

RLS 기반의 Natural Actor-Critic 알고리즘을 이용한 터널 환기제어기 설계

주백석* (고려대 기계공학과), 김동남(고려대 대학원 기계공학과), 홍대희(고려대 기계공학과),
박주영(고려대 제어계측공학과), 정진택(고려대 기계공학과), 김태형(한국건설기술연구원)

Tunnel Ventilation Controller Design Employing RLS-Based Natural Actor-Critic Algorithm

B. Chu* (Dept. of Mech. Eng., Korea Univ.), D. Kim(Dept. of Mech. Eng. Grad. School, Korea Univ.),
D. Hong(Dept. of Mech. Eng., Korea Univ.), J. Park(Dept. of Con. & Inst. Eng., Korea Univ.),
J. T. Chung(Dept. of Mech. Eng., Korea Univ.), T. -H. Kim(KICT)

ABSTRACT

The main purpose of tunnel ventilation system is to maintain CO pollutant and VI (visibility index) under an adequate level to provide drivers with safe driving condition. Moreover, it is necessary to minimize power consumption used to operate ventilation system. To achieve the objectives, the control algorithm used in this research is reinforcement learning (RL) method. RL is a goal-directed learning of a mapping from situations to actions. The goal of RL is to maximize a reward which is an evaluative feedback from the environment. Constructing the reward of the tunnel ventilation system, two objectives listed above are included. RL algorithm based on actor-critic architecture and natural gradient method is adopted to the system. Also, the recursive least-squares (RLS) is employed to the learning process to improve the efficiency of the use of data. The simulation results performed with real data collected from existing tunnel are provided in this paper. It is confirmed that with the suggested controller, the pollutant level inside the tunnel was well maintained under allowable limit and the performance of energy consumption was improved compared to conventional control scheme.

Key Words : Tunnel ventilation control (터널 환기제어), Reinforcement learning (RL, 강화학습), Actor-critic architecture, Natural gradient, Recursive least-squares (RLS, 회귀적 최소자승법)

1. 서론

터널 환기시설의 적절한 운전은 터널을 통과하는 운전자에게 쾌적한 환경과 안전한 운전조건을 제공한다. 그러나 이러한 터널 환기시설을 구동하기 위해서는 상당히 많은 양의 에너지가 소비되기 때문에 효과적인 운전 알고리즘이 도입된다면 에너지 감소 측면에서도 바람직한 결과를 얻을 수 있다.

터널 환기제어기 설계를 위해 이 연구에서는 강화학습 (RL)이라는 새로운 방법을 도입한다. 강화학습은 환경의 모델에 대한 정보 없이 주어진 상황에 대한 적절한 대응 방법을 찾는 목적 지향적 학습 방법이다. 강화학습의 목표는 외부로부터의 평가 정보인 '보상 (reward)' 신호를 최대화하는 것에 있다. 이 연구에서 보상 신호의 형태는 제어기의 목적과 직접적으로 연관이 되도록 터널 내부의 오염도 감소와 에너지 소비 최소화 사이의 적절한 조합으로 구성된다. 강화학습의 여러 분야 중에서 본 논문은 'actor-critic' 구조에 기초하고 있다. 그리고 데이터 사용의 효율성을 증가시키기 위해 학습 과정에 'recursive least-squares (RLS)' 방법을 적용하였

다. 특히 이 연구에 사용된 제어기는 'RLS-based natural actor-critic algorithm'에 기반하고 있다.[1]

2. 터널 환기시스템

본 연구의 대상이 되는 둔내 터널은 대한민국의 영동 고속도로에 위치하고 있다. 터널 내부의 상태를 측정하기 위하여 CO 센서와 VI 센서 그리고 풍속센서가 적절한 간격으로 배치되어 있으며 터널 입구에는 차량 계수기 (traffic counter)가 장착되어 있다. 터널 내 오염도를 감소시키기 위한 제어요소로는 총 32 대의 제트팬이 길이방향으로 설치되어 있다. 터널 내부의 오염도분포는 차량에 의한 오염물 유입을 오염원으로 가지는 1 차원 확산 이류 방정식으로 표현된다.[2]

3. 'Actor-critic' 구조를 이용한 강화학습

'Actor-critic' 모델은 'actor'와 'critic'으로 구분되는 두개의 모듈로써 이루어진다.[3] 'critic'은 성능평가를 위한 부분으로써 외부로부터의 평가 정보인 'reward' 신호와 현재 상태의 정보를 받아서 'actor'

의 성능개선을 위한 정보로 변형시켜 전달한다. 그리고 정확하고 빠르게 이 정보가 계산될수록 제어기의 성능은 더욱 신속히 개선된다. 이 연구에서는 빠른 수렴특성으로 인해 온라인 추정과 제어 분야에 적합하다고 알려진 RLS 방법을 적용하여 학습 데이터 사용의 효율을 향상시켰다.[1]

‘Actor’는 실제 제어신호를 출력하는 부분으로써 ‘critic’으로부터 얻은 정보와 상태정보를 이용하여 현재 상태에서부터 향후에 획득하게 될 ‘reward’ 신호의 총합을 최대화하는 방향으로 제어신호를 출력하고 또한 최적의 성능을 위하여 자신의 형태를 지속적으로 개선시킨다. 이 때 ‘policy gradient’ 방법을 이용하여 제어목표를 향해 가장 빠른 기율기로 추종하도록 제어가 조절된다. ‘Actor’의 제어 출력부를 개선하기 위한 더욱 효율적인 방법으로 이 연구에서는 ‘natural gradient’의 개념을 도입하여 가장 효율적이라고 알려진 조절식을 이용하였다.[4]

4. 시뮬레이션 결과

이 연구에서 제안된 알고리즘은 실제 터널에서 획득한 데이터를 이용했으며 컴퓨터 시뮬레이션을 통하여 증명되었다. 제어기의 입력 데이터로는 터널 내부의 CO 오염도와 차량의 대수 정보가 사용되었다. 출력은 터널 내부에 장착된 32 대의 제트팬으로써 가동대수를 적절히 조절하여 오염도를 적정 수준 (40ppm)으로 유지하고 동시에 불필요한 에너지를 최소화하게 된다.

Fig.1 은 제어입력이 가해지지 않은 상황의 터널 내부 오염도에 대한 3-D 결과를 나타낸다. 이 경우에는 최대 CO 오염도가 목표 값인 40ppm 을 상당히 초과하는 모습을 관찰할 수 있다. 반면에 이 연구에서 제안한 강화학습에 기반한 제어입력이 시스템에 인가되면 CO 오염도의 최소화과 에너지 소비의 감소라는 두 가지 제어목표가 효과적으로 달성됨을 Fig.2 를 통하여 확인할 수 있다. 이 때 충분한 학습이 이루어진 후 오염도가 목표 값인 40ppm 을 초과할 경우에는 제트팬의 가동대수를 증가시켜 오염도를 허용농도 이하로 낮추고, 반대로 오염도가 충분히 낮은 경우에는 가동대수를 감소시켜 에너지의 소비를 최소화하는 양상을 볼 수 있다.

5. 결론

터널 환기제어를 효과적으로 수행하기 위해 ‘actor-critic’ 구조와 ‘natural gradient’ 방법에 기초한 강화학습 알고리즘을 사용하였다. 또한, RLS 방법을 적용하여 학습 데이터 사용의 효율을 더욱 향상시켰다. 이로 인해 터널 내부의 오염도 감소와 에너지 소비 최소화라는 터널 환기시스템의 두 가지 제어 목표를 충분히 수행하였다. 그리고 기존의 많은

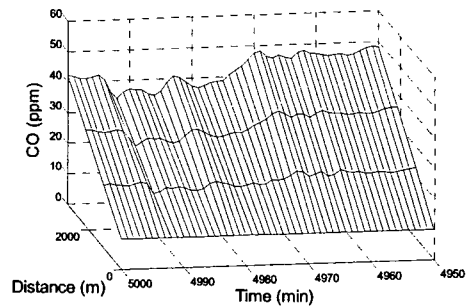


Fig. 1 3-D plot of the pollutant distribution for the last 50 time steps about ‘uncontrolled case’

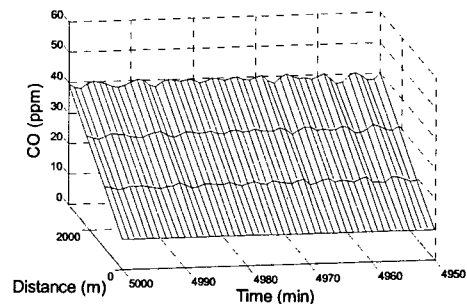


Fig. 2 3-D plot of the pollutant distribution for the last 50 time steps about ‘controlled case by RL-based controller’ when reference CO level is 40ppm

강화학습에 대한 연구가 전형적 문제에만 그 적용 분야가 국한된 반면 이 연구는 터널 환기시스템과 같은 실제적인 분야로 그 영역을 확대시켰다.

후 기

이 연구는 건설교통부가 출연하고 한국건설교통기술평가원에서 위탁시행한 2003 년도 건설핵심기술연구개발사업인 ‘지하공간 환경개선 및 방재기술 연구사업’ (03 산학연 C03-01)에 의한 것임

참고문헌

1. Park, J., Kim, J., and Kang, D., “An RLS-based natural actor-critic algorithm for locomotion of a two-linked robot arm,” LNAI, Vol. 3801, pp. 65-72, 2005.
2. Chu, B., Kim, D., Hong, D., Park, J., Chung, J. T., and Kim, T. -H., “GA-based fuzzy controller design for tunnel ventilation systems,” Proc. of ISARC, 2005.
3. Kaelbling, L. P., Littman, M. L., and Moore, A. W., “Reinforcement learning: a survey,” J. of Artificial Intelligence Research, Vol. 4, pp. 237-285, 1996.
4. Peters, J., Vijayakumar, S., and Schaal, D., “Reinforcement learning for humanoid robotics,” Proc. of IEEE-RAS Int. Conf. on Humanoid Robots, 2003.