

분해된 보행 패턴을 활용한 뱀 로봇 강화학습의 탐색 공간 축소

송봉섭*, 권우경, 남승우, 정운수
한국전자통신연구원 지능로봇시스템연구실

*sbs@etri.re.kr, wkwon@etri.re.kr, swnam@etri.re.kr, yoonsu@etri.re.kr

Reducing Exploration Space in Snake Robot RL using Decomposed Gaits

Bongsu Song*, Woogyong Kwon, Seung Woo Nam, Yun Su Chung
Electronics and Telecommunications Research Institute (ETRI)

요 약

본 연구는 뱀 로봇의 자연스러운 움직임에서 영감을 받아 강화학습의 탐색 공간을 축소하고 학습 효율을 향상시키는 Gait Decomposition(GD) 기법을 제안한다. 제안된 방법은 기존 곡선 기반 게이트를 시계열 토크 방향 모션 행렬로 분해하여, 생체 모방의 장점을 유지하면서도 초기 학습의 수렴 속도를 개선한다. 시뮬레이션을 통해 GD의 학습 효율성과 성능 향상 효과를 검증하였다.

I. 서 론

뱀 로봇은 모듈형 구조를 바탕으로 다양한 환경에서 높은 기동성과 유연성을 발휘한다 [1]. 최근에는 강화학습 기반 제어기 설계가 활발히 연구되고 있으나 [2], 고차원의 상태 및 행동 공간으로 인해 학습 속도가 저하되고 안정적 수렴이 어려운 한계가 존재한다. 본 연구는 생물학적 움직임에서 관찰되는 반복적·대칭적 게이트(gait) 특성을 수학적으로 분해함으로써 [3], 강화학습의 탐색 공간을 효과적으로 축소하는 새로운 접근법을 제안한다.

II. 제안된 방법론 (Methods)

$$\theta = \begin{cases} i = \text{odd}, & \theta_i = \beta_{\text{odd}} + A_{\text{odd}} \sin(\xi_{\text{odd}}) \\ i = \text{even}, & \theta_i = \beta_{\text{even}} + A_{\text{even}} \sin(\xi_{\text{even}} + \eta) \end{cases} \quad (1)$$

$$\begin{aligned} \xi_{\text{odd}} &= \psi_{\text{odd}}i + \nu_{\text{odd}}t \\ \xi_{\text{even}} &= \psi_{\text{even}}i + \nu_{\text{even}}t \end{aligned} \quad (2)$$

기존 Serpenoid 곡선 기반 게이트는 수식 (1), (2)와 같이 Dorsal 및 Lateral 방향의 관절 궤적을 시간에 따라 정의한다. 이에 반해 본 연구에서는 곡선을 시간-관절 토크 방향 모션 행렬로 변환하였다. 변환된 Gait Motion Matrix는 각 시점별 토크 방향을 -1 또는 +1로 이산화하여 표현하며, 이를 기반으로 토크 크기를 출력하는 신경망을 학습시킨다. 그림 1은 Serpenoid 함수와 토크 기반 임베딩 행렬 간의 변환 과정을, 그림 2는 Slithering gait를 패턴화하여 표현한 예시를 보여준다. 이 구조는 복잡한 함수 최적화 대신 간단한 토크 프로파일 학습만으로 다양한 게이트 유형을 유지할 수 있다는 장점을 갖는다.

III. 실험 결과

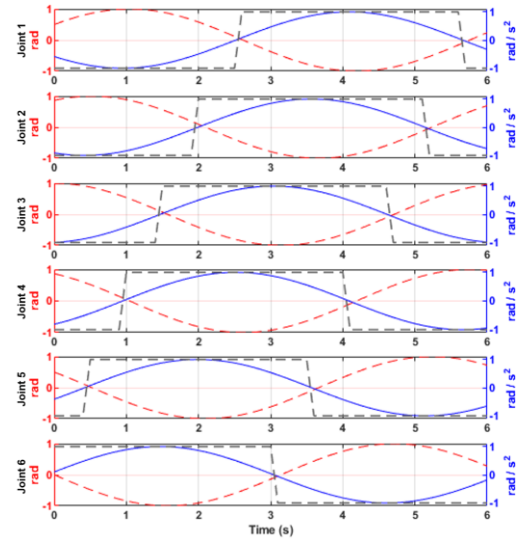


그림 1 Serpenoid 함수로부터 얻어진 관절의 토크 임베딩

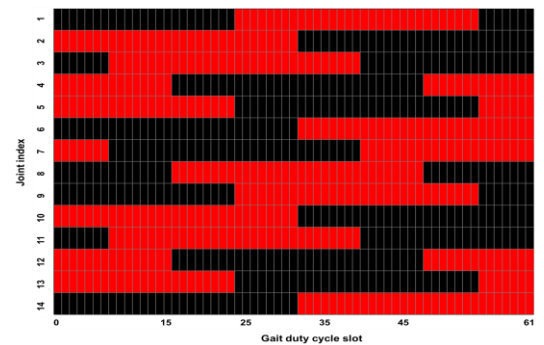


그림 2 Slithering Gait의 패턴 표현

그림 3은 임베딩 행렬을 기반으로 뱀 로봇의 토크 신호를 생성하는 알고리즘 개요를 제시한다. PPO와 SAC 알고리즘을 이용해 학습한 결과, GD를 적용한 정책은 기존 Serpenoid 기반 정책에 비해 빠르게 수렴하였으며, 그림 4에서 확인되듯이 에피소드 길이가

빠르게 증가하여 안정적인 제어가 가능함을 보였다. 또한 그림 5 와 같이 GD 기반 정책은 더 높은 보상 값을 달성하여 학습 효율이 개선됨을 확인하였다.

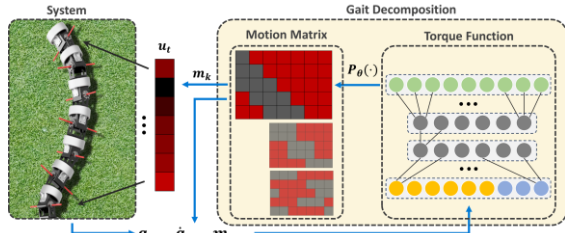


그림 3 GD 를 통해 뱀 로봇을 제어하는 알고리즘의 개요도

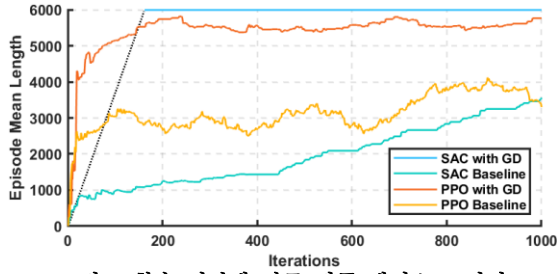


그림 4 학습 시간에 따른 평균 에피소드 길이

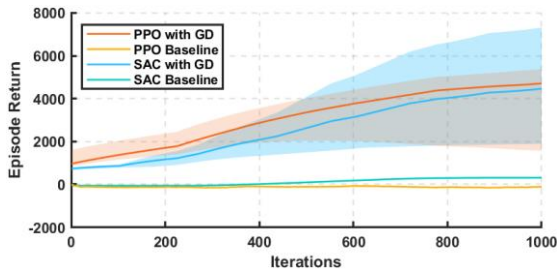


그림 5 학습 시간에 따른 평균 보상 값의 크기

IV. 결론

본 연구는 뱀 로봇의 생체 모방적 특성을 활용하여 강화학습의 효율성을 높이는 토크 기반 Gait Decomposition 기법을 제안하였다. 제안된 방법은 탐색 공간을 효과적으로 축소하여 학습 속도와 안정성을 향상시키며, 실제 적용 가능성을 높일 수 있다. 향후에는 다양한 지형에서의 일반화 성능 검증과 실시간 정책 업데이트 가능성에 대해 추가 연구를 진행할 예정이다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 과학기술사업화진흥원의 지원을 받아 수행된 연구임 ('학연협력플랫폼구축 시범사업' RS-2023-00304776).

참 고 문 헌

- [1] Wright, Cornell, et al. "Design and architecture of the unified modular snake robot." 2012 IEEE international conference on robotics and automation. Ieee, 2012.
- [2] Zhang, Dandan, et al. "One-shot domain-adaptive imitation learning via progressive learning applied to

robotic pouring." IEEE Transactions on Automation Science and Engineering 21.1 (2022): 541-554.

[3] Song, Bongsub, Insung Ju, and Dongwon Yun.

"Optimizing snake robot locomotion with decomposed gait pattern representation." Integrated Computer-Aided Engineering 32.2 (2025): 196-225.