

VAE-MPPI: 양팔 로봇 조작을 위한 폐쇄 연쇄 제약 궤적 최적화

VAE-MPPI: Closed-Chain Constrained Trajectory Optimization for Dual-Arm Manipulation

이 슬 찬¹ · 김 상 현[†]

Seulchan Lee¹, Sanghyun Kim[†]

Abstract: Coordinated dual-arm manipulation, where two robots jointly interact with a shared object, forms a closed-chain system that imposes strict kinematic equality constraints on the robot configuration. Conventional Model Predictive Path Integral (MPPI) control addresses such constraints only through soft penalty terms, making strict feasibility guarantees fundamentally unachievable. This paper presents a novel MPPI-based framework that enforces hard equality constraint satisfaction for closed-chain dual-arm manipulation. A Variational Autoencoder (VAE) learns a compact latent representation of the constraint manifold, enabling efficient near-feasible trajectory sampling. KKT-based projection is then applied in parallel across all samples on GPU, guaranteeing exact constraint satisfaction during trajectory optimization. A hierarchical control architecture decouples manifold-constrained trajectory optimization at 10 Hz from joint-level PD control at 1000 Hz, enabling real-time execution. Validated in MuJoCo simulation on a dual Franka Emika Panda platform, the proposed framework achieves stable constraint satisfaction and near real-time trajectory optimization across both goal-tracking and obstacle avoidance tasks.

Keywords: Trajectory Optimization, Model Predictive Control, Multi-Robot Control Systems

1. 서 론

로봇 시스템에서 기구학적·동역학적 제약 조건의 만족은 안전하고 실행 가능한 동작을 보장하기 위한 핵심 요건이다. 특히

양팔 로봇^[1]이 공유 물체를 함께 파지하는 협력 조작 태스크^[2,3]에서는 폐쇄 연쇄 기구학적 구조(Closed kinematic chain)가 형성되어 관절 공간에 엄격한 등호 제약 조건이 부과된다. 이 제약 조건을 위반하는 궤적은 물리적으로 실행 불가능하며, 파지 물체의 이탈이나 로봇 손상으로 이어질 수 있다. 따라서 폐쇄 연쇄 양팔 조작에서는 제약 조건을 만족하는 궤적을 실시간으로 생성하고 추종하는 것이 핵심 과제이다.

기존에는 리더-팔로워^[4], 임피던스 제어^[5], QP 또는 Null-Space 기반 방법^[6,7], NMPC^[8] 등이 폐쇄 연쇄 양팔 조작에 적용되어 왔으나, 이들은 예측 지평을 고려한 궤적 최적화 능력이 없거나 고차원 비선형 시스템에서의 연산 비용으로 인해 실시간 궤적 재계획이 어렵다는 공통적 한계를 가진다. 궤적 최적화(Trajectory Optimization, TO) 관점에서 이를 보다 직접적으로 다루려는 시도로 CHOMP^[9], TrajOpt^[10], iLQR^[11], DDP^[12,13] 등이 연구되었다. 이들은 전체 궤적을 대상으로 제약 조건을 명시적으로 다루며 고품질 궤적을 생성할 수 있으나, 경사 기반 최

Received : Mar. 17. 2026; Revised : Apr. 23. 2026; Accepted : May. 4. 2026

※ This research was supported by the Ministry of Trade, Industry & Energy (MOTIE, Republic of Korea) through the Industrial Strategic Technology Development Program (No. RS-2025-25449039) and the Korea Planning & Evaluation Institute of Industrial Technology (KEIT) (No. RS-2025-11082970); by the Ministry of Science and ICT through a Korea Basic Science Institute (National Research Facilities and Equipment Center) grant (No. RS-2025-00564593); and by the Ministry of Education and the Regional Innovation System & Education (RISE) program through the Gyeonggi RISE Center, funded by the Ministry of Education (MOE) and the Gyeonggi-do, Republic of Korea (2026-RISE-09-A07).

1. Graduate Student, Mechanical Engineering, Kyung Hee University, Yongin, Korea (lee081847@khu.ac.kr)

† Assistant Professor, Corresponding author: Mechanical Engineering, Kyung Hee University, Yongin, Korea; Adjunct Researcher, Advanced Institute of Convergence Technology, Suwon, Korea (kim87@khu.ac.kr)

적화에 의존하여 초기 해의 품질에 민감하고 매 제어 주기마다 재최적화를 수행하기에는 연산 비용이 과도하다. 이러한 한계는 실시간 갱신이 가능한 샘플링 기반 접근법에 대한 동기를 제공한다. 샘플링 기반 궤적 최적화 중 가장 주목받는 방법은 MPPI (Model Predictive Path Integral)^[14,15]이다. MPPI는 GPU 기반 대규모 병렬 샘플링을 통해 경사 계산 없이 비선형 비용 함수를 최적화할 수 있으며, 수천 개의 궤적 후보를 실시간으로 평가하고 가장 평균을 통해 최적 궤적을 갱신한다. STORM^[16]은 이를 고차원 로봇 매니퓰레이터에 성공적으로 적용한 대표적 사례이다. 그러나 표준 MPPI는 등호 제약 조건을 비용 함수 내 소프트 패널티로만 처리하기 때문에 엄격한 제약 만족을 수학적으로 보장할 수 없다. 이를 극복하기 위해 π -MPPI^[17]는 KKT 조건 기반 업데이트를, CSC-MPPI^[18]는 DBSCAN 기반 안전 필터링을 도입하였으나, 이들은 주로 부등호 제약 조건을 대상으로 설계되어 매니폴드 기반 등호 제약 조건에는 계산 비용 폭발과 수치적 불안정성으로 인해 실시간 적용이 사실상 불가능하다.

본 논문의 핵심 관찰은 다음과 같다. 등호 제약 조건 $h(\mathbf{q}) = 0$ 은 전체 관절 공간 내에 저장된 실현 가능 매니폴드 $\mathcal{M}$ 을 정의하며, 이 매니폴드 위에서 직접 샘플링이 이루어진다면 모든 궤적 후보가 하드 제약을 원천적으로 만족할 수 있다. 그러나 매니폴드 위에서의 샘플링을 야코비안 역행렬 기반 사영으로 구현하면 샘플 수에 비례하는 막대한 연산량이 요구된다. 이를 해결하기 위해 본 논문은 VAE (Variational Autoencoder)^[19-21]를 통해 제약 매니폴드의 저차원 잠재 표현을 학습하여 근실현 가능 (near-feasible) 샘플을 효율적으로 생성하고, GPU 병렬화된 KKT 기반 사영을 적용하여 각 샘플의 엄격한 등호 제약 만족을 보장하는 VAE-MPPI 프레임워크를 제안한다. 최적화된 궤적은 저수준 PD 제어기에 의해 추종된다.

본 논문의 주요 기여는 다음과 같다. 첫째, 폐쇄 연쇄 양팔 시스템에 대해 등호 제약 조건을 엄격히 만족하는 새로운 MPPI 기반 실시간 궤적 최적화 프레임워크를 제안한다. 둘째, VAE 기반 잠재 공간 샘플러와 KKT 기반 사영을 결합하여 매니폴드 불일치 문제를 해결하고 각 샘플의 엄격한 제약 조건 만족을 달성한다. 셋째, 제안한 알고리즘을 MuJoCo^[22] 시뮬레이션 환경에서 목표 추종 및 장애물 회피 태스크를 통해 검증한다.

2. 사전 지식

2.1 변분 오토인코더

변분 오토인코더(VAE)는 생성형 딥러닝 모델로, 입력 데이터의 확률적 잠재 표현을 학습하는 것을 목표로 한다. 기존의 오토인코더가 입력을 고정된 잠재 벡터로 압축하는 것과 달리, VAE는 잠재 공간을 평균 μ 와 분산 σ^2 으로 매개변수화된 확률

분포로 모델링한다. 이를 통해 잠재 공간에서의 샘플링이 가능해지며, 학습된 분포로부터 새로운 데이터를 생성할 수 있다.

VAE는 인코더와 디코더로 구성된다. 인코더는 입력 데이터 x 를 잠재 분포의 매개변수 μ, σ 로 변환하며, 디코더는 잠재 벡터 z 로부터 원본 데이터를 복원한다. 학습은 증거 하한(Evidence Lower Bound, ELBO)을 최대화하는 방식으로 이루어지며, 이는 아래와 같이 재구성 손실과 KL 발산의 합으로 표현된다.

$$\mathcal{L} = \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z})] - D_{\text{KL}}(q_\phi(\mathbf{z}|\mathbf{x})||p(\mathbf{z})) \quad (1)$$

여기서 첫 번째 항은 디코더의 재구성 정확도를 높이는 재구성 손실이며, 두 번째 항은 잠재 분포를 표준 정규분포 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ 에 가깝게 정규화하는 KL 발산 항이다. VAE의 잠재 공간은 연속적이고 구조화된 특성을 가지므로, 데이터가 분포하는 저차원 매니폴드를 효과적으로 근사할 수 있다.

2.2 모델 예측 경로 적분

모델 예측 경로 적분은 샘플링 기반의 모델 예측 제어기로, 자유 에너지와 상대 엔트로피 간의 정보 이론적 관계를 기반으로 유도된다. 이산 시간 확률 동역학 시스템 $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t) + \epsilon_t$, $\epsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$ 를 고려할 때, MPPI는 최적 제어 분포 $\mathbb{Q}^*$ 와 샘플링 분포 $\mathbb{Q}$ 사이의 KL 발산을 최소화하는 문제로 정식화된다.

$$\min_q D_{\text{KL}}(\mathbb{Q}||\mathbb{Q}^*) = \min_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}} \left[\log \frac{\mathbb{Q}(\mathbf{u})}{\mathbb{Q}^*(\mathbf{u})} \right] \quad (2)$$

이를 풀면 최적 제어 업데이트는 비용 가중 평균으로 표현되며, k 번째 샘플 제어 시퀀스 $\mathbf{U}^{(k)} = \mathbf{U} + \epsilon^{(k)}$, $\epsilon^{(k)} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ 에 대해 다음과 같이 업데이트된다.

$$\mathbf{U} \leftarrow \mathbf{U} + \sum_{k=1}^K w^{(k)} \epsilon^{(k)} \quad (3)$$

이때, 가중치 $w^{(k)}$ 는 다음과 같이 계산된다.

$$w^{(k)} = \frac{\exp\left(-\frac{1}{\lambda} S^{(k)}\right)}{\sum_{j=1}^K \exp\left(-\frac{1}{\lambda} S^{(j)}\right)} \quad (4)$$

여기서 $S^{(k)}$ 는 k 번째 샘플 궤적의 총 비용이며, λ 는 온도 매개변수로 비용이 낮은 샘플에 더 높은 가중치를 부여하는 정도를 조절한다. MPPI는 GPU 기반 병렬 연산을 통해 수천 개의 샘플을 실시간으로 평가할 수 있어, 고차원 비선형 시스템에 효과적으로 적용 가능하다.

3. 방법론

이 절에서는 폐쇄 연쇄 양팔 로봇 제어를 위한 방법론을 소개한다. 본 논문은 양팔 로봇이 공유 물체를 파지한 채 수평을 유지하며 목표 위치로 이동시키는 태스크를 대상으로 한다. 이를 위해 해당 태스크의 등호 제약 조건을 만족하는 매니폴드를 VAE 기반으로 학습하고, 이를 MPPI 기반 궤적 최적화와 통합하는 프레임 워크를 제안한다. 전체 구조는 [Fig. 1]에 도식화하였다.

3.1 VAE 기반 매니폴드 학습

폐쇄 연쇄 양팔 시스템에서 두 팔이 공유 물체를 함께 파지하는 경우, 왼팔 끝단과 오른팔 끝단의 상대적 변환은 일정하게 유지되어야 한다. 이를 강제 파지 제약(rigid grasp constraint)으로 표현하면 다음과 같다.

$$\phi(\mathbf{q}) = ({}^w\mathbf{T}_l(\mathbf{q}) \cdot {}^l\mathbf{T}_r)^{-1} \cdot {}^w\mathbf{T}_r(\mathbf{q}) - \mathbf{I} = \mathbf{0} \quad (5)$$

여기서 ${}^w\mathbf{T}_l(\mathbf{q}) \in SE(3)$ 와 ${}^w\mathbf{T}_r(\mathbf{q}) \in SE(3)$ 는 각각 월드 프레임에 대한 왼팔과 오른팔 끝단의 동차 변환 행렬이며, ${}^l\mathbf{T}_r$ 은 두 끝단 사이의 고정된 상대 변환이다.

폐쇄 연쇄 강제 파지 제약 이외에도, 물체를 수평으로 유지하기 위한 제약조건을 만족해야하고, 이를 수식으로 표현하면 다음과 같다.

$$\psi(\mathbf{q}) = \begin{bmatrix} roll({}^w\mathbf{R}_o(\mathbf{q})) \\ pitch({}^w\mathbf{R}_o(\mathbf{q})) \end{bmatrix} = 0 \quad (6)$$

결론적으로, 폐쇄 연쇄 양팔 로봇의 등호 제약 조건은 다음과 같이 정의된다.

$$\mathbf{h}(\mathbf{q}) = \begin{bmatrix} \phi(\mathbf{q}) \\ \psi(\mathbf{q}) \end{bmatrix} = \mathbf{0} \in \mathbb{R}^8 \quad (7)$$

이 등호 제약 조건은 전체 관절 공간 $Q \subset \mathbb{R}^n$ 내에 저차원 실현 가능 매니폴드 $\mathcal{M}$ 을 정의한다.

$$\mathcal{M} = \{\mathbf{q} \in Q | \mathbf{h}(\mathbf{q}) = \mathbf{0}\} \quad (8)$$

이 매니폴드는 영측도(zero measure)를 가지는 비선형 부분 공간으로, 표준 가우시안 샘플링으로는 실현 가능한 샘플을 얻기 매우 어렵다.

이를 해결하기 위해 VAE를 활용하여 매니폴드 $\mathcal{M}$ 위의 데이터로부터 저차원 잠재 표현을 학습한다. 학습 데이터는 제약

조건 $\mathbf{h}(\mathbf{q}) = \mathbf{0}$ 을 만족하는 관절 구성으로 구성되며, 이는 역기구학 기반 폐쇄 루프 최적화를 통해 수집된다. 수집된 데이터를 기반으로 VAE의 인코더 $q_\phi(\mathbf{z}|\mathbf{q})$ 와 디코더 $p_\theta(\mathbf{q}|\mathbf{z})$ 식(1)의 손실함수를 사용하여 학습한 후, 디코더 $p_\theta(\mathbf{q}|\mathbf{z})$ 는 잠재 벡터 $\mathbf{z}$ 로부터 매니폴드에 근접한 관절 구성을 생성하는 생성기로 활용된다.

3.2 매니폴드 기반 MPPI

학습된 VAE의 잠재 공간을 MPPI의 샘플링 공간으로 활용한다. 표준 MPPI가 관절 공간 또는 작업 공간에서 직접 샘플링하는 것과 달리, 본 프레임워크는 잠재 속도 $\dot{\mathbf{z}}$ 를 제어 입력으로 정의하고 잠재 공간에서 동역학을 전개한다. 잠재 상태의 전이는 다음과 같이 정의된다.

$$\mathbf{z}_{t+1} = \mathbf{z}_t + \dot{\mathbf{z}}_t \cdot \Delta t \quad (9)$$

MPPI는 잠재 속도 시퀀스 $\dot{\mathbf{Z}} = \{\dot{\mathbf{z}}_0, \dot{\mathbf{z}}_1, \dots, \dot{\mathbf{z}}_{T-1}\}$ 를 제어 입력으로 삼아 K 개의 샘플을 다음과 같이 생성한다.

$$\dot{\mathbf{Z}}^{(k)} = \dot{\mathbf{Z}} + \boldsymbol{\epsilon}^{(k)}, \boldsymbol{\epsilon}^{(k)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_z) \quad (10)$$

각 샘플에 대해 잠재 상태 궤적을 전개하면 다음과 같다.

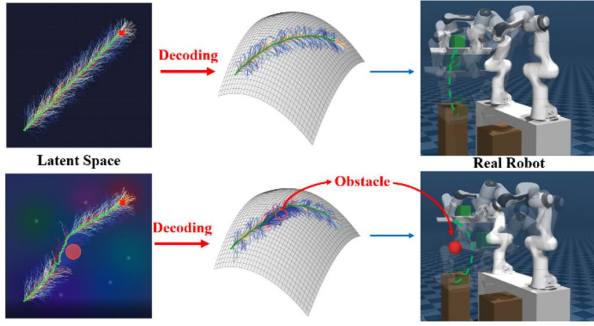
$$\mathbf{z}_{t+1}^{(k)} = \mathbf{z}_t^{(k)} + \dot{\mathbf{z}}_t^{(k)} \cdot \Delta t \quad (11)$$

그러나, 장애물 회피와 같은 부등식 제약 조건이나, 목표 추종과 같은 작업은 여전히 로봇의 작업 공간 내에서 이루어지기 때문에, 잠재공간 상에서 직접적으로 이러한 경로를 평가할 수 없다. 따라서 본 논문에서는 비용함수를 로봇의 작업공간에서 정의하고, 전개된 잠재 궤적의 각 상태 $\mathbf{z}_t^{(k)}$ 를 미리 학습된 디코더를 통해 관절 구성으로 변환되어 비용 함수 평가에 활용된다.

$$\tilde{\mathbf{q}}_t^{(k)} = \mu_\theta(\mathbf{z}_t^{(k)}) \quad (12)$$

이때, $\mu_\theta(\mathbf{z})$ 와 $\tilde{\mathbf{q}}_t^{(k)}$ 는 학습된 디코더와 디코더를 통해 복원된 k 번째 t 시간대 조인트 값을 의미한다.

그러나 VAE는 매니폴드 $\mathcal{M}$ 을 근사적으로만 학습하므로, 디코더 출력 $\tilde{\mathbf{q}}_t^{(k)}$ 는 실제 제약 조건을 엄밀히 만족하지 못하는 매니폴드 불일치(manifold mismatch) 문제가 발생한다. 이를 해결하기 위해서, $\tilde{\mathbf{q}}_t^{(k)}$ 에 대해 KKT 조건 기반 사영을 다음과 같이 수행하여 최종 관절 샘플 $\hat{\mathbf{q}}_t^{(k)}$ 을 계산한다. 사영 문제는 다음과 같이 정식화된다.



[Fig. 1] Overview of the Manifold-Based MPPI framework. (Top) In an obstacle-free environment, trajectories sampled in the latent space are decoded and projected onto the constraint manifold, which are then applied to real dual-arm robot control. (Bottom) In the presence of obstacles, since the latent space cannot explicitly represent inequality constraints, obstacle avoidance is handled through the MPPI cost function while equality constraint satisfaction is maintained via the same projection process

$$\hat{\mathbf{q}}_t^{(k)} = \arg \min_{\mathbf{q}} \frac{1}{2} \|\mathbf{q} - \tilde{\mathbf{q}}_t^{(k)}\|^2, \text{ subject to } h(\mathbf{q}) = \mathbf{0} \quad (13)$$

이 문제의 KKT 조건은 다음과 같다.

$$\mathbf{q} - \hat{\mathbf{q}}_t^{(k)} + \mathbf{J}_h(\mathbf{q})^\top \boldsymbol{\lambda} = \mathbf{0}, \quad h(\mathbf{q}) = \mathbf{0} \quad (14)$$

여기서 $\mathbf{J}_h(\mathbf{q}) = \partial h / \partial \mathbf{q}$ 는 제약 함수의 야코비안이며, $\boldsymbol{\lambda}$ 는 라그랑주 승수이다. 이를 뉴턴 방법으로 반복적으로 풀면 다음과 같다.

$$\begin{bmatrix} \mathbf{q}^{i+1} \\ \boldsymbol{\lambda}^{i+1} \end{bmatrix} = \begin{bmatrix} \mathbf{q}^i \\ \boldsymbol{\lambda}^i \end{bmatrix} - \begin{bmatrix} \mathbf{I} & \mathbf{J}_h^T \\ \mathbf{J}_h & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{q}^i - \hat{\mathbf{q}}_t^{(k)} + \mathbf{J}_h^T \boldsymbol{\lambda}^i \\ h(\mathbf{q}^i) \end{bmatrix} \quad (15)$$

이 사영은 디코더를 통해 복원된 모든 샘플들에 대해서 수행해야 하므로, 일반적으로 높은 계산 비용이 요구된다. 이를 해결하기 위하여, 해당 사영 과정을 GPU 기반으로 병렬화하여 계산 시간을 단축하였다.

사영된 샘플을 이용하여 비용함수 $S^{(k)} = C\left(\left\{\hat{\mathbf{q}}_t^{(k)}\right\}_{t=0}^{T-1}\right)$ 를 평가한 후, MPPI 가중치를 통해 최적 잠재 속도 시퀀스와 최적 관절 값을 계산한다.

$$\dot{\mathbf{z}} \leftarrow \dot{\mathbf{z}} + \sum_{k=1}^K w^{(k)} \dot{\epsilon}^{(k)} \quad (16)$$

$$\mathbf{q}_t^* \leftarrow \sum_{k=1}^K w^{(k)} \mathbf{q}_t^{(k)} \quad (17)$$

이때, 가중치 $w^{(k)}$ 는 식 (4)와 같이 계산된다. 업데이트된 최적

잠재 속도 $\dot{\mathbf{z}}_0^*$ 를 적용하여 현재 잠재 상태를 전진시킨다.

$$\mathbf{z}_t^* = \mathbf{z}_{t-1} + \dot{\mathbf{z}}_{t-1}^* \cdot \Delta t \quad (18)$$

최적 잠재 상태 $\mathbf{z}^*$ 는 다음 경로 최적화의 초기 값으로 업데이트됨으로써 연속적인 관절 최적화를 진행한다.

3.3 계층적 제어 프레임워크

3.2절에서 제안한 매니폴드 기반 MPPI를 실제 로봇 제어에 적용하기 위해, 본 논문은 계층적 제어 프레임워크를 구성한다. 상위 계층에서는 매니폴드 기반 MPPI를 통해 최적 관절 궤적을 계산하며, 전체 샘플에 대한 KKT 사영의 계산 부하를 고려하여 10 Hz 주기로 동작한다. 하위 계층에서는 상위 계층에서 생성된 예측 궤적의 일부 구간을 참조 궤적으로 활용하여, PD 제어를 통해 각 관절에 필요한 토크를 1000 Hz 주기로 계산 및 인가한다. 이러한 계층적 구조를 통해 계산 효율성과 실시간 제어 성능을 동시에 달성한다.

4. 실험

4.1 실험 환경

제안한 알고리즘의 유효성을 검증하기 위해 MuJoCo 물리 시뮬레이션 환경에서 듀얼 Franka Emika Panda 로봇을 이용한 폐쇄 연쇄 양팔 조작 실험을 수행하였다. 두 로봇은 각각 7 자유도를 가지며, y축 방향으로 0.6 m 간격으로 대칭 배치되었다. 양팔은 트레이(폭 0.4 m, 핸들 길이 0.05 m, 핸들 각도 0.8 rad)를 공동으로 파지하여 14 자유도 시스템에 8차원 등호 제약 조건(6차원 폐쇄 연쇄 제약 + 2차원 트레이 수평 제약)을 부과하는 구속 시스템을 구성한다.

VAE는 MLP [512, 512] 구조에 GELU 활성화함수를 사용하며 잠재 공간 차원은 6으로 설정하였다. 학습 데이터는 역기구학 기반으로 수집한 100,000개의 제약 만족 관절 구성으로 구성된다. MPPI는 배치 크기 100, 시간 지평 15 스텝, 온도 매개변수 $\lambda = 0.005$, 잠재 공간 샘플링 표준편차 $\sigma_z = 0.005$ 로 설정하였다. KKT 사영의 수렴 허용 오차는 2×10^{-3} 로 설정하였다. 비용함수는 트레이 중심의 위치 및 자세 추종 비용, 중립 자세 유지 비용, 그리고 구 (sphere) 기반 장애물 회피 비용으로 구성된다.

4.2 목표 추종 실험

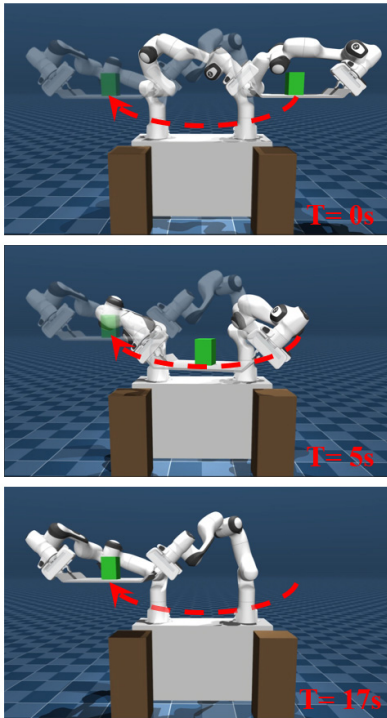
첫 번째 실험에서는 장애물이 없는 환경에서 트레이를 초기

자세로부터 y 축으로 -1.0 m 이동시키는 과제를 수행하였다. 대표 스냅샷은 [Fig. 2]에 제시하였다. [Fig. 3]은 목표 추종 오차, 등호 제약 위반 크기, 연산 시간을 시간에 따라 나타낸다.

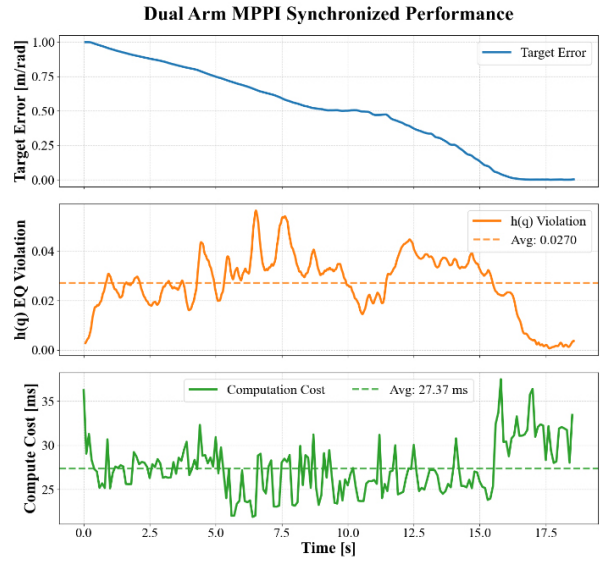
상단의 목표 추종 오차는 트레이 중심의 위치 및 자세 오차의 결합 지표로, 초기 약 1.0에서 시작하여 약 18초에 걸쳐 단조 감소하며 0에 수렴하였다. 약 8~10초 구간에서 오차 감소 속도가 일시적으로 완만해지는 구간이 관찰되는데, 이는 MPPI가 잠재 공간에서 비용 지형을 탐색하며 점진적으로 최적 궤적을 개선하는 특성에 기인한다.

중단의 등호 제약 위반은 8차원 제약 잔차의 노름 $\|h(q)\|$ 로 측정하였다. KKT 사영의 수렴 허용 오차는 2×10^{-3} 으로 설정되었으나, 실제 측정된 제약 위반은 평균 0.0270으로 이보다 크게 나타났다. 이는 KKT 사영이 계획된 참조 궤적에 대해 수행되는 반면, 실제 측정값은 하위 계층 PD 제어기의 참조 추종 오차가 반영된 실제 관절 구성 q 에서 평가되기 때문이다. 전 구간에 걸쳐 최대 약 0.06 이내로 제한되었으며 이는 VAE 기반 잠재 공간 샘플링이 제약 매니폴드를 효과적으로 근사하고 있음을 보여준다. 특히 목표 수렴 후(약 16초 이후)에는 잠재 공간에서의 이동이 감소함에 따라 제약 잔차가 급격히 감소하여 0에 근접하였다.

하단의 연산 시간은 MPPI 1회 반복당 평균 27.37 ms를 기록하였으며, 이는 약 36 Hz의 경로 최적화 주기에 해당한다. GPU 기반 병렬 연산을 통해 VAE 디코더 추론, KKT 사영, 비용 평가를 포함한 전체 파이프라인이 실시간에 준하는 속도로 동작함을 확인하였다.



[Fig. 2] Snapshots for target tracking experiments



[Fig. 3] Target tracking experiment results in an obstacle-free environment. (Top) Target tracking error of the tray center. (Middle) Equality constraint violation $\|h(q)\|$. (Bottom) Computation time per MPPI iteration

4.3 장애물 회피 실험

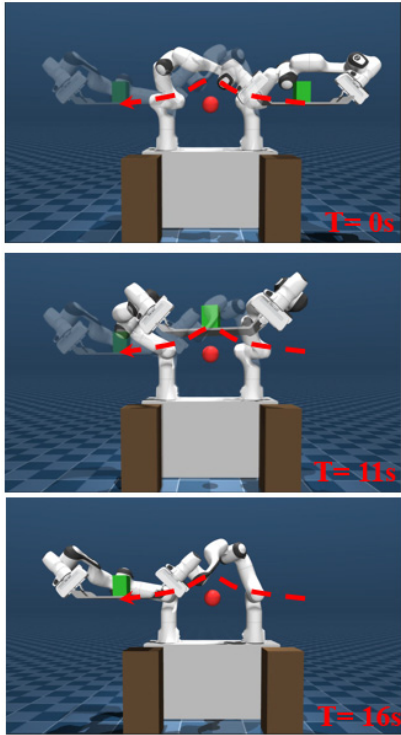
두 번째 실험에서는 Tray의 초기 위치에서 y 축 방향으로 -1.1 m 이동하는 작업을 수행하였다. 대표 스냅샷은 [Fig. 4]에 제시하였다. 특히 경로 상에 구형 장애물(반지름 0.05 m)을 배치하여, 목표 위치를 추종하면서도 장애물을 회피하는 과제를 수행하였다. [Fig. 5]는 [Fig. 3]과 동일한 성능 지표를 나타낸다. 목표 추종 오차는 초기 1.1에서 시작하여 약 18초에 목표에도 도달하였다. 장애물 회피를 위한 우회 경로 탐색으로 인해 약 3~5초 구간에서 오차 감소가 일시적으로 정체되는 현상이 관찰되었으며, 약 10초 부근에서 오차가 소폭 증가하는 구간은 장애물 인접 영역에서의 경로 재탐색에 기인한다. 그럼에도 불구하고 최종적으로 목표에 수렴하여, 장애물 존재 하에서도 안정적인 추종 성능을 확인하였다.

등호 제약 위반은 장애물 회피 시에도 평균 0.0260을 유지하였으며, 이는 장애물이 없는 실험(0.0270)과 동등한 수준이다. 이 결과는 장애물 회피를 위한 궤적 변경이 폐쇄 연쇄 제약 만족에 부정적 영향을 미치지 않음을 의미하며, 잠재 공간 샘플링이 제약 매니폴드 위에서의 탐색을 일관되게 수행하고 있음을 뒷받침한다.

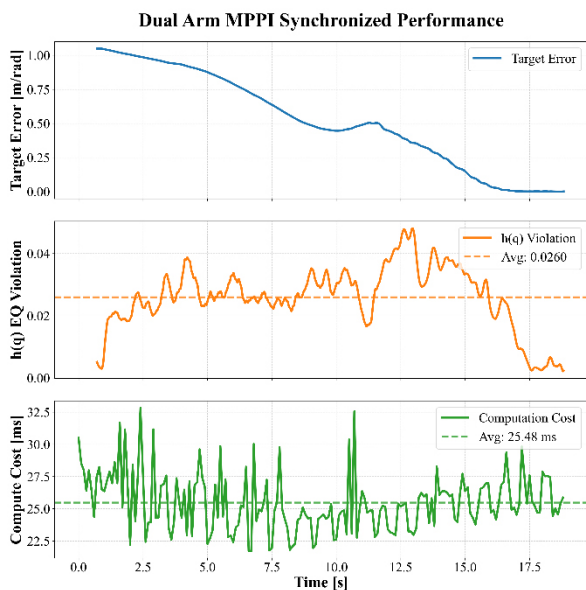
연산 시간은 평균 25.48 ms로, 장애물이 없는 실험(27.37 ms) 대비 오히려 소폭 감소하였다. 이는 장애물 회피 비용의 추가 연산 부담이 무시할 수 있는 수준이며, GPU 병렬 연산의 효율성에 의해 전체 연산 시간이 안정적으로 유지됨을 보여준다.

두 실험 결과를 종합하면, 제안한 VAE-MPPI 프레임워크는 장애물 유무와 무관하게 일관된 제약 만족 성능(평균 잔차 0.027

이하)과 안정적인 실시간 연산 성능(평균 26 ms 이내)을 달성하였으며, 폐쇄 연쇄 양팔 시스템에서 제약 조건을 만족하는 경로 최적화 성능을 입증하였다.



[Fig. 4] Snapshots for obstacle avoidance task



[Fig. 5] Obstacle avoidance experiment results in an obstacle environment. (Top) Target tracking error of the tray center. (Middle) Equality constraint violation $\|h(q)\|$. (Bottom) Computation time per MPPI iteration

5. 결론 및 논의

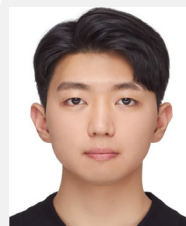
본 논문에서는 폐쇄 연쇄 양팔 로봇 조작을 위한 매니폴드 기반 모델 예측 경로 적분 제어 프레임워크를 제안하였다. 제안한 프레임워크는 VAE 기반 잠재 공간 샘플러를 통해 제약 매니폴드를 근사하고, KKT 기반 사영을 통해 각 샘플의 등호 제약 조건 만족을 보장한다. 또한, 계층적 제어 구조를 통해 10 Hz의 궤적 최적화와 1000 Hz의 관절 토크 제어를 분리함으로써 실시간 로봇 제어를 실현하였다. MuJoCo 시뮬레이션 환경에서의 실험을 통해, 장애물 유무와 무관하게 일관된 제약 만족 성능과 안정적인 실시간 연산 성능을 검증하였다.

본 연구는 몇 가지 한계점을 가진다. 첫째, MPPI는 샘플링 기반의 지역 최적화기로서 전역 경로 계획 능력이 제한적이며, 복잡한 장애물 환경에서는 지역 최솟값에 수렴할 가능성이 있다. 이를 해결하기 위해 전역 경로 계획기와의 통합이 향후 연구 과제로 남아있다. 둘째, VAE는 매니폴드를 근사적으로 학습하므로, 학습 데이터 분포를 벗어난 관절 구성에 대해서는 디코더 출력의 정확도가 저하될 수 있다. 향후 연구에서는 온라인 학습 또는 조건부 VAE (C-VAE)를 활용하여 이를 보완할 수 있을 것으로 기대된다. 셋째, 본 연구에서는 등호 제약 조건을 엄밀히 만족하는 기존 방법이 존재하지 않아 적합한 비교 기준선 선정이 어려웠으며, 이로 인해 정량적 비교 실험을 수행하지 못하였다. 넷째, 본 연구는 파지 구성이 고정되고 양팔 간 내부 응력이 발생하지 않는 이상적인 조건을 가정한다. 실제 로봇 적용 시에는 파지 위치 변화에 따른 VAE 일반화 성능 저하 및 KKT 사영 궤적 추종 오차로 인한 내부 응력 문제가 발생할 수 있으며, 이는 어드미턴스 제어 또는 힘/토크 센서 기반 피드백과의 통합을 통해 향후 연구에서 다룰 예정이다.

References

- [1] D. I. Park, C. Park, D. Kim, and J. Kyung, "Analysis and Design of the Dual Arm Manipulator for Rescue Robot," *Journal of Korea Robotics Society*, vol. 11, no. 4, pp. 235-241, Dec., 2016, DOI: 10.7746/jkros.2016.11.4.235.
- [2] T. Liu, Y. Lei, L. Han, W. Xu, and H. Zou, "Coordinated Resolved Motion Control of Dual-arm Manipulators with Closed Chain," *International Journal of Advanced Robotic Systems*, vol. 13, no. 3, Art. no. 80, May, 2016, DOI: 10.5772/63430.
- [3] M. Abbas, J. Narayan, and S. K. Dwivedy, "A systematic review on cooperative dual-arm manipulators: modeling, planning, control, and vision strategies," *International Journal of Intelligent Robotics and Applications*, vol. 7, no. 4, pp. 683-707, Jul., 2023, DOI: 10.1007/s41315-023-00292-0.
- [4] Y. Cui, Z. Xu, L. Zhong, P. Xu, Y. Shen, and Q. Tang, "A Task-Adaptive Deep Reinforcement Learning Framework for Dual-Arm Robot Manipulation," *IEEE Transactions on Auto-*

- tion Science and Engineering, vol. 22, pp. 466-479, Jan., 2024, DOI: 10.1109/TASE.2024.3352584.
- [5] L. Qian, L. Hao, S. Cui, X. Gao, X. Zhao, and Y. Li, "Research on Motion Trajectory Planning and Impedance Control for Dual-Arm Collaborative Robot Grinding Tasks," *Applied Sciences*, vol. 15, no. 2, pp. 819, Jan., 2025, DOI: 10.3390/app15020819.
- [6] J. Steen, G. Brandt, N. Wouw, J. Kober, and A. Saccon, "Quadratic programming-based reference spreading control for dual-arm robotic manipulation with planned simultaneous impacts," *IEEE Transactions on Robotics*, vol. 40, pp. 3341-3355, 2024, DOI: 10.1109/TRO.2024.3420800.
- [7] K. Jang, S. Kim, S. Park, and J. Park, "A Unified Framework for Overcoming Motion Constraints of Robots Using Task Transition Algorithm," *Journal of Korea Robotics Society*, vol. 13, no. 2, pp. 129-141, Jun., 2018, DOI: 10.7746/jkros.2018.13.2.129.
- [8] X. Zhao, Y. Zhang, W. Ding, B. Tao, and H. Ding, "A Dual-Arm Robot Cooperation Framework Based on a Nonlinear Model Predictive Cooperative Control," *IEEE/ASME Transactions on Mechatronics*, vol. 29, no. 5, pp. 3993-4005, Oct., 2024, DOI: 10.1109/TMECH.2023.3263357.
- [9] N. Ratliff, M. Zucker, J. A. Bagnell, and S. Srinivasa, "CHOMP: Gradient optimization techniques for efficient motion planning," *2009 IEEE International Conference on Robotics and Automation*, pp. 489-494, Jun., 2009, DOI: 10.1109/ROBOT.2009.5152817.
- [10] J. Schulman, J. Ho, A. X. Lee, I. Awwal, H. Bradlow, and P. Abbeel, "Finding locally optimal, collision-free trajectories with sequential convex optimization," in *Robotics: Science and Systems*, vol. 9, no. 1, pp. 1-10, 2013, DOI: 10.15607/RSS.2013.IX.031.
- [11] J. van den Berg, "Iterated LQR smoothing for locally-optimal feedback control of systems with non-linear dynamics and non-quadratic cost," in *2014 American Control Conference*, Portland, USA, pp. 1912-1918, 2014, DOI: 10.1109/ACC.2014.6859404.
- [12] Y. Wang, H. Li, Y. Zhao, X. Chen, X. Huang, and Z. Jiang, "A Fast Coordinated Motion Planning Method for Dual-Arm Robot Based on Parallel Constrained DDP," *IEEE/ASME Transactions on Mechatronics*, vol. 29, no. 3, pp. 2350-2361, Jun., 2024, DOI: 10.1109/TMECH.2023.3323798.
- [13] J. Lee, J. Lee, B. Yu, J. Kim, B. Cho, and M. Kim, "Trajectory generation algorithm based on differential dynamic programming considering center of mass for a 3-link mobile manipulator," *Journal of Korea Robotics Society*, vol. 20, no. 3, pp. 511-519, Sep., 2025, DOI: 10.7746/jkros.2025.20.3.511.
- [14] G. Williams, A. Aldrich, and E. A. Theodorou, "Model predictive path integral control: From theory to parallel computation," *J. Guid. Control Dyn.*, vol. 40, no. 2, pp. 344-357, Jan., 2017, DOI: 10.2514/1.G001921.
- [15] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, "Aggressive driving with model predictive path integral control," in *Proceedings of 2016 IEEE International Conference on Robotics and Automation*, Stockholm, Sweden, pp. 1433-1440, 2016, DOI: 10.1109/ICRA.2016.7487277.
- [16] M. Bhardwaj, B. Sundaralingam, A. Mousavian, N. D. Ratliff, D. Fox, F. Ramos, and B. Boots, "STORM: An Integrated Framework for Fast Joint-Space Model-Predictive Control for Reactive Manipulation," in *Proceedings of Conference on Robot Learning*, PMLR, London, UK, 2022, pp. 750-759, [Online] <https://proceedings.mlr.press/v164/bhardwaj22a.html>.
- [17] E. M. Andrejev, A. Manoharan, K.-E. Unt, and A. K. Singh, " π -MPPI: A projection-based model predictive path integral scheme for smooth optimal control of fixed-wing aerial vehicles," *IEEE Robotics and Automation Letters*, vol. 10, no. 6, pp. 344-357, Jan., 2025, DOI: 10.1109/LRA.2025.3568323.
- [18] L. Park, K. Jang, and S. Kim, "CSC-MPPI: A novel constrained MPPI framework with DBSCAN for reliable obstacle avoidance," in *Proceedings of 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Hangzhou, China, 2025, DOI: 10.48550/arXiv.2506.16386.
- [19] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proceedings of 2nd International Conference on Learning Representations (ICLR)*, Banff, Canada, 2014, DOI: 10.48550/arXiv.1312.6114.
- [20] A. L. Mitchell et al., "First steps: Latent-space control with semantic constraints," in *Proceedings of IEEE International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, pp. 5343-5350, 2020, DOI: 10.1109/IROS45743.2020.9340737.
- [21] J. Yun, M. Na, and J. Song, "Variational Autoencoder-based Assembly Feature Extraction Network for Rapid Learning of Reinforcement Learning," *Journal of Korea Robotics Society*, vol. 18, no. 3, pp. 352-357, Aug., 2023, DOI: 10.7746/jkros.2023.18.3.352.
- [22] E. Todorov, T. Erez, and Y. Tassa, "MuJoCo: A physics engine for model-based control," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura, Portugal, pp. 5026-5033, 2012, DOI: 10.1109/IROS.2012.6386109.



이슬찬

2024 경희대학교 기계공학과(공학학사)
2025~현재 경희대학교 기계공학과 석사박사
통합과정

관심분야: Robotic Manipulation, Optimal Control



김상현

2012 서울대학교 기계공학과(공학학사)
2020 서울대학교 기계공학과(공학박사)
2020~2023 한국기계연구원 선임연구원
2023~현재 경희대학교 기계공학과 조교수
2024~현재 차세대융합기술원 겸임연구원

관심분야: Optimal Control, Whole-body Control, Reinforcement Learning