본 논문은 시작점- 목표점- 장애물이 주어진 격자 환경에서, 조건부 확산모델(diffusion)을 통해 전역 탐색 히트맵  $H_{pred}$  을 생성하고 이를 기반으로 경로를 추출하는 아모타이즈드(amortized) 경로탐색 프레임워크를 제안한다. 기존 히트맵  $H_{opt}$  은 최단경로 전수 기반의 동적계획(DAG)으로 구성하고, 예측 히트맵과의 분포 차이를 평가한다. 결과적으로  $H_{pred}$  는 전통적 A\*/Dijkstra 의 반복 재탐색 없이 저비용 로컬 물로도 안정적인 추종이 가능하며, 장애물 이동/폐색 시에도 즉시 우회 경향을 나타낸다. 또한  $\Delta H = H_{pred} - H_{opt}$  및 신뢰도 지표를 통해 설명가능성을 제공한다.

### I. 서 론

로봇/자율주행에서의 전통적 경로탐색(A\*, Dijkstra)은 환경 변화가 발생할 때마다 전역 재탐색이 요구되어  $O(E \log V)$  수준의 비용과 싱글 솔루션(불확실성 부재)이라는 한계를 가진다. D\*, LPA\* 같은 증분형 변형이 존재하지만, 빈번한 장애물 변화나 부분 폐색이 누적되는 실 환경에서는 여전히 재계산 오버헤드와 급격한 경로 변동을 피하기 어렵다. 특히 동적 장애물이나 부분 폐색 상황에서는 잦은 재계획으로 지연이 누적되며, 경로가 단번에 급변하는 문제가 발생한다.

본 논문은 이러한 문제를 완화하기 위해, 경로를 먼저 분포로 그려놓고(히트맵) 그 분포를 따라가는 역(逆) 파이프라인을 제안한다. 구체적으로 시작/목표/장애물/좌표 인코딩을 조건으로 하는 조건부 확산모델이 전 공간의 선호도 필드  $H_{pred}$  를 생성한다. 이후 플래너는  $-\log(H_{pred})$ 를 비용으로 하여 A\* 혹은 간단한 로컬 정책으로 경로를 추종한다.

이 접근은 (i) 환경 변화에 대한 재탐색 비용을 대폭 줄이는 아모타이즈드(Amortized) 플래닝, (ii) 단일 최단 경로가 아닌 다중 경로의 분포와 불확실성을 내재화, (iii) 분포/신뢰도 지표를 통한 설명가능성(XAI) 제공이라는 장점을 갖는다. 본 연구는 제안 프레임워크의 설계·학습·평가 절차를 제시하고, Grid-World 실험에서 정량(예: JSD/EMD/Top-K/ECE, 길이비) 및 정성(히트맵·경로 시각화) 결과로 그 유효성을 보고한다.

### II. 본론

본 연구는 경로를 한 번에 확정하지 않고, 전 격자에 대한 선호도 분포(heatmap)를 먼저 생성한 뒤 이를 비용으로 변환해 저비용으로 추종하는 아모타이즈드 경로탐색을 목표로 한다. 파이프라인은 그림 1 과 같이 두 갈래로 구성된다. Ground-Truth 분기는 시작·목표·장애물로 부터 최단경로들의 분포를 계산해 학습 타깃 히트맵  $H_{opt}$  을 만든다. Diffusion 분기는 조건(시작 S, 목표 G, 장애물 O, 좌표 인코딩 X,Y)을 입력으로 조건부 DDPM 을 통해 예측 히트맵  $H_{pred}$  을 생성한다. 마지막으로  $H_{pred}$  를 비용으로 바꾸어 A 또는 로컬 정책으로 경로를 얻는다. 이 순서는 “경로를 먼저 그리지 않고, 분포를 먼저 그린 후 경로를 따라간다”는 본 연구의 핵심 아이디어를 반영한다.

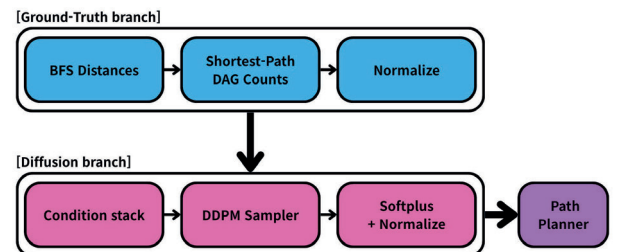


그림 1. 제안 파이프라인(Heatmap-Amortized Planning)

먼저 시작점과 목표점으로부터의 최단거리 지도를 BFS 로 얻으면, 임의의 셀  $(i,j)$ 이 최단경로 위에 놓일 필요충분조건은  $ds(i,j) + dg(i,j) = D (=ds(g))$ 이다. 이 제약을

만족하는 셀들만 모아 최단경로 DAG 를 구성하면, 시작 → 셀의 경로 수(Cs)와 셀→목표의 경로 수(Cg)를 동적계획으로 독립 집계할 수 있다.

$$\text{최단거리 정의: } \begin{cases} d_s(i, j) = \text{dist}(s, (i, j)), \\ d_g(i, j) = \text{dist}((i, j), g) \end{cases}$$

$$\text{전체 최단거리: } D = d_s(g) = \min_{\pi: s \rightarrow g} \text{len}(\pi)$$

$$\text{최단경로 제약: } d_s(i, j) + d_g(i, j) = D$$

$$\text{정규화: } H_{\text{opt}}(i, j) = \frac{n_{\text{sp}}(i, j)}{\sum_{p, q} n_{\text{sp}}(p, q)}$$

H<sub>opt</sub> 을 “데이터”로 보고, 노이즈가 섞인 x<sub>t</sub> 에서 노이즈 ε 를 예측해 제거하도록 U-Net 을 학습한다(ε -예측). 입력 조건 c 는 [S, G, O, X, Y]로 구성된다. S, G 는 원-핫, O 는 장애물 마스크, X, Y 는 좌표 인코딩으로, 모델이 “어디가 시작-목표이고 어디가 금지영역인지”와 “격자 위치”를 즉시 파악하게 돕는다.

$$\text{스케줄: } \beta_t, \quad \alpha_t = 1 - \beta_t, \quad \bar{\alpha}_t = \prod_{k=1}^t \alpha_k$$

$$\text{조건 스택: } \mathbf{c} = [S, G, O, X, Y]$$

$$\text{학습 목적(MSE): } L(\theta) = \mathbb{E}_{t, x_0, \varepsilon} \left[ \|\varepsilon - \varepsilon_\theta(x_t, t, \mathbf{c})\|_2^2 \right]$$

추론은 무작위 x<sub>T</sub>~N(0, I)에서 시작해 조건을 고정 한 채 역과정을 내려오며 x<sub>0</sub> 를 얻는다. 이때 조건/무조건 예측을 가중합해(가이던스 w) 분포의 집중도를 제어한다. 최종 x<sub>0</sub> 는 softplus로 비음수화한 뒤 총합 1로 정규화해 확률 히트맵 H<sub>pred</sub> 로 사상한다. 이렇게 얻은 H<sub>pred</sub> 는 “전역 선호도 필드”로 해석할 수 있어, 이후 플래닝 단계에서 비용으로 바로 변환된다. 장애물 셀에 대해서도 H<sub>pred</sub> 는 완전 0 이 아니라 미소 질량을 두는데, 이는 환경이 바뀔 때(장애물 이동/해제) 분포가 급격히 끊기지 않고 부드럽게 재분배되도록 하는 장점이 있다.

$$\text{가이던스 결합: } \hat{\varepsilon} = (1 + w) \varepsilon_\theta(x_t, t, \mathbf{c}) - w \varepsilon_\theta(x_t, t, \mathbf{0})$$

$$\text{비용 격자: } \text{cost}(i, j) = -\log(H_{\text{pred}}(i, j) + \epsilon) + M \mathbf{1}\{O(i, j) = 1\}$$

$$\text{국소 최대화 정책: } \text{next}(x_t) = \arg \max_{u \in \mathcal{N}(x_t)} H_{\text{pred}}(u)$$

확률을 곱셈적 선호도로 보면, -log 변환을 통해 가산 비용으로 만들 수 있고 이는 A 에 즉시 호환된다. 우리는 cost=-log(H<sub>pred</sub>+ε)+M·1[장애물]로 정의하여, 현재 장애물은 큰 패널티 M 으로 확실히 배제하지만, 분포 자체는 장애물 주변에 미소 질량을 유지한다. 결과적으로 동적 장애물이나 부분 폐색이 발생해도 전역 재탐색 없이(혹은 매우 짧은 국소 재탐색만으로) H<sub>pred</sub> 가 가리키는 “확률 경사”를 따라 신속히 우회할 수 있다(아모타이즈드 플래닝).

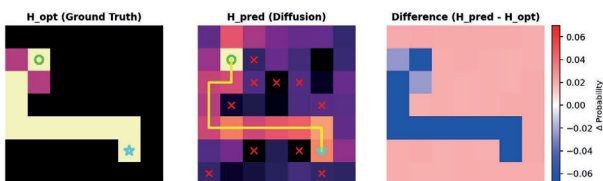


그림 2. Ground Truth vs. Diffusion Heatmap

### III. 결론

본 연구는 소형 Grid-World 에서 조건부 DDPM 이 생성한 H<sub>pred</sub> 와 최단경로 전수 기반 H<sub>opt</sub> 를 비교한 결과, TopK@20% 커버리지가 0.84±0.18 로 높아 핵심 경로 대역을 안정적으로 포착했고, 길이비(len-ratio)는 거의 1.0 에 수렴하여 A\*가 -log(H<sub>pred</sub>) 위에서 최단경로에 근접한 해를 찾음을 확인했다.

지표(metric)	JSD	EMD (approx)	TopK@20% coverage	ECE (cell)
값	0.38±0.09	1.82±0.60	0.84±0.18	0.22±0.14

표 1. 제안 방법의 정량 평가

이 분포적 성질은 동적 장애물·부분 폐색 환경에서 특히 유효했다. H<sub>pred</sub> 가 전역 선호도 필드로 기능하기 때문에, 장애물이 이동하거나 새로 나타나도 전역 재탐색 없이 분포의 경사를 따라 즉시 우회하는 경향을 보였다. 이는 전통적 A\*/Dijkstra 가 환경 변화마다 O(E log V)의 비용을 지불하던 한계를 아모타이즈드(planning-by-distribution) 방식으로 실질적으로 완화한 결과다.

한편 한계도 분명하다. 첫째, 분해능 부족 문제로 그리드 크기를 키우면 히트맵이 과도하게 평탄화되어 EMD 가 커지고 모서리에서 오차가 집중된다. 둘째, 데이터 다양성 문제로 시작/목표 배치와 장애물 패턴의 다양성이 부족한 경우 장면별 편차(TopK 표준편차↑)가 커진다. 이는 대규모·다양한 지도 데이터와 더 깊은 네트워크로 스케일링할수록 분포 정합과 신뢰도가 함께 개선될 것으로 보인다.

종합하면, 제안한 Heatmap-Driven(Amortized) 경로 탐색은 재탐색 비용 절감, 다중 경로·불확실성의 자연스러운 내재화, XAI 친화적 평가/시각화를 동시에 달성했다. 본 프레임워크는 2D 격자에서의 유효성이 확인되었고, 이후 연속 좌표/대형 맵·3D 환경(지형, 시야 차단)·MPC/RL 과의 하이브리드로 확장될 여지가 크다. 본 접근은 동적·복잡 환경에서 안정적이고 설명 가능한 경로계획의 실질적 대안이 될 수 있다.

### ACKNOWLEDGMENT

This research was supported by the BK21 FOUR Program, Sungkyunkwan University.

### 참고 문헌

- [1] Ho J., Jain A. N., Abbeel P., “Denoising Diffusion Probabilistic Models,” Advances in Neural Information Processing Systems 33 (NeurIPS), 2020.
- [2] Ho J., Salimans T., “Classifier-Free Diffusion Guidance,” arXiv preprint arXiv:2207.12598, 2022.
- [3] Janner M., Du Y., Tenenbaum J. B., Levine S., “Planning with Diffusion for Flexible Behavior Synthesis,” Proc. ICML, PMLR 162:9902-9915, 2022.