

강화학습 기반 차량 냉방 제어 알고리즘 개발

한 서 희 · 최 완 식 · 안 창 선*

부산대학교 기계공학부

Development of Vehicle Cabin Cooling Controller Based on Reinforcement Learning

Seohee Han · Wansik Choi · Changsun Ahn*

School of Mechanical Engineering, Pusan National University, Busan 46241, Korea

(Received 26 July 2023 / Revised 29 August 2023 / Accepted 31 August 2023)

Abstract : To increase an electric vehicle's mileage, the energy efficiency of vehicles grows in importance every day. To this end, this study proposes an automotive air conditioner control system based on DQN(Deep Q-Learning), a reinforcement learning algorithm. This control system aims to reduce energy consumption with equivalent temperature error by controlling the compressor RPM and blower output. Also, the control system must improve its performance if there is a change in vehicle velocity profile. For this purpose, the state signals are designed to include thermal information and driving velocity, and the reward function is designed to reduce temperature error and power consumption and to suppress the undesirable behavior of the compressor. Training results show an improvement in energy efficiency when simulating with trained driving velocity profiles and other driving velocity profiles compared to the target A/C system.

Key words : Reinforcement learning(강화학습), Energy efficiency(에너지 효율), Air-conditioner(에어컨), Optimal control(최적제어), Electric vehicle(전기자동차)

1. 서 론

환경오염 및 지구온난화가 심각해짐에 따라 온실가스 배출 규제가 강화되면서 친환경 자동차에 대한 관심이 높아지고 있다.¹⁾ 또한 배터리 기술의 발전과 각 정부의 적극적인 지원에 따라 전기자동차의 판매량이 최근 들어 급격히 증가하고 있다.^{2,4)} 하지만 내연기관 자동차 대비하여 불편한 에너지 재충전 및 짧은 주행거리로 인하여 많은 소비자들이 여전히 구입을 망설이고 있다.

따라서 자동차 제작업체들은 주행거리 증가를 위해 자동차 구성품들의 에너지 효율을 향상시키려 노력하고 있다.⁵⁾ 특히 냉난방 시스템은 전기자동차에서 많은 전력을 소모하는 시스템이기 때문에, 자동차 업계에서는 냉난방 시스템의 설계 및 제어에 더욱 많은 연구 역량을 투입하고 있다. 학계에서도 탑승자의 냉난방 수요를 충족시키면서도 배터리 사용량을 최소화하는 제어 알고리즘을 개발하려는 연구가 급격히 늘어나고 있다. 많은 수의

연구가 기존에 널리 쓰이는 제어 방식을 답습하며 변수 조절을 통하여 효율을 향상시키려 하고 있다. 기존의 일반적으로 널리 쓰이는 Fuzzy logic,^{6,7)} PID,⁸⁾ 및 규칙 기반 제어기는 설계 및 구현이 간단하고 연산량이 적다는 장점을 가지고 있지만, 설계자의 튜닝 실력에 따라 성능이 달라지며 최적 성능을 보장하지 못한다는 단점을 가지고 있다.

설계자의 경험 및 능력에 영향을 많이 받는 제어 설계 방법론 보다, 체계적인 기법을 통한 성능향상이 가능한 최적 제어 기법을 이용하려는 많은 연구가 진행되고 있다. 동적 계획법(Dynamic programming)은 최적 제어 입력 값을 계산 가능하다는 장점을 가지고 있지만, 완벽한 미래정보가 필요하기 때문에 실시간 제어기로는 사용할 수 없다. 최근에는 실시간 제어기로 사용 가능하면서도 최적에 준하는 성능을 보장하는 모델 예측 제어 기법이 널리 쓰이고 있다.⁹⁻¹¹⁾ 모델 예측 제어는 최적 제어를 기

*Corresponding author, E-mail: sunahn@pusan.ac.kr

[†]This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

만으로 작동하여 규칙기반 제어기보다 뛰어난 성능을 보인다. 하지만 모델 예측 제어는 모델 정확도와 미래 정보의 예측 정확도에 따라 성능이 달라지며 최적화 연산을 위한 계산량이 많다는 단점이 있다.

최근 인공지능 기법의 발달로 제어기 설계에 여러가지 인공지능 기법을 활용한 사례가 보고되고 있다. 특히 최적 제어 기반 제어기 설계에는 단순하지만 정확성이 높은 제어모델이 요구되는데, 이를 인공지능 기법을 활용하여 모델을 수립한 경우가 있다.^{12,13)} 인공지능 기법을 활용하여 제어모델을 설계하는 것 이외에도 제어기 자체를 인공 신경망으로 설계하는 시도가 늘어나고 있다. 강화학습의 경우 마르코프 결정 프로세스(Markov decision process)를 기반으로 최적 제어 규칙을 추출하는 방법론으로서, 제어기 설계에 제어 모델을 필요로 하지 않는다는 장점이 있다.¹⁴⁾ 강화학습 기반의 제어기 설계를 위한 명시적인 미래 정보나, 제어 모델이 필요로 하지 않기 때문에 시스템 복잡도가 높은 열관리 시스템 제어기 설계에 적용하기 수월하다.

본 논문에서는 차량용 냉방 시스템의 압축기와 팬 속도를 조절하는 제어기를 강화학습 기반으로 설계하고 검증하도록 하겠다. 강화학습 관점에서 제어기 설계에 필요한 상태 변수, 보상함수(비용함수), 행동(제어입력)을 설계하는 과정을 소개하고 학습된 제어 알고리즘을 시뮬레이션 기반으로 검증하도록 하겠다.

2. 에어컨 시스템

본 연구를 위한 대상 냉방 시스템으로는 미국 National Renewable Energy Laboratory이 전기자동차용 냉방 시스템을 연구용으로 제작하고 이의 컴퓨터 시뮬레이션 모델로 공개한 CoolSim Electric A/C System이 사용되었다.^{15,16)} 해당 모델은 Simulink 기반으로 제작되었으며 실험데이터와 비교하여 유효성이 검증되었다.¹⁷⁾ 해당 모델은 Fig. 1과 같이 냉매사이클, 압축기, 경계 조건, 승객 공간으로 구성되어 있다.

냉매 사이클 모델은 응축기, 팽창밸브, 증발기로 구성 되어 있다. 압축기는 냉방 시스템의 주요 에너지 소비 장치이며, 차량 냉방 사이클의 중요한 역할을 한다. 경계 조건 모델에서는 차량 속도, 외기 온도 및 습도와 같은 주변 조건을 제공한다. CoolSim은 시스템 모델과 함께 시스템 작동을 위한 기본적인 제어기를 포함하고 있다. 기본 제어기는 Anti-windup 로직이 있는 비례-적분(PI) 형식으로 설계되어 있다. 냉각 대상인 승객실은 여러가지 열원이 존재하며, Fig. 2와 같이 외부와 차체 사이의 열전달, 차체와 실내 공기 사이의 열전달, 실내 인테리어

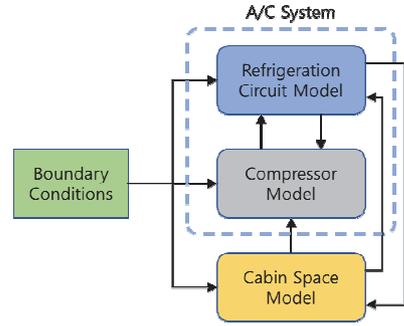


Fig. 1 Scheme of the cabin thermal model

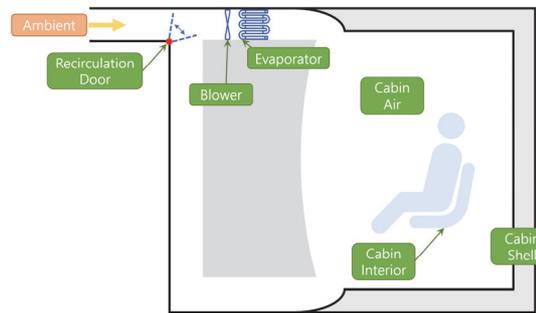


Fig. 2 Scheme of the cabin thermal model

와 실내 공기 사이의 열전달 등을 고려하여 모델링 되어 있다.

3. 강화학습 기반 제어기 설계

3.1 강화학습

강화학습은 기계학습의 한 분야로, 에이전트(Agent)가 환경(Environment)과 상호작용하면서 보상(Reward)을 최대화하기 위한 최적의 행동(Action)을 스스로 학습하는 방법이다. 에이전트는 상태(State)값을 관찰하고 환경으로부터 보상 값을 받으며 행동을 선택한다. 선택된 행동에 따라 환경이 변하게 되고 이에 따라 새로운 상태값과 보상 값이 결정된다. 이러한 상호작용을 통해 에이전트는 보상을 최대화하기 위한 최적의 행동을 학습하게 된다. 제어기 설계의 관점에서는 강화학습의 각 요소들을 다음과 같이 바라볼 수 있다. 에이전트는 제어기, 보상은 비용함수, 상태 값은 상태변수, 최적 행동은 최적 제어 입력의 역할을 하게 된다.

강화학습의 여러가지 방법 중 Deep Q Network(DQN)는 유한한 행동 공간을 가지는 환경에서 어떤 상태에서 특정 행동의 가치를 반환하는 함수인 Q 함수를 인공신경망이 근사하도록 경험 데이터를 통해 학습한다. 이를 통해 얻어진 제어기는 매시간마다 현재 상태값을 인공

신경망에 입력하여 모든 가능한 행동의 가치의 근사값을 반환 받은 다음 최대 가치를 가지는 행동을 선택하는 방식으로 동작한다. 여기서 가치는 어떤 상태에서 어떤 행동을 했을 때, 먼 미래까지 얻을 수 있는 보상 누적합의 기댓값이다. DQN은 제어 분야에 다양한 응용 사례가 존재하고, 행동 공간이 유한하여 학습 결과가 전역해에 가까운 장점이 있다. 따라서 본 연구에서도 DQN을 적용하였다.

DQN을 학습 할 때 행동을 결정하는 방법은 ϵ 의 확률로 무작위 행동을 선택하고 $(1-\epsilon)$ 의 확률로 학습된 가치를 바탕으로 가장 높은 가치를 가지는 행동을 선택하는 ϵ -greedy 방법을 사용하였다. 여기서 ϵ 값은 다음과 같은 규칙으로 매 스텝 감소하도록 설정하였다.

$$\epsilon = \epsilon(1 - \epsilon_{decay}) \quad \text{when } \epsilon \geq \epsilon_{min} \quad (1)$$

3.2 State, Action, Reward 설계

강화학습을 위해 상태 값 벡터는 실내 온도, 인테리어 온도, 증발기의 벽의 온도와 증발기를 통과하는 공기의 온도, 차량 속도, 지난 5초 동안의 압축기 회전속도의 변화 횟수로 구성하였다. 실내 온도와 응축기의 온도는 이후 차량 실내 온도에 직접적으로 영향을 줄 수 있는 요소이다. 차량의 속도는 응축기 주변의 공기속도에 영향을 주기 때문에 선정되었다. 압축기 회전 속도 변화 횟수는 보상함수에 사용하기 위하여 상태 값으로 선정하였다.

행동은 압축기의 회전속도와 송풍기(Blower) 유량으로 정의하였다. 압축기는 1500 ~ 7000 rpm 구간에서 작동하며, 송풍기 유량은 5.5 ~ 10.6 liter/min 작동 범위를 갖는다. 두 행동 값을 각각 6개, 4개로 이산화하여 Action set을 정의하였다.

강화학습의 결과물인 최적 행동 규칙은 보상함수 설계에 매우 큰 영향을 받는다. 개발하려는 제어기는 온도 성능을 유지하면서도 소모 에너지를 최소화하는 것이 목표이므로 식 (2)와 같이 보상 함수를 설계하였다.

$$R = -\alpha\Delta T^2 - \beta P - N_{rpm} \quad (2)$$

위 식에서 α 와 β 는 가중치 계수, ΔT 는 실내 온도의 목표 값과의 오차, P 는 소모 동력 N_{rpm} 은 압축기 속도의 변화 횟수이다. 이 보상함수의 첫 번째 항과 두 번째 항은 냉방 제어기의 냉방 성능 및 소모 전력과 관련된 항으로, 최적 제어 기반 냉방 제어기에 널리 사용되는 비용함수를 기반으로 설계하여 온도 오차가 감소하고 소모 전력이 작을수록 보상이 증가한다. 마지막 항은 이산화된

압축기 속도로 인해 발생할 수 있는 압축기의 심한 요동을 억제하기 위한 항으로, 압축기가 요동칠 경우 5초 동안 압축기 속도 변화 횟수 N_{rpm} 이 증가하여 보상이 감소한다. 즉, 위 보상 함수는 강화 학습을 통하여 압축기의 작동의 요동을 최소화하면서 에너지 소비를 줄이며 온도 요구조건을 맞추는 규칙을 찾도록 설계되었다.

3.3 에이전트 설계 및 학습

학습을 위해 에이전트는 인공신경망으로 설계되었다. 인공신경망은 3개의 층이 있으며 각 16, 64, 64개의 노드로 구성되어 있다. 또한 각 층 사이에는 Relu 활성화함수를 위치시켰다. 강화학습을 위해 필요한 초 매개변수(Hyper-parameter) 값은 Table 1을 통해 확인할 수 있다. 강화 학습 과정에 사용된 주행 사이클은 SC03 사이클을 사용하였고, 외기 온도는 35 °C, 상대 습도는 50 %로 설정하였다.

Table 1 Hyperparameter and values for reinforcement learning

Hyperparameter	Value
Sampling time	0.5 s
Minibatch size	128
Experience buffer	1,000,000
Target network update frequency	12,000
Discount factor (γ)	0.99
Initial epsilon (ϵ)	1
Epsilon decay (ϵ_{decay})	0.000001
Final epsilon (ϵ_{min})	0.002

4. 시뮬레이션 검증

학습된 제어기의 성능을 확인하기 위해, 학습에 사용된 주행 사이클 이외에 두개의 사이클을 추가하여 총 세 개의 주행 사이클을 사용하였다. 검증용 사이클은 Fig. 3에 나타낸 것과 같이 SC03과 다른 경향을 가지면서 길이가 비슷한 두 주행 사이클, 고속도로 주행 조건인 HWFET 주행사이클과 고속 및 급 가속 주행 조건인 US06 주행사이클을 추가적으로 선정하였다. 이 때, 외기 온도와 상대 습도는 학습 시 조건과 동일하게 35 °C 및 50 %로 설정하였다.

강화학습을 통하여 개발된 제어기의 성능을 검증하기 위하여, 학습된 제어기로 작동된 결과와 CoolSim에 장착된 PI 기반 기본 제어기로 작동된 결과를 비교하였다. 주행 사이클 SC03에서 검증된 결과를 Fig. 4에 도시하였다. 공정한 비교를 위하여 제어기 설계 변수를 달리 하면서 설계된 제어기들을 다수 사용하였다. 기본 제어기는 PI

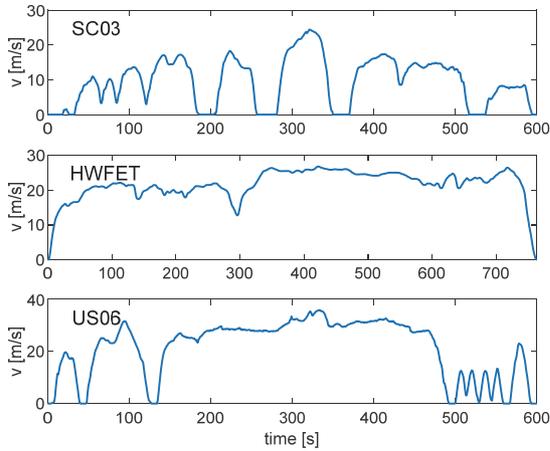


Fig. 3 Velocity profile for controller validation

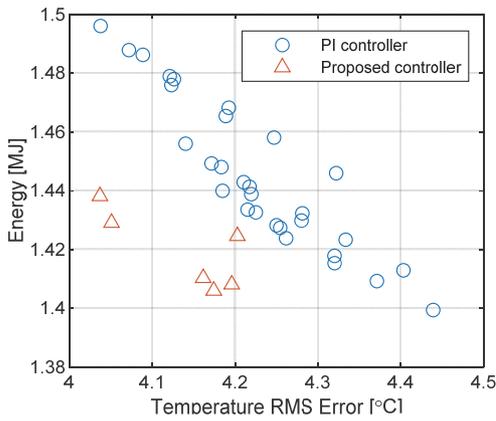


Fig. 4 Controller comparisons on SC03 cycle

제어 이득을 여러가지로 변경하며 성능을 평가하였으며, 강화학습을 통하여 개발된 제어기도 보상함수의 가중치 α 와 β 를 바꿔가면서 학습 후 검증하였다. 기본 제어기는 동그라므로, 강화학습으로 설계된 제어기는 삼각형으로 표시하였다. 가로 축은 목표 온도와 실내 온도와의 차이 값의 제곱평균제곱근(Root Mean Square, RMS)을 표시하고, 세로 축은 해당 주행 사이클에서 실내 냉방에 소모된 전기에너지를 표시한다. 즉 왼쪽 아래로 갈수록 작은 에너지를 소모하며 온도관리를 잘 수행하는 좋은 제어기이며 오른쪽 위로 갈수록 에너지 소모 및 성능 면에서 열세인 제어기를 의미한다.

기본 제어기는 PI 제어 이득을 변경하더라도 일정 수준 내에서 에너지 소모량과 온도 오차가 Trade-off 되며 총 성능의 변화가 거의 없다는 것을 확인할 수 있다. 강화학습을 통해 설계한 제어기는 기본 제어기보다 비슷한 온도오차 대비 적은 에너지를 사용하는 것을 확인할 수 있다. 동일한 온도 성능에서 두 제어기 사이의 소모

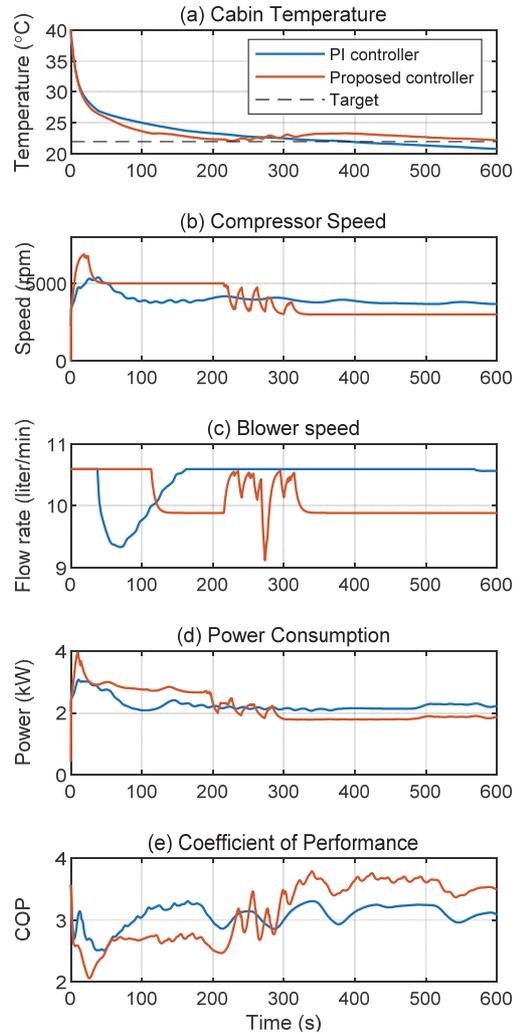


Fig. 5 Controller validation results on SC03 cycle

동력을 비교하면 제안하는 제어기는 기본 제어기 대비하여 에너지 소모량을 최대 4% 정도 감소시키고 있다.

SC03 사이클에서 대략적으로 4.2도 정도의 동일한 온도 오차 성능을 가지는 PI 제어기를 사용하는 기본제어기들과 RL 기반 제어기들 중 소모 전력이 낮은 제어기를 각각 선정하여, 두 제어기로 SC03 사이클 상에서 제어성능 검증 실험을 실시하였다. Fig. 5는 해당 검증 결과를 보여준다. Fig. 5(a)에서 보면, RL 기반으로 설계된 제어기는 빠르게 온도를 낮춘 후 22°C 부근에서 온도를 유지하는 것을 확인할 수 있다. 즉 일정 온도 이하로 내려간 경우에는 온도를 낮추는 것보다 전력을 적게 소모하는 것에 더 큰 비중을 두는 것을 확인할 수 있다. Fig. 5(b)에서는 제어 변수인 압축기 회전속도와 Blower 속도를 확인할 수 있다. 전체 냉방 시스템의 동력 소모량은 Fig. 5(d)를 통해 확인할 수 있으며 시스템의 효율을 확인하기

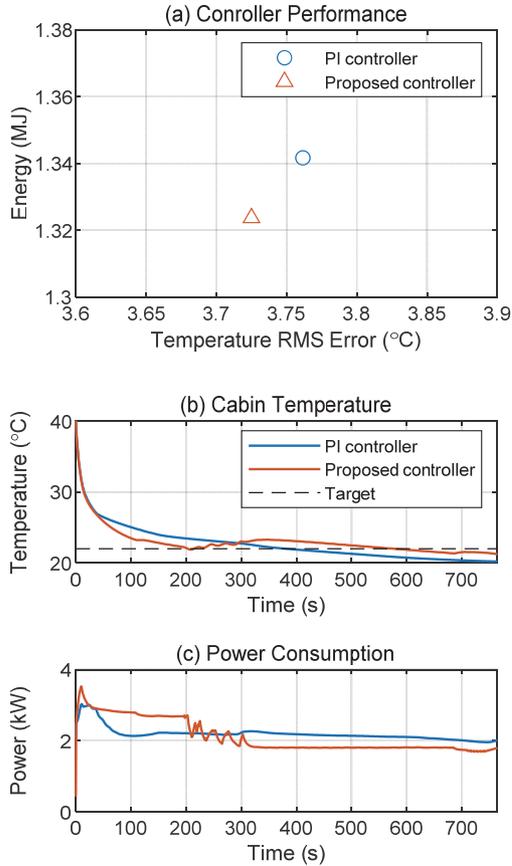


Fig. 6 Controller validation results on HWFET cycle

위해 PI 제어시스템과 RL 제어 시스템의 COP(Coefficient of performance)를 확인하였다. PI 제어기과 강화학습 제어기의 COP는 Fig. 5(e)에서 확인할 수 있다. 강화학습 제어 시스템의 경우 시뮬레이션 초반에 압축기의 회전속도를 높이고 일정 시간 이후에는 기존 제어기 대비하여 낮은 회전속도를 유지한다. 이러한 냉방방식은 초반 동력 소모량이 크고, 초반 COP는 낮지만, 일정 시간 이후로는 높은 COP를 보이며 일정 온도를 유지한다.

검증 신뢰도를 높이기 위해 학습에 사용되지 않은 추가적인 사이클에 사용하였다. SC03 사이클에서 동일한 온도 오차 성능을 보이는 기본제어기와 RL 기반 제어기를 HWFET 사이클에서의 검증한 결과가 Fig. 6에 나타나 있다. HWFET 사이클은 SC03에 비하여 가감속이 적고 고속 주행하는 시간이 많아, 에너지 소모량 감소 효과가 1.34%로 다소 적게 나타났다.

동일한 제어기들을 US06 사이클에서도 검증하였다. 시뮬레이션 결과를 Fig. 7에 도시하였다. 해당 주행사이클의 경우 SC03과 HWFET 사이클의 특성을 모두 지니고 있다. 따라서 에너지 사용량 감소량이 SC03과 HWFET 사이클의 경우의 사이 값인 1.83%로 측정되었

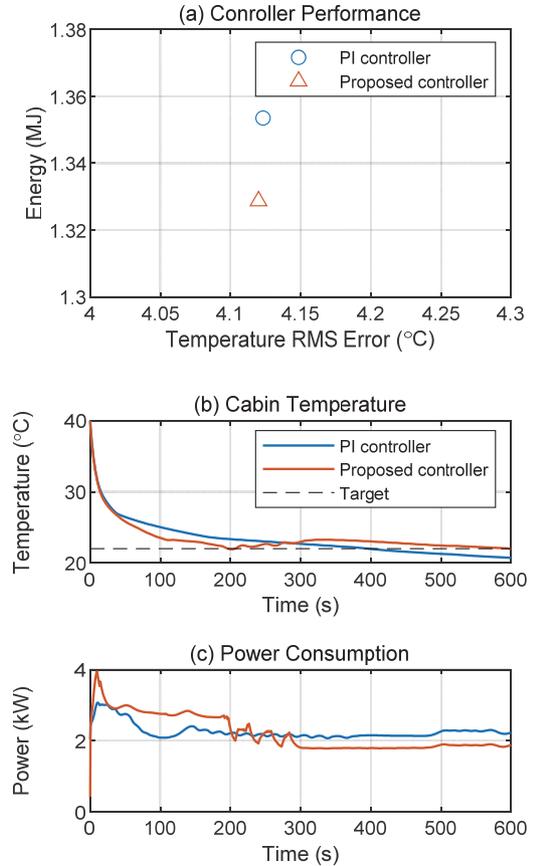


Fig. 7 Controller validation results on US06 cycle

다. 주행 속도의 차이로 인해 냉각 시스템의 성능 차이가 발생하게 되고 이로 인해 전체적으로 사용되는 에너지와 온도 오차가 달라진다. SC03에 비해 HWFET와 US06 사이클 모두 속도가 빠르고 정지 구간이 짧기 때문에 응축기로 들어가는 공기의 평균 속도가 빨라 전체적인 냉각 성능이 우수해 적은 에너지를 사용하는 것으로 분석된다.

5. 결론

- 1) 강화학습 방법 중 DQN 기반 전자동차의 실내 냉방 시스템 제어 알고리즘을 개발함
- 2) DQN의 이산화된 행동 공간으로 인해 발생할 수 있는 압축기 요동을 억제하면서 냉방 성능을 향상시킬 수 있는 보상함수를 설계함
- 3) 동등한 온도 제어 성능의 PI 제어기 대비 에너지 효율 향상

- SC03(학습 주행 사이클): 4% 향상
- HWFET(검증 주행 사이클1): 1.3% 향상
- US06(검증 주행 사이클2): 1.8% 향상

- 4) 더 다양한 외기 조건, 목표 온도 및 주행 사이클 조건에 대해 학습을 진행하면 에너지 효율이 높고 강건한 제어기 설계가 가능할 것으로 기대됨

후 기

이 과제는 부산대학교 기본연구지원사업(2년)에 의하여 연구되었음.

References

- 1) L. Nicoletti, M. Brönnner, B. Danquah, A. Koch, A. König, S. Krapf, A. Pathak, F. Schockenhoff, G. Sethuraman, S. Wolff and M. Lienkamp, "Review of Trends and Potentials in the Vehicle Concept Development Process," 2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), pp.1-15, 2020.
- 2) G. Harper, R. Sommerville, E. Kendrick, L. Driscoll, P. Slater, R. Stolkin, A. Walton, P. Christensen, O. Heidrich, S. Lambert, A. Abbott, K. Ryder, L. Gaines and P. Anderson, "Recycling Lithium-ion Batteries from Electric Vehicles," Nature, Vol.575, No.7781, pp.75-86, 2019.
- 3) J. W. Kim and Y. G. Shin, "Temperature Management of EV Battery Cell by Optimal Operation Scheduling," Transactions of KSAE, Vol.27, No.7, pp.509-519, 2019.
- 4) M. Jafari, A. Gauchia, S. Zhao, K. Zhang and L. Gauchia, "Electric Vehicle Battery Cycle aging Evaluation in Real-World Daily Driving and Vehicle-to-grid Services," IEEE Transactions on Transportation Electrification, Vol.4, No.1, pp.122-134, 2017.
- 5) J. Gu, M. Ouyang, D. Lu, J. Li and L. Lu, "Energy Efficiency Optimization of Electric Vehicle Driven by In-wheel Motors," Int. J. Automotive Technology, Vol.14, No.5, pp.763-772, 2013.
- 6) B. S. K. K. Ibrahim, M. A. N. Aziah, S. Ahmad, R. Akmeliawati, H. M. I. Nizam, A. G. A. Muthalif, S. F. Toha and M. K. Hassan, "Fuzzy-based Temperature and Humidity Control for HV AC of Electric Vehicle," Procedia Engineering, Vol.41, pp.904- 910, 2012.
- 7) H. Nasution, "Development of Fuzzy Logic Control for Vehicle Air Conditioning System," TELKOMNIKA (Telecommunication Computing Electronics and Control), Vol.6, No.2, pp.73-82, 2008.
- 8) J. Wang, C. Zhang and Y. Jing, "Application of an Intelligent PID Control in Heating Ventilating and Air-conditioning System," 2008 7th World Congress on Intelligent Control and Automation, pp.4371-4376, 2008.
- 9) A. Afram and F. Janabi-Sharifi, "Theory and Applications of HVAC Control Systems—A Review of Model Predictive Control (MPC)," Building and Environment, Vol.72, pp.343-355, 2014.
- 10) R. Shah, B. P. Rasmussen and A. G. Alleyne, "Application of a Multivariable Adaptive Control Strategy to Automotive Air Conditioning Systems," International Journal of Adaptive Control and Signal Processing, Vol.18, No.2, pp.199-221, 2004.
- 11) Q. Zhang, L. Fiorentini and M. Canova, "H ∞ Robust Control of an Automotive Air Conditioning System," 2014 American Control Conference, pp. 5675-5680, 2014.
- 12) B. C. Ng, I. Z. M. Darus, H. Jamaluddin and H. M. Kamar, "Application of Adaptive Neural Predictive Control for an Automotive Air Conditioning System," Applied Thermal Engineering, Vol.73, No.1, pp.1244-1254, 2014.
- 13) A. Afram, F. Janabi-Sharifi, A. S. Fung and K. Raahemifar, "Artificial Neural Network(ANN) based Model Predictive Control(MPC) and Optimization of HVAC Systems: A State of the Art Review and Case Study of a Residential HVAC System," Energy and Buildings, Vol.141, pp.96-113, 2017.
- 14) V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra and M. Riedmiller, "Playing Atari with Deep Reinforcement Learning," arXiv Preprint arXiv:1312.5602, 2013.
- 15) H. Wang, I. Kolmanovsky, M. R. Amini and J. Sun, "Model Predictive Climate Control of Connected and Automated Vehicles for Improved Energy Efficiency," 2018 Annual American Control Conference(ACC), pp.828-833, 2018.
- 16) T. Kiss, L. Chaney and J. Meyer, "New Automotive Air Conditioning System Simulation Tool Developed in Matlab/simulink," National Renewable Energy Lab.(NREL), Golden, CO (United States), 2013.
- 17) T. Kiss and J. Lustbader, "Comparison of the Accuracy and Speed of Transient Mobile A/C System Simulation Models," National Renewable Energy Lab.(NREL), Golden, CO (United States), 1946-4002, 2014.