

노이즈가 포함된 사람의 시연 데이터 활용을 위한 생성 모델 기반 경로 획득 방법론

A Generative Model-based Approach to Obtain a Robust Trajectory against Noisy Human Demonstration Data

박찬진¹.이정우¹.현승민¹.김송우¹.정호진¹.윤한얼²

Chanjin Park, Jungwoo Lee, Seungmin Hyun, Songwoo Kim, Ho-Jin Jung and Han Ul Yoon

¹연세대학교 일반대학원 전산학과

E-mail: {cjpark04, jw.lee, smhyun, swkim, hojinj}@yonsei.ac.kr

²연세대학교 소프트웨어학부

E-mail: huyoon@yonsei.ac.kr

요 약

Learning from demonstration에서 사용하는 사람의 시연 데이터는 개개인의 heuristic policy로부터 샘플링된 결과물이다. 하지만, 생성된 경로에는 외부 요소들에 의한 노이즈가 포함된다. 본 논문에서는 사람의 시연 데이터를 활용하기 위해, 노이즈가 포함된 데이터로부터 Diffusion을 사용하여 경로를 생성해 내고자 한다. 먼저, 직접 교시를 통해 로봇 매니퓰레이터를 원하는 도착점으로 이동하는 사람의 시연 데이터를 획득한다. 다음으로, 시연 데이터를 전처리하고, Diffusion 모델을 학습한다. 최종적으로, 학습된 Diffusion 모델에서 경로를 생성하여, 로봇 매니퓰레이터의 움직임을 가상환경에서 재현한다. 제안하는 방법론을 통해, 노이즈가 포함된 사람의 시연 데이터도 로봇 시스템에 대한 LfD에서의 활용을 기대할 수 있다.

키워드 : 생성 모델, 디퓨전, 모션 플래닝, 경로 생성

1. 서 론

Learning from demonstration(이하 LfD)는 사람의 시연 데이터를 기반으로, 사람의 정밀한 태스크 수행을 모방하는 시스템 구현을 위한 방법론이다. 자세히 말하면, 사람의 시연 데이터를 정답 값으로 제공해 줌으로써, 모델이 사람처럼 행동하도록 학습하는 것이다. 따라서, 사람의 정교한 행동을 잘 재현하는 모델을 얻기 위해서는, 태스크를 성공적으로 수행함과 동시에 손 떨림과 같은 노이즈가 최소화된 일관성(consistency)이 있는 데이터를 모델에 제공해야 한다. 그러나, 시연 데이터는 개개인의 특성이 반영된 휴리스틱 정책(heuristic policy)의 결과물이므로, 모든 반복 시행에 대해 일관성 있는 데이터를 생성(획득)하기는 어렵다.

본 논문은 사람의 시연 데이터를 로봇 매니퓰레이터 경로 생성에 이용하기 위해, Diffusion을 적용하여 사람이 수행한 프로세스를 모델링하고자 한다. 먼저, 직접 교시를 통해 로봇 매니퓰레이터를 시작점에서 도착점까지 이동시키면서 사람의 시연 데이터를 얻는다. 다음으로, 획득된 시연 데이터에 전처리를 진행한다. 이후, 전처리가 진행된 시연 데이터를 활용하여 Diffusion 모델을 학습시킨다. 학습된 모델로부터, 사람의 시연을 모방하는 경로를 생성하여, 가상환경에서 로봇 매니퓰레이터 움직

임을 재현해 보고자 한다.

2. 제안하는 방법론

2.1 생성 모델 Diffusion

Diffusion 모델은 크게 forward 과정과 reverse 과정으로 이루어져 있다. Forward 과정은 초기 데이터 x_0 에서 매 스텝 가우시안 노이즈를 더하여 x_t 로 표현하는 과정을 의미한다. 최종 스텝을 T 라고 한다면, x_T 는 표준 정규 분포를 따르게 된다. Diffusion 학습 시간 때는 임의의 x_t 에서 $x_T \sim N(0, I)$ 로 표현하기 위해 요구되는 가우시안 노이즈를 예측하도록 학습된다.

다음으로 Reverse 과정은 최종 스텝 데이터인 $x_T \sim N(0, I)$ 를 시작으로, 노이즈 제거 과정 통해 \hat{x}_0 를 획득하는 과정을 말한다. 따라서, 데이터를 생성하고자 할 때, 노이즈 벡터 x_T 를 샘플링하여 reverse 과정을 통해 \hat{x}_0 를 획득하게 된다.

2.2 Diffusion 기반 경로 획득 방법론

$$s_t = [q_{1,t}, \dot{q}_{1,t}, q_{2,t}, \dot{q}_{2,t}, \dots, q_{n,t}, \dot{q}_{n,t}]^T. \quad (1)$$

본 논문에서 생성하고자 하는 데이터는 로봇이 원하는 도착점에 도달하기까지의 움직임을 담은 위치, 속도 프로파일 경로이다. 시작점으로부터 도착점까지의 타임 스텝에 따른 경로는 식(1)과 같이 n 개의 조인트에 대한 각도 $q_{n,t}$ 와 각속도 $\dot{q}_{n,t}$ 로 정의한다.

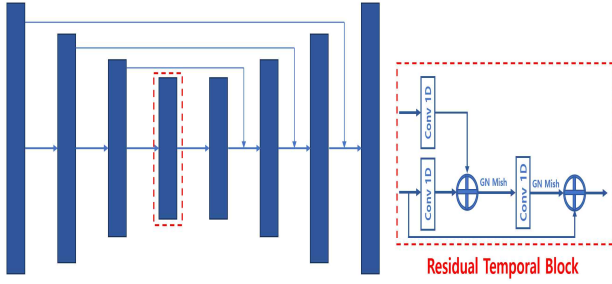


그림 1. 로봇 패스플래닝 문제를 Diffusion으로 풀기 위한 Residual Temporal U-Net 구조[3].

사람이 로봇을 움직이면서 얻은 사람의 시연 데이터를 활용하여 경로를 생성하기 위해서는 forward 과정에서의 가우시안 노이즈 ϵ 을 예측해야 한다. 그림 1은 로봇 패스 플래닝 문제를 Diffusion으로 해결하기 위해 제안된 Janner et al.의 신경망 구조로, ϵ 을 예측하기 위해 Residual Temporal U-Net 구조를 사용하였다[3]. 이미지 생성에 있어 2차원 공간 convolution 연산을 적용한 기존에 제안된 구조들과는 다르게, 1차원 시간 convolution 연산으로 구현되어 있다.

3. 실험 및 결과

3.1 실험 구성 및 진행 프로토콜

사람의 시연 데이터를 획득하기 위해, 로봇 매니플레이터(Kinova Gen2, Kinova, Boisbriand QC, Canada)를 직접 교시를 통해 조작해 주었다. 시연 데이터 획득을 위한 진행 프로토콜은 다음과 같다:

- 1) 로봇 매니플레이터를 원하는 방향으로 조작하기 위해 모드를 admittance control로 변경
- 2) 미리 정해놓은 시작점부터 도착점까지 최대한 동일한 경로를 유지하면서, 직접 교시를 통해 로봇 매니플레이터를 이동 (총 이동 소요시간 30초). 데이터셋 1개는 7384x2n개의 데이터 샘플로 구성 (n =조인트의 개수)
- 3) 과정 2)를 30회 반복
- 4) 과정 3)을 완료하여 얻어진 데이터셋들의 평균에 white noise(평균: 0, 분산: 1)를 더하여 20,000개로 데이터셋 갯수를 증강
- 5) 과정 4)에서 얻어진 데이터셋들을 각각 7384x2n에서 64x2n으로 다운 샘플링

위의 실험을 로봇 매니플레이터의 2-DoF($n=2$)와 3-DoF($n=3$)의 움직임에 대해 각각 시행하였다. 최종적으로, 2-DoF와 3-DoF에 대해 얻어진 20,000개의 데이터셋을 이용하여 사람의 시연을 모방하는 경로를 생성하기 위한 Diffusion 모델을 학습하였다.

3.2 사람의 시연 데이터와 Diffusion이 생성한 경로 비교

그림 2는 $n=2$ 에 대해 사람의 시연과 3.1절에 소개된 과정으로 학습된 Diffusion 모델에 의해 생성된 로봇 매니플레이터의 경로들을 보여준다. 각 점의 색깔 밝기는 시간 스텝이 증가함에 따라 밝아지도록 표현하였다. 그림 2에서 볼 수 있는 것과 같이, Diffusion으로 생성한 경로가 사람의 시연 데이터 경로와 유사하게 생성된 것을 볼 수 있다. 그림 3은 $n=3$ 에 대한 결과를 보여준다. 시연 데이터와 Diffusion으로 생성된 경로가 3차원

에서도 유사하게 생성되는 것을 볼 수 있다. 그림 4는 2-DoF의 경우 Diffusion에 의해 생성된 경로를 Unity에서 재현한 스냅샷을 보여준다.

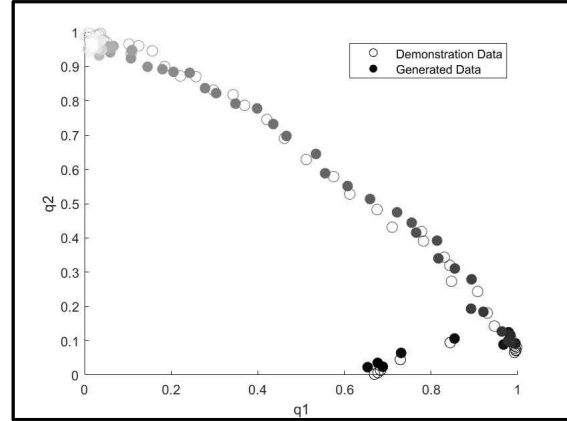


그림 2. 로봇 매니플레이터 2-DoF($n=2$)에 대해 사람 시연 데이터 경로와 Diffusion에 의해 생성된 경로 비교

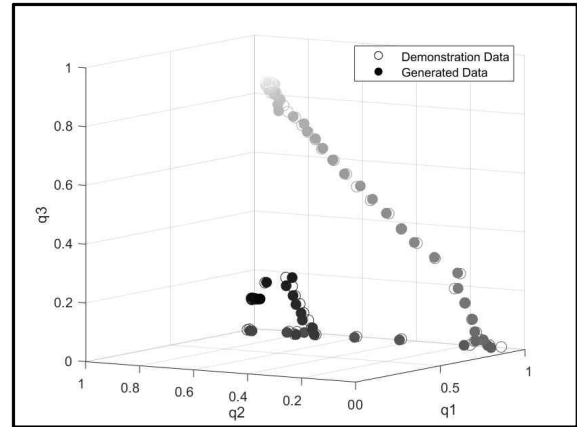


그림 3. 로봇 매니플레이터 3-DoF($n=3$)에 대해 사람 시연 데이터 경로와 Diffusion에 의해 생성된 경로 비교

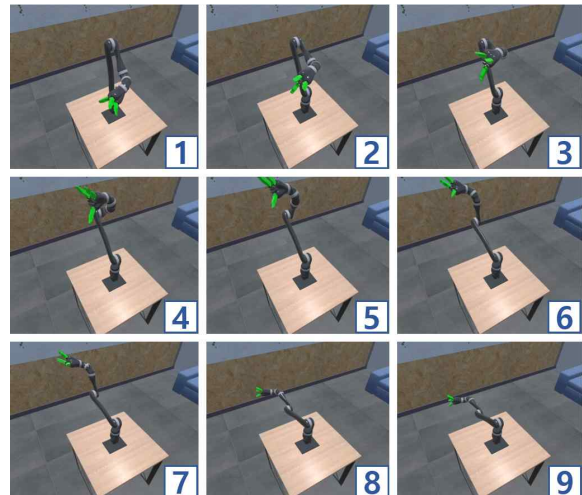


그림 4. Diffusion에 의해 생성된 경로를 가상환경에서 구현한 모습

4. 결론 및 향후 연구

본 논문은 노이즈가 포함된 시연 데이터를 사용하여 LfD 학습에 유용한 데이터를 생성하는 방법론에 대해 논하였다. 사람의 시연 데이터는 사람이 휴리스틱 정책 (heuristic policy)에 의해 구현된 데이터이지만, 시연이 반복되는 과정에서의 오차가 노이즈로 더해진다. 이 노이즈가 포함된 데이터도 LfD에 활용하기 위하여, 생성 모델인 Diffusion을 통해 사람의 시연 데이터를 학습하였다. Diffusion이 생성한 경로를 가시화하여 사람의 시연 데이터와 비교하고 Unity에서의 재현을 통해, 제안된 방법론의 타당성을 검증하였다.

향후 연구로는 Stanford의 Mobile ALOHA와 같이 양팔 로봇을 위한 사람의 시연 데이터를 Diffusion으로 학습하여, 양팔 로봇의 작업 수행을 위한 경로를 생성해 볼 계획이다[4].

참 고 문 헌

- [1] S. Schaal, "Learning From Demonstration," in *Proceedings of the 9th International Conference on Neural Information Processing Systems(NeurIPS)*, Denver, USA, December 2-5, pp. 1040-1046, 1996
- [2] J. Ho et al., "Denoising Diffusion Probabilistic Models," in *Proceedings of the 34th International Conference on Neural Information Processing Systems(NeurIPS)*, Vancouver, Canada, December 6-12, pp. 6840-6851, 2020.
- [3] M. Janner et al., "Planning with Diffusion for Flexible Behavior Synthesis," *arXiv preprint arXiv:2205.09991*, 2022
- [4] Z. Fu et al., "Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation," *arXiv preprint arXiv:2401.02117*, 2024