creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

공학석사 학위논문

조도 예측 모델 기반의 DDPG 강화학습을 이용한 실내 조명 시스템 제어

Optimal Control of Lighting System based on Illuminance Prediction Model with Deep Deterministic Policy Gradient (DDPG)

2021년 2월

서울대학교 대학원

건축학과

김 영 섭

조도 예측 모델 기반의 DDPG 강화학습을 이용한 실내 조명 시스템 제어

Optimal Control of Lighting System based on Illuminance Prediction Model with Deep Deterministic Policy Gradient(DDPG)

지도교수 박 철 수

이 논문을 공학석사 학위논문으로 제출함
2021년 2월

서울대학교 대학원
건축학과
김 영 섭

김영섭의 석사 학위논문을 인준함
2021년 2월

위 원 장 _____ 여 명 석 _____ (인)

부위원장 _____ 박 철 수 _____ (인)

위 원 _____ 김 선 숙 _____ (인)

국문초록

실내 공간에서는 작업의 효율과 안전을 위해 적정 수준의 조도 환경을 확보해야 한다. 채광창을 통해 자연채광이 유입될 경우, 부족한 만큼만 전기조명을 통해 공급함으로 조명 에너지를 절감할 수 있다. 자연채광은 시간별 태양의 위치 및 기상환경에 따라 동적으로 변한다. 즉, 전기조명 시스템의 제어변수(디밍 제어의 경우: 0~1)를 결정하기 위해서는, 제어 시점마다 자연채광으로 인한 작업면 조도의 동적 거동을 독립적으로 파악할 수 있어야 한다. 또한, 각 지점에 부족한 만큼의 전기조명 조도를 공급하기 위해서는 제어변수에 따라 여러 조명기구에서 도달하는 광속을 알아야 한다. 조명기구별 출력과 작업면 조도의 관계를 시스템 모델로 작성하면, 자연채광 현황에 따라 권장 조도 환경을 만족하는 동시에 전력사용량을 최소화하는 전기조명 시스템의 제어변수를 결정할 수 있다.

작업면의 조도는, 센서를 설치하거나 시뮬레이션 모델로 예측하여 알 수 있다. 조도 센서를 이용하는 방법은 권장 조도 수준의 만족 여부에 대한 정확도가 높고, 직관적이다. 하지만, 센서의 측정값은 자연채광과 여러 전기조명기구의 조도가 합해진 것으로, 이를 분리하는 것이 어렵다. 또한, 조도 센서는 작업면에 설치해야 하지만, 간섭으로 인해 주변의 벽이나 천장에 설치하게 되며, 이로 인해 정확도가 떨어진다. 조도 시뮬레이션을 이용하면 이와 같은 센서의 물리적인 한계를 극복할 수 있다. 본 연구에서는 인천의 한 공장 건물을 대상으로 빛 환경 시뮬레이션 도구인 Radiance 기반의 빛 환경 시뮬레이션 도구인 DIVA for Grasshopper를 이용하여 자연채광과 전기조명 조도를 예측하는 물리법칙 기반의 모델을 각각 개발하였다.

제작된 모델을 제어에 이용하기 위해서는 실시간으로 예측 결과를 도출할 수 있어야 한다. 자연채광 모델은 기상청의 수평면 총일사량을 입력으로 이용하는데, 이를 실시간으로 이용하기 어렵고, 이용할지라도 기상대와 대상 건물의 지역적 차이로 인한 오차가 있다. 모델의 연산시간 또한 해결해야 한다. Radiance에서는 Backward ray-tracing 방법으로 측정지점으로 유입

되는 다중 반사광(interreflection)을 계산한다. 빛을 광선으로 가정하고, 추적 방향을 샘플링하여 확률적으로 계산하기에 샘플 개수와 반사 횟수가 많을수록 정확도가 높아지고 연산시간이 길어진다. 자연채광 모델의 연산시간을 단축하고, 지역적 특징을 반영하는 참조 조도를 모델의 입력변수로 활용하기 위해 기계학습 대리모델을 개발하였다. 전기조명 모델의 경우, 조명기구 그룹별 시뮬레이션을 수행한 결과로 디밍 제어변수에 대한 선형시스템 모델을 제작하였다. 각 모델은 작업면에 임시로 조도 센서를 설치하여 수집한 데이터로 보정 및 검증하였다.

제어 최적화는 강화학습으로 구현하였다. 강화학습은 인공지능의 한 분야로, 주어진 환경에서 제어변수를 결정하고, 그 결과에 따라 보상을 받으며 최적제어 방법을 탐색한다. 본 연구에서 적용한 심층 확정적 정책 그래디언트(DDPG) 알고리즘은 디밍 제어와 같은 연속 제어변수를 결정할 수 있다. 또한, 제어변수가 많을 경우, 확률적 알고리즘보다 우수하다. 훈련 기간을 통해 도출한 최적 정책 신경망(Optimal Policy Network)으로 제어를 수행한 결과, 평균조도와 균제도를 권장수준으로 확보할 수 있었다. 전력 사용량의 경우, 기존의 스케줄 기반 제어 대비 38%를 절감할 수 있었다.

강화학습으로 결정된 제어 시점별 변수 패턴은 선형계획법 최적화 알고리즘의 패턴과 유사하다. 자연채광이 부족한 지점에 집중하여 전기조명을 공급하였으며, 시간 및 날짜별 자연채광 수준에 따라 제어를 수행하였다. 하지만, 선형계획법은 실내 자연채광이 부족한 아침과 저녁 시간의 경우, 제어 패턴에 불안정성을 보였으며, 각 지점의 조도를 최대한 확보하기 위해 보수적으로 제어가 수행된다. 해당 시점에서는 어떤 방법의 조명 시스템 제어로도 선형계획법의 제약조건을 만족할 수 없었기 때문이다. 한편, 강화학습은 이와 같은 환경에서도 일관적인 제어를 수행하였으며, 전력사용량과 조도 환경의 균형이 유지되었다. 결과로 선형계획법보다 17% 더 많은 전기에너지를 절감할 수 있었다.

본 연구는 실제 건물을 대상으로 조도예측모델을 제작하여 실시간 전기조명 제어를 수행하였다. 모델을 통해 작업면의 조도 센서를 대체할 수 있었으며, DDPG 강화학습 제어의 최적화 성능 및 안정성을 확인하였다. 이

를 통해 조도예측모델에 기반한 실내 조명 시스템 제어가 유용함과 동적 자연채광 환경 및 연속 제어변수에 대한 DDPG 강화학습의 가능성을 전달하고자 한다.

주요어 : 조도 예측 모델, Radience 시뮬레이션, 실시간 조명 제어, 강화학습, 심층 확정적 정책 그래디언트(DDPG)
학 번 : 2019-21516

목 차

제 1 장 서 론 ····· 1
1.1 연구의 배경 및 목적 ····· 1
1.2 연구의 범위 및 방법 ····· 4

제 2 장 조명 시스템 제어 ····· 6
2.1 조도 센서 기반 조명 시스템 제어 ····· 6
2.2 모델 기반 조명 시스템 제어 및 최적화 ····· 9
2.2.1 조도 예측 모델 ····· 9
2.2.2 시스템 제어 및 최적화 ····· 12

제 3 장 강화학습 알고리즘 ····· 16
3.1 강화학습 개요 ····· 16
3.1.1 마르코프 결정 과정(MDP) ····· 16
3.1.2 가치함수 ····· 18
3.2 정책 그래디언트 강화학습 ····· 20
3.2.1 Reinforce 알고리즘 ····· 20
3.2.2 Actor-Critic 구조 ····· 22
3.3 심층 확정적 정책 그래디언트(DDPG) ····· 23
3.4 소결 ····· 27

제 4 장 조도 예측 모델 개발 ····· 29
4.1 대상 건물 및 조명 시스템 ····· 29
4.2 자연채광 조도 예측 모델 개발 ····· 32
4.2.1 자연채광 모델 작성 ····· 32

4.2.2 자연채광 모델 보정 및 검증 ····· 35
4.2.3 자연채광 대리모델 작성 ····· 40
4.3 전기조명 조도 예측 모델 개발 ····· 43
4.4 모델 통합 및 검증 ····· 48

제 5 장 DDPG 강화학습 조명 제어 ····· 53
5.1 강화학습 환경(environment) 제작 ····· 53
5.1.1 에피소드 ····· 53
5.1.2 보상함수 ····· 55
5.2 DDPG 훈련 및 검증 ····· 59
5.3 DDPG를 이용한 조명 시스템 제어 결과 ····· 61
5.3.1 DDPG 제어 결과 ····· 61
5.3.2 기존 제어 방법 및 DDPG 제어 성능 비교 ····· 70

제 6 장 결론 ····· 79

참고문헌 ····· 83
Abstract ····· 89

표 목 차

[표 4-1] 그룹별 조명기기 개수 및 전력 사용량 ····· 31
[표 4-2] 미지변수 추정 범위 및 보정값 ····· 38
[표 4-3] 모델 검증 기간, 천공 상태 및 모델 정확도 ····· 48
[표 5-1] 3월 13일-16일, 천공 상태 ····· 65
[표 5-2] 3월 13일, 각 제어 시점별 강화학습 제어변수 ····· 66
[표 5-3] 3월 14일, 각 제어 시점별 강화학습 제어변수 ····· 67
[표 5-4] 3월 15일, 각 제어 시점별 강화학습 제어변수 ····· 68
[표 5-5] 3월 16일, 각 제어 시점별 강화학습 제어변수 ····· 69
[표 5-6] 3월 13일-16일, 여러 가지 제어 방법의 성능 비교 70
[표 5-7] 3월 13일, 각 제어 시점별 선형계획법 제어변수 ·· 75
[표 5-8] 3월 14일, 각 제어 시점별 선형계획법 제어변수 ·· 76
[표 5-9] 3월 15일, 각 제어 시점별 선형계획법 제어변수 ·· 77
[표 5-10] 3월 16일, 각 제어 시점별 선형계획법 제어변수 78

그 림 목 차

[그림 1-1] 하루 동안 실내·외에 머무는 시간 비율 ····· 1
[그림 1-2] 연구의 절차 및 구성 ····· 5
[그림 2-1] 조도센서 기반 전기조명 제어 방법 ····· 7
[그림 2-2] 다수의 조명기기 및 작업면 사이의 관계 ····· 8
[그림 2-3] 빛 환경 시뮬레이션 알고리즘 ····· 10
[그림 3-1] 마르코프 결정 과정 ····· 17
[그림 3-2] Deep Deterministic Policy Gradient 학습흐름도 25

[그림 4-1] 대상 건물 ····· 29
[그림 4-2] 조명기기 그룹(#1-#10) 및 평면도 ····· 30
[그림 4-3] DIVA 조도 예측 모델 ····· 34
[그림 4-4] 모델 보정 및 검증용 조도 센서 설치 위치 ····· 36
[그림 4-5] 대상 구역 부위별 재질 및 유리관리 현황 ····· 37
[그림 4-6] 자연채광 조도 예측 모델 제작 방법 ····· 39
[그림 4-7] ANN 모델의 구조 ····· 40
[그림 4-8] 자연채광 조도 센서(E_r) ····· 41
[그림 4-9] DIVA 시뮬레이션 vs ANN 대리 모델 비교 ···· 42
[그림 4-10] 조명기구별 배광 곡선 ····· 43
[그림 4-11] 총 조도의 실측값 vs 예측값 비교 ····· 50
[그림 4-12] 자연채광 조도의 실측값 vs 예측값 비교 ····· 52
[그림 5-1] 환경과 에이전트의 상호작용 ····· 54
[그림 5-2] 각 지점 조도에 따른 획득 점수 ····· 56
[그림 5-3] DDPG 학습 곡선(에피소드별 반환) ····· 59
[그림 5-4] 3월 13일 12시 광원별 작업면 조도 ····· 62
[그림 5-5] 3월 13일, DDPG 최적 정책의 조명기기 그룹별 디밍 제어변수 및 작업면 조도 ····· 64
[그림 5-6] 3월 13일, 선형계획법의 조명기기 그룹별 디밍 제어변수 및 작업면 조도 ····· 74

제 1 장 서 론

1.1 연구의 배경 및 목적

인간의 무분별한 개발 행위 및 석유 에너지 이용으로 인해 지구 환경이 파괴되고 있다. 기후 변화로 인한 자연재해와 식량자원 감소 등 그 피해는 인간에게 돌아오고 있다. 이에 전 세계 국가들은 기후변화협약을 통해 온실가스 농도를 지구 환경이 지속 가능한 수준으로 유지하는 데에 합의하였다. 2019 국가 온실가스 인벤토리 보고서에 따르면 온실가스의 86.8%가 에너지 분야에서 발생하며, 건물은 산업, 수송과 함께 국가 에너지 사용량의 가장 큰 부분을 차지하고 있다. 사람은 86.9%의 시간을 건물 안에서 지내며 (그림 1-1), 활동에 필요한 환경(온도, 습도, 조도 CO2 농도 등)을 쾌적한 수준으로 확보하는 데에 에너지를 사용한다.

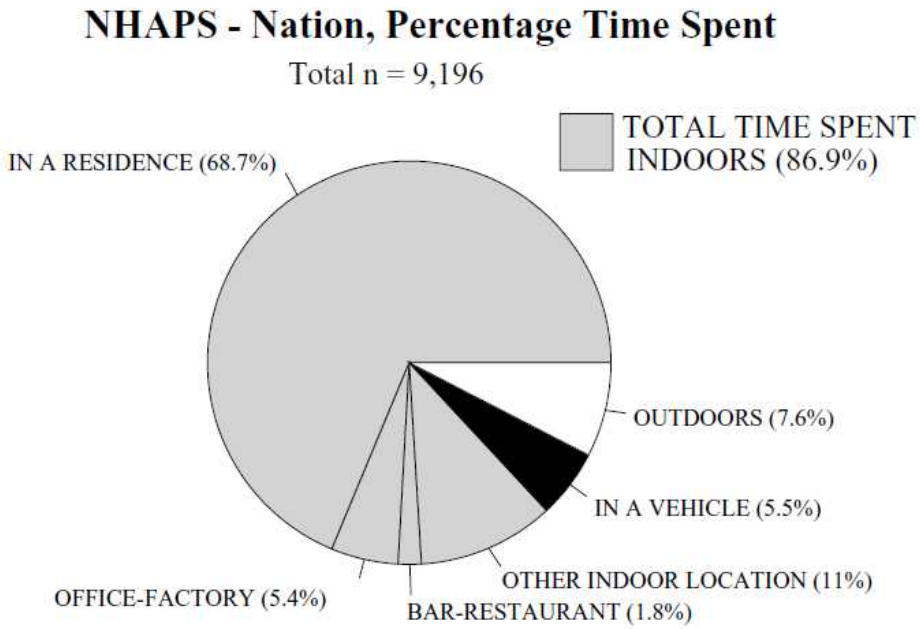


그림 1-1. 하루 동안 실내·외에 머무는 시간 비율 (Klepeis et al., 2001)

조도 환경은 작업 효율 및 안전을 위해 확보되어야 한다. 공간의 용도 및 작업 유형에 따라 필요한 조도 환경이 다르며, 수평, 수직면 조도 및 균제도로 이를 평가할 수 있다. 조도는 평가 지점 주변의 물리적 형태, 재질의 물성값(반사율, 투과율)으로 인해 같은 실에서도 측정지점마다 모두 다르다. 즉, 실질적인 작업이 이루어지는 지점의 조도 환경을 확보하는 것이 중요하다. 실내 공간의 조도 환경은 창을 통해 유입되는 자연채광과 전기조명의 조도로 구성된다. 자연채광 조도는 태양과 측정지점의 위치 관계 및 천구의 운량에 따라 지속해서 변한다. 이에 비해, 전기조명 조도는 조명기기의 배치 및 배광곡선, 보수율에 따라 결정되며, 시간에 따른 변화가 작다.

건물 실내 각 지점의 조도를 광원에 따라 예측할 수 있으면, 여러 가지 제어 방법(디밍 제어, 온·오프 제어, 격등 제어 등)을 적용하여 조명 에너지를 절감할 수 있다. 자연채광으로 필요한 조도 환경을 만족하지 못할 경우, 부족만 만큼을 전기조명 시스템을 통해 제공하는 것이다. 이때, 주간 자연채광 조도는 시간에 따라 지속적으로 변하기 때문에 실시간으로 파악해야 한다. 또한, 전기조명 시스템의 제어 방법 역시 실시간으로 결정할 수 있어야 한다.

작업면의 조도를 파악하기 위해, 조도 센서를 설치하는 방법과 시뮬레이션 모델을 이용하는 방법이 있다. 조도 센서를 이용하는 방법은 간편하고 정확하다. 하지만, 측정된 값은 자연채광과 전기조명 조도가 합해진 것으로 이를 분리하는 것이 어렵다. 또한, 센서의 설치 지점에 주의해야 한다. 조도 센서를 작업면에 설치하는 것이 가장 좋지만, 사람의 행동으로 인한 간섭으로 인해 주변의 벽이나 천장에 설치한 다음 보정하여 이용한다. 한편, 조도 시뮬레이션을 수행하기 위해서는 물리법칙 기반의 3D 모델을 제작해야 한다. 모델을 이용하면, 자연채광 및 인공조명 등의 광원에 따른 조도를 분리할 수 있다. 또한, 원하는 모든 지점의 조도를 작업 활동의 간섭 없이 구할 수 있다.

자연채광 조도 현황에 따라 제어변수를 결정하는 방법으로는 최적제어와 강화학습 등이 있다. 최적제어는 작업면의 조도 환경을 제약조건으로 하여 에너지 사용량을 최소화하는 알고리즘을 적용하는 것이다. 한 번에 제어변

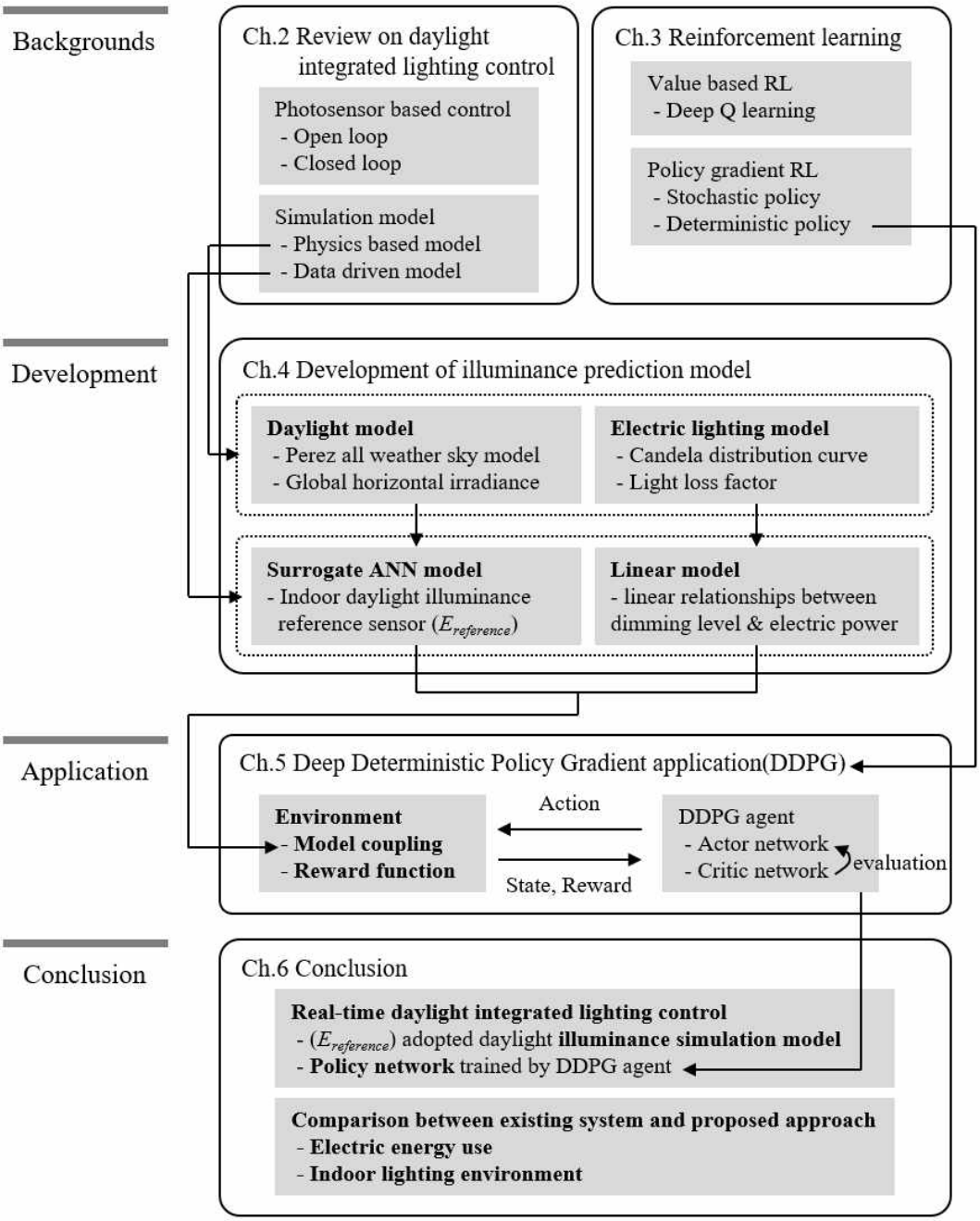
수를 결정할 수 없고, 수렴 조건에 도달할 때까지 반복하여 제어변수를 탐색한다. 반면, 강화학습은 제약조건과 목적함수 대신, 제어 결과를 바탕으로 변수의 성능을 평가하는 보상함수를 이용한다. 학습하는 동안, 일정 주기를 반복하면서 다양한 환경에 대해 정책에 기반하여 제어를 수행하고, 그에 따른 보상을 얻으며 궤적 데이터를 수집한다. 이를 이용하여, 주기(에피소드)의 총 보상을 가장 크게 만드는 최적 정책을 구하는 것이 강화학습의 목표이다. 실제 제어를 수행할 때는 최종 정책을 이용하여 실시간으로 제어변수를 도출할 수 있다.

본 연구에서는 작업면에 조도 센서를 설치하지 않고, 시뮬레이션 모델을 이용하여 자연채광 및 전기조명 조도를 실시간으로 구하는 방법을 제안한다. 또한, 제작된 모델을 기반으로 강화학습을 이용하여 조명 시스템을 최적으로 제어하는 방법을 제시하고자 한다.

1.2 연구의 범위 및 방법

대상 건물은 인천광역시 동구에 위치한 중장비 엔진 공장이며, 내부는 작업 유형에 따라 가벽으로 구획되어 있다. 본 연구에서는 검사 및 운송 작업이 수행되는 구역을 대상으로 한다. 3D 모델은 Rhinoceros로 제작하였으며, Grasshopper 플랫폼에서 DIVA를 이용하여 조도 시뮬레이션을 수행하였다. 자연채광 조도 예측 모델은 기상청의 수평면 총 일사량(Global Horrizontal Irradiance, GHI) 데이터를 입력으로 받아 작업면의 조도를 예측한다. 하지만, 기상청의 일사량 데이터는 실시간으로 받기 어렵고, 건물이 위치한 지역적 특징을 반영하지 못한다. 실시간 제어를 수행하기 위해 실내에 설치된 자연채광 조도 센서의 측정값을 입력으로 이용하는 기계학습 대리모델을 개발하였다. 기계학습 알고리즘은 인공신경망(ANN)을 이용하였다. 전기조명의 경우, 조명기구 제조사에서 제작한 배광파일(ies 파일)을 이용하였다. 조명 기기 그룹별로 수행한 DIVA 시뮬레이션의 결과를 바탕으로 디밍 수준에 따른 선형 모델을 제작하였다. 모델 제작 방법은 4장에서 상술되며, 각 모델은 실측 및 제작, 측정 데이터 기반 보정 및 검증, 모델 확장 순서로 개발되었다.

제작된 모델과 강화학습 알고리즘을 연계하기 위해 파이썬 환경을 이용하였다. 연속 변수인 디밍 제어를 수행하기 위해 정책 그래디언트 강화학습을 검토하였으며, 고차원 제어변수 결정에 효과적인 심층 확정적 정책 그래디언트(Deep Deterministic Policy Gradient, DDPG) 알고리즘을 선정하였다. 훈련 및 검증을 위해 자연채광 모델과 전기조명 모델을 결합하여 강화학습 환경을 제작하였다. OpenAI의 Gym을 참고하였으며, 현재 시점의 제어변수(행동, action)를 입력하면 제어 결과를 평가한 보상(Reward)과 다음 시점의 시스템 상태(state)를 돌려준다. 강화학습 제어에 따른 작업면의 조도 및 균제도를 권장 조도 및 균제도와 비교하여 본 연구에서 제안한 제어 방법의 유용성을 제시하였다. 결과로 기존의 제어 방법과 비교하여, 에너지 절감량을 산출하였다. 그림 1-2는 본 논문의 전체 흐름을 보여준다.



* Bold : Key processes and works

그림 1-2. 연구의 절차 및 구성

제 2 장 조명 시스템 제어

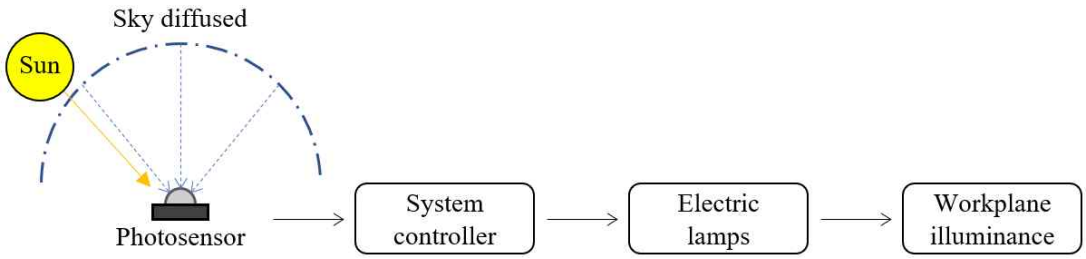
본 장에서는 문헌조사를 통해 자연채광이 도입된 실내 조도 환경에서 기존에 사용된 전기조명 시스템의 제어 방법을 고찰한다. 대부분 조도 센서를 기반으로 제어변수를 결정하는 방법이며, 센서의 설치 위치를 선정하는 데에 한계점을 지닌다.

2.1 조도 센서 기반 조명 시스템 제어

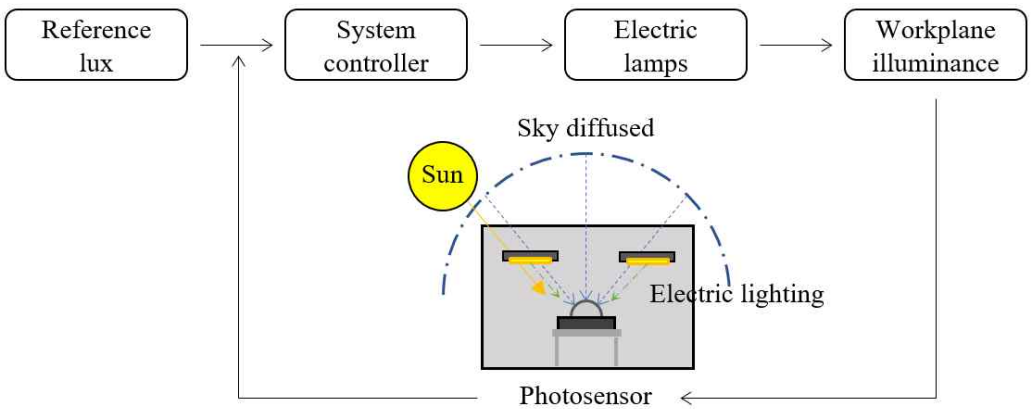
조도 센서를 이용하여 자연채광 조도에 따라 전기조명을 제어하는 알고리즘은 open-loop 제어와 closed-loop 제어로 구분된다. open-loop 제어의 경우, 조도 센서는 자연채광이 잘 드는 동시에 전기조명의 영향이 미치지 않는 지점에 설치된다(그림 2-1의 (a)). 센서의 위치와 작업면에 도달하는 자연채광 사이의 관계를 보정하여 이를 기반으로 제어 방법을 결정해야 한다. 대부분의 경우, 거리에 따른 선형 관계를 가정하는데, 이는 천구 상태를 반영하지 못한다. 실내 공간의 형태로 인해, 태양의 위치 및 운량에 따라 각 지점에서의 자연채광 변화율은 모두 다르다. 즉 단순한 선형 가정으로는 작업면의 자연채광을 정확히 알 수 없으며, 이로 인해 제어 결과에 오차가 생긴다. 이 문제를 해결하기 위해서는 작업면과 동일한 자연채광 특성을 가진 지점에 센서를 설치해야 한다. 하지만 그 위치를 찾는 것이 어렵고, 찾았다고 할지라도 해당 작업면에만 한정되어, 다른 자연채광 특성을 가진 작업면에 대해 범용적으로 이용될 수 없다.

closed-loop 제어는 작업면에 조도 센서를 설치하여 측정한 조도(자연채광 조도와 전기조명 조도의 합)를 피드백으로 조명기기의 광속을 조절하는 방법이다 (그림 2-1의 (b)). 조명기구마다 대응하는 조도 센서를 이용하여 대상 지점의 조도를 측정한다. 측정값과 목표 조도의 차이만큼 해당 조명기구의 광속을 조절하여 정확도를 확보할 수 있다. closed-loop 제어에서 조

도 센서는 작업면에 설치하는 것이 가장 효과적이다(Di Louie, 2008; Haq et al., 2014). 하지만, 이는 작업자의 활동에 방해가 될 수 있으며, 반대로 사람이 센서를 가릴 수도 있다. 이와 같은 이유로 작업면에 센서를 설치하는 대신, 벽 또는 천장에 설치하고 센서를 보정해야 한다 (Caicedo, Pandharipande, & Willems, 2014). open-loop 제어의 경우와 같이, 센서의 측정 정확도 확보를 위해 설치 지점을 적절하게 선정해야 한다. 또한 closed-loop 제어를 위해서는 조명기구 간의 조도 간섭을 해결해야 한다. 각 조명기구는 여러 작업면에 동시에 영향을 준다. 한 지점의 조도를 맞추기 위해 대응하는 조명기구의 광속을 변경하는 순간 다른 지점에 설치된 조도 센서의 측정값이 함께 달라진다.



(a) open-loop 제어



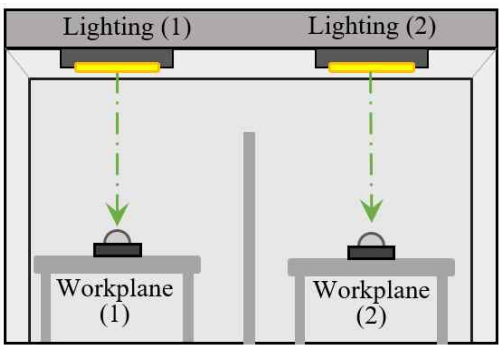
(b) closed-loop 제어

그림 2-1. 조도센서 기반 전기조명 제어 방법

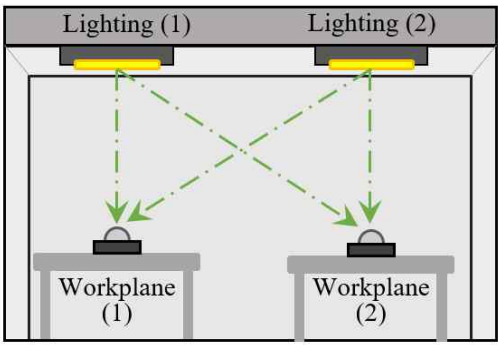
Mahdavi et al(1995)는 실내 조명시스템 설계단계에서 작업면을 여러 구역으로 나누고, 각 구역을 담당하는 조명기구를 그룹으로 묶어서 한 번에 제어하는 방법을 제시하였다. 결과로 조도 센서 설치 개수를 줄일 수 있었으며, 해당 구역의 전기조명 조도를 평균화하여 간섭을 해결하는 효과가 있었다. 하지만, 구역과 구역 사이의 전기조명 간섭 문제는 여전히 존재하였으며, 구역이 많은 대형 공간일수록 전체 조명 시스템의 제어변수를 찾는 것이 어려워진다.

Koroglu, & Passino(2014)는 제어 단위별 조명기구를 번갈아 조절하면서 모든 지점의 조도 환경을 확보하는 알고리즘(illumination balancing algorithm)을 제안하였다. 이 방법은 알고리즘에 대한 수학적 해석이 필요하고, 센서 보정(calibration)도 수행해야 한다. 또한, 모든 조도 환경이 정상 상태에 도달할 때까지 반복해야 하고, 자연채광 조도가 변할 때마다 다시 수행해야 한다.

이처럼, 작업면과 조명기기 사이에 교호작용이 있는 경우, 조도를 확보하기 위해 각 조명 기구의 디밍 수준을 개별로 결정하는 방법(Single Input Single Output, SISO)(그림 2-2의 (a))으로는 전체 조명 시스템의 최적 제어변수를 찾는 것이 어렵다. 각 작업면에 도달하는 광속이 어떤 조명기구에서 출력된 것인지 추적할 수 있어야 각 조명기구의 제어 방법을 결정할 수 있다(그림 2-2의 (b)).



(a) 조명과 작업면 사이 간섭 無



(b) 조명과 작업면 사이 간섭 有

그림 2-2. 다수의 조명기기 및 작업면 사이의 관계

2.2 모델 기반 조명 시스템 제어 및 최적화

2.2.1 조도 예측 모델

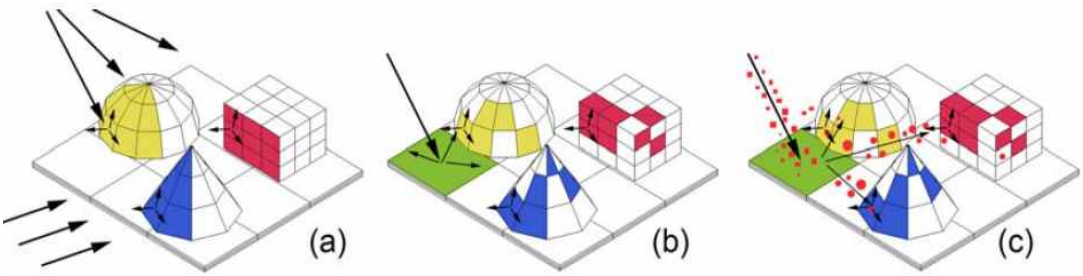
각 조명기구의 제어변수를 한 번에 결정하여 전기조명 조도를 공급하는 방법(Mulitple Input Multiple Output, MIMO)은 조명기구 사이에 발생하는 간섭 문제를 해결할 수 있다. 이를 위해서는 각 광원에 따라 작업면의 조도를 출력하는 실내 조명 시스템 모델이 필요하다. 물리법칙 혹은 이를 기반으로 한 시뮬레이션 프로그램을 이용하거나 작업면에서 측정된 조도 데이터를 수집하여 광원에 따른 조도 예측 모델을 제작할 수 있다.

시뮬레이션 모델은 물리법칙에 기반한 White-box 모델과 데이터에 근거한 Black-box 모델로 나뉘며, 이 둘을 결합한 Gray-box 모델도 있다. 조도 예측을 위한 White-box 모델(혹은 Physics based 모델) 제작을 위해서는 건물의 형상 및 재료의 물성치 정보와 광원에 대한 해석(Sky model, 조명기구 배광정보)이 필요하다. 미지변수의 경우 가정해야 하고 광속을 계산하는 물리법칙을 선정해야 하는 등 시뮬레이션 수행자의 주관적인 판단이 개입된다. 이로 인해 시뮬레이션 결과가 달라질 수 있으며, 모델의 재연성과 반복성 및 신뢰성을 확보하기 어렵다 (박철수, 2006). 모델의 결과와 실재 환경의 차이를 최대한 줄이기 위해, 측정값에 기반하여 모델을 보정하고 검증하는 과정이 필요하다 (Kim and park, 2016).

조도 예측 모델은 광원에 따라 자연채광 모델과 전기조명 모델로 구분된다. 두 모델은 3d 형상과 재료의 물성치를 공유하되, 빛을 설명하는 변수에 차이가 있다. 자연채광은 태양에서 바로 도달하는 직달일사와 천구에 반사되어 유입되는 확산일사로 구성되며, 태양의 방위각과 고도각, 일사량, 운량에 따라 지속적으로 변한다. 이를 묘사하기 위해 sky-model을 이용하여 천구를 패치로 나누고 각각의 패치에 광량을 분배한다. 한편, 전기조명의 경우, 배광곡선과 조명보수율(Light Loss Factor, LLF)을 이용하여 작업면에 도달하는 조도를 구할 수 있다. 조명기구마다 광속이 출력되는 방향과 세기가 다르다. 제조사에서는 이를 측정하여 배광곡선(Candela Distribution

Curve, CDC)을 제작하고, 사용전력, 광도 등의 스펙 정보와 함께 ies파일 포맷으로 제공한다. 조명보수율은 설치 환경 및 유지보수로 인한 효율 손실을 반영하는 요소이며, 최적 상태일 때의 광도 대비 실제 광도의 비율을 의미한다.

광원에서 대상 지점에 도달하는 조도를 계산하는 대표적인 알고리즘으로 Radiosity 방법과 Ray-tracing 방법이 있다. Radiosity 방법은 장면(Scene)에 따라 빛을 계산하는 방법으로 복사 열전달에서 개발되었다. 한 장면에 포함된 표면의 형상 계수(Form factor)에 기반하여 각 면 사이의 반사광을 계산한다. 실제 건물 환경에서는 형상 계수를 구하는 계산이 복잡하고, 정확도가 떨어진다. Ray-tracing 방법은 관점(View)에 따라 빛을 계산하는 방법이다. 추적 계산을 시작하는 지점에 따라 광원에서 시작하는 Forward ray-tracing, 측정 위치에서 시작하는 Backward ray-tracing으로 구분된다. 한 지점에 도달하는 광선은 모든 방향에서 유입된다. 태양이나 조명기기와 같이 비중이 크고, 중요한 광원에서 도달하는 빛은 확정적으로 추적하여 계산할 수 있다. 하지만, 천구나 다른 표면에 반사되어 유입되는 빛의 경우, 모든 방향을 추적하는 것은 불가능하기에 확률적 샘플링 방법을 이용한다. 샘플링 방향 수와 반사 횟수가 많을수록 정확한 시뮬레이션이 가능하지만, 이에 따라 연산시간도 함께 늘어난다.



(a) ray-tracing (b) radiosity (c) photon map

그림 2-3. 빛 환경 시뮬레이션 알고리즘

(Ochoa, Aries, & Hensen, 2010)

여러 가지 시뮬레이션 프로그램은 목적에 따라 다음 두 방법을 취사선택하거나 융합하여(그림 2-3.) 조도 계산을 수행한다. 현재까지 Backward ray-tracing 기반의 Radiance가 가장 많이 이용되고 있으며, 수많은 연구를 통해 그 성능이 검증되었다 (Ochoa et al. 2010). Radiance는 사용자 인터페이스가 없기에 3d 모델링 프로그램에 엔진으로 탑재되어 활용된다. 본 연구에서 자연채광 및 전기조명 모델 제작에 이용한 DIVA는 Rhinoceros & Grasshopper 3d 모델링 프로그램 환경에서 제작된 빛 환경 시뮬레이션 도구로서 Radiance를 엔진으로 이용한다 (Solemma LLC, 2020).

Black-box 모델(혹은 Data-driven 모델)은 물리법칙을 이용하지 않고, 입·출력 데이터의 관계를 확률적, 통계적으로 묘사한다. 모델 제작을 위해 전문적인 지식이 필요하지 않으며, 매우 빠른 시뮬레이션이 가능하다. 기계학습(Machine learning) 모델은 대표적인 Data-driven 모델로서 데이터의 구조 및 패턴을 발견하는 방법이다. 이를 위해서는 데이터의 양이 충분해야 하며, 질적으로도 다양성을 확보해야 한다. 즉, 모델의 인과성(causality)을 고려하여 입력변수와 출력변수를 적절히 선정해야 하며, 노이즈 및 이상치를 검출하는 등의 데이터 전처리가 필요하다. 결과로 완성된 데이터 세트의 양이 기계학습 알고리즘을 훈련 및 검증하기에 충분해야 한다 (박성호, 2019).

Data-driven 방법으로 조도 시뮬레이션 모델을 제작하기 위해서는 광원에 따라 입·출력 변수를 선정하고, 이를 수집해야 한다. White-box 모델의 변수를 그대로 사용하는 것이 모델의 인과성(causality)을 확보하는 측면에서 가장 좋지만, 목적에 따라 그 외의 변수를 사용할 수도 있다. 한편, 조도 시뮬레이션의 출력변수는 목표 지점의 조도이다. 정확한 측정값을 얻기 위해서는 목표 지점에 조도 센서를 설치하고 주변 환경의 간섭을 통제해야 한다. 그 후, 입력변수(자연채광: 시간에 따른 태양의 위치, 직달일사량, 산란일사량, 운량 등, 전기조명: 조명 제어 값(ON·OFF, 디밍 등))를 다양하게 변경하면서 측정된 값을 수집해야 한다.

여러 기계학습 알고리즘은 모델 구조와 예측 방법이 서로 다르며, 데이터의 형태에 따라 성능에 차이가 있다. 하지만, 모든 기계학습 알고리즘은 매

개변수(parameter)를 이용하여 입·출력 데이터 사이의 관계를 설명하고, 최적화 과정을 통해 실측값과 모델의 예측값의 차이를 줄이는 방향으로 매개변수를 조절하는 것은 같다 (안기언 2019). 건물 에너지 시뮬레이션 분야에서는 비선형 시스템에 대한 높은 수준의 정확성을 구현해 낼 수 있는 인공신경망(Artificial Neural Network, ANN)이 많이 이용되고 있다 (Afram et al., 2017).

본 연구에서는 DIVA로 작성 및 보정된 자연채광 조도 예측 모델(White box model)을 모사하는 대리모델을 인공신경망(ANN)으로 제작하였다. 기계학습 모델을 제작하는 과정에서 모델의 인과성을 고려하여 입력변수를 변경하였다. 자연채광 조도에 대한 실시간 시뮬레이션 수행을 위해 기상청에서 제공하는 수평면 총 일사량(GHI) 대신, 실내에서 측정된 자연채광 조도(E_r)를 사용하였다.

2.2.2 시스템 제어 및 최적화

조명시스템 제어에 필요한 전기조명 모델은, 디밍 제어변수에 따라 작업면의 조도를 출력할 수 있어야 한다. 또한, 목적함수에 해당하는 전력사용량도 계산할 수 있어야 최적화 알고리즘을 적용할 수 있다. 조명기기의 출력과 작업면 조도 수준 그리고 전력사용량과 조명기기의 출력 사이의 관계는 선형으로 근사할 수 있으며, 이에 따라 선형 모델을 제작할 수 있다. 조명기기의 출력이 최대일 때를 기준으로 1법칙 기반의 전기조명 조도예측모델을 제작하고 시뮬레이션을 수행하여 각 조명기기와 각 작업 지점의 조도 관계를 알 수 있다.

시뮬레이션 결과를 매트릭스 형태로 정리하여 식 (2-1)과 같이 표현할 수 있다. 여기에 조명 기기별 디밍 수준을 벡터 형태로 곱하면 디밍 제어변수에 따른 작업 지점별 조도를 구할 수 있다(식 (2-2)). 한편, 조명기기의 출력이 최대일 때의 전력량과 디밍 제어변수를 곱하여 전체 시스템의 전력량을 구할 수 있다(식 (2-3)).

$$A = \begin{bmatrix} a_{11}\ a_{12}\ a_{13} \cdots a_{1j} \\ a_{21} \qquad\qquad a_{2j} \\ a_{31} \qquad \ddots \qquad a_{3j} \\ \vdots \qquad\qquad \vdots \\ a_{i1}\ a_{i2}\ a_{i3} \cdots a_{ij} \end{bmatrix} \tag{2-1}$$

여기서,

a_{ij} : j번째 조명기구에 따른 i지점의 조도

i : 작업 지점의 수

j : 조명 기구의 수

$$E_{elec} = A \times D \tag{2-2}$$

여기서,

$E_{elec} = [e_1, e_2, e_3, \cdots, e_i]^T$, e_i : i지점(node)의 전기조명 조도

$D = [d_1, d_2, d_3, \cdots, d_j]^T$, d_j : j번째 조명기구의 디밍 수준

$$SP = P^T \times D \tag{2-3}$$

여기서,

SP : 전체 시스템의 전력량

$P = [p_1, p_2, p_3, \cdots, p_j]^T$, p_j : j번째 조명기구의 전력량

$D = [d_1, d_2, d_3, \cdots, d_j]^T$, d_j : j번째 조명기구의 디밍 수준

기존의 모델 기반 조명 제어 연구에서는 이와 같은 시스템 모델을 제작하고, 최적화 알고리즘을 적용하여 에너지 절감량을 도출하였다. Wen & Agogino(2008)는 물리법칙 기반의 시뮬레이션 도구인 Radiance를 이용하여, 각 전기 조명기구에 대한 작업면 조도 분포를 구한 다음, 디밍 수준에 따라 선형결합으로 시스템 모델을 제작하였다. 최적화는 선형계획법으로 수행되었으며, 작업면 목표 조도를 만족하는 동시에 에너지 사용량을 최소화하도록 적용되었다. Tran & Tan(2014)은 작업면에 임시로 조도 센서를 배치하고, 개별 등에 대한 제어를 수행하여 조도 데이터를 수집하였다. 수집한 데이터를 기반으로 기계학습 모델을 제작하여 최적화에 이용하였다. 상술한 두 연구는 자연채광의 영향이 없거나 적은 실내환경을 대상으로 수행되었으며, 전기 조명기구의 조도만 고려한다. 시스템 모델을 이용하면 각 제어변수 조합에 따른 작업 지점의 조도를 예측할 수 있으므로, 권장 조도 환경을 만족하면서 에너지 사용량을 가장 적게 사용하는 제어변수를 최적화 알고리즘으로 탐색한 것이다.

Pandharipande, & Caicedo(2010)와 Seyedolhosseini et al(2020)의 연구에서는 자연채광을 함께 고려하기 위해 조도 센서를 이용하였다. 먼저 전기조명 시스템을 설명하는 모델을 제작하여 전기조명 조도를 예측하였다. 이를 센서의 측정값과 비교하여 그 차이만큼을 자연채광 조도로 추정하였다. 건물 내에서 측정된 조도는 자연채광 조도와 전기조명 조도의 합이므로, 센서의 측정값에서 모델로 구한 전기조명 조도를 제하여 자연채광 조도를 구한 것이다. 권장 조도수준에서 자연채광 조도로 확보되는 만큼을 뺀 것을 제약조건으로, 선형계획법을 적용하여 전력사용량을 최소화하였다.

전기조명 모델과 같이 자연채광 조도를 예측하는 모델을 제작하면, 자연채광이 유입되는 공간에서도 조도 센서 없이 전기조명 제어를 수행할 수 있다. 김영섭 외(2019)의 연구에서는 자연채광 조도 예측 모델과 전기조명 조도 예측 모델을 각각 제작하였다. 자연채광의 동적 거동을 baseline으로 하여 전기조명 시스템의 제약조건을 설정하고, 메타 휴리스틱스 전역 최적화 방법인 유전알고리즘을 적용하여 제어변수를 탐색하였다.

본 연구의 대상 구역은 매우 넓고, 내벽이 거의 없는 열린 공간(open

space)이다. 북측 삼각 천장 3열의 채광창을 통해 산란 일사가 많이 유입되며, 작업 지점마다 자연채광 조도의 수준이 달랐다. 이와 같은 이유로 고정된 조도센서의 설치 지점을 찾는 데에 한계가 있으므로, 광원별로 조도 예측 모델을 제작하여 이를 대체하였다. 모델 제작 방법 및 과정은 4장에서 상술된다.

한편, 현재까지 조명 제어에 적용된 최적화 알고리즘은 대부분 선형계획법에 기반한 경사하강법 혹은 메타-휴리스틱스 방법을 적용한 것이다. 경사하강법의 경우, 빠른 속도로 변수를 탐색할수 있다. 하지만, 설정 파라미터(초기값, 제약조건 등)에 따라 국지 최적해에 빠질 수 있으며. 이로 인해 최적화 성능에 편차가 생긴다. 유전알고리즘과 같은 메타-휴리스틱스 방법은 비선형인 해공간에서 전역 최적해를 찾는 문제에 효과적이다. 하지만, 많은 변수 조합을 샘플링하여 해공간을 탐색하는 방법으로, 최적해를 찾기 위해 많은 연산시간이 필요하다.

실시간 조명 시스템 제어를 위해서는 각각의 제어 시점마다 빠른 속도로 최적 제어변수를 도출할 수 있어야 한다. 본 연구에서는 심층 확정적 정책 그래디언트 강화학습을 이용하여 실시간 조명 시스템 제어를 수행하였다. 3장에서 강화학습에 대해 간략히 살펴보고, 5장에서 조도 예측 모델과 강화학습을 연계한 방법 및 제어 결과를 소개한다.

제 3 장 강화학습 방법

3.1 강화학습 개요

강화학습은 정해진 기간 안에 목표를 최적으로 달성하기 위한 순차적인 의사결정 문제를 푸는 방법이다. 최근 다양한 제어 분야에서 많이 활용되고 있으며, 기존의 최적제어(optimal control)와 같이 행동에 대한 보상(최적제어에서는 비용)에 기반한다. 강화학습에서는 의사결정 주체를 에이전트(agent)라 하며, 시스템을 환경(environment)이라고 한다. 에이전트는 환경의 상태(state)를 측정하거나 관측하고, 행동(action)을 선택하여 환경에 전달한다. 환경은 에이전트의 행동에 따라 변화하고, 이를 평가하여 보상(reward)을 돌려준다. 에이전트의 의사결정(행동)은 일정한 정책(policy)에 근거하여 이루어지는데, 학습 기간에 다양한 행동을 선택하고 획득한 보상을 이용하여 이를 구축한다. 지정된 기간(episode)이 끝날 때까지 누적된 보상을 최대화하는 정책을 구하는 것이 강화학습의 목표이다.

3.1.1 마르코프 결정 과정(MDP)

강화학습에서의 시간은 이산적(discrete-time)이다. 즉, 기간을 일정 간격으로 나누고, 시점마다 독립적으로 에이전트와 환경이 상호작용하게 된다. 이와 같은 흐름을 수학적으로 표현한 것이 마르코프 결정 과정(Markov Decision Process, MDP)이며, 그 효율이 과거 모든 정보에 기반한 의사결정만큼 우수함이 확인되고 있다 (Sutton and Barto, 1998). MDP에서는 상태, 행동 및 보상 등 각 요소의 관계를 확률적으로 표현한다. 어떤 상태(s_t)에서 행동(a_t)을 취했을 때, 다음 상태(s_{t+1})의 확률밀도함수를 상태전이함수, $p(s_{t+1} \mid s_t, a_t)$라 한다. 여기서 행동(a_t)은 정책, $\pi(a_t \mid s_t)$에 의해 결정되

며, $\pi(a_t \mid s_t)$는 상태(s_t)에 따라 선택할 수 있는 행동의 확률을 나타내는 조건부 확률밀도함수, $p(a_t \mid s_t)$이다. 정책으로 결정된 행동을 상태전이함수에 입력하여 다음 상태로 이동하면, 환경으로부터 보상, $r(s_t, a_t)$이 주어진다(그림 3-1).

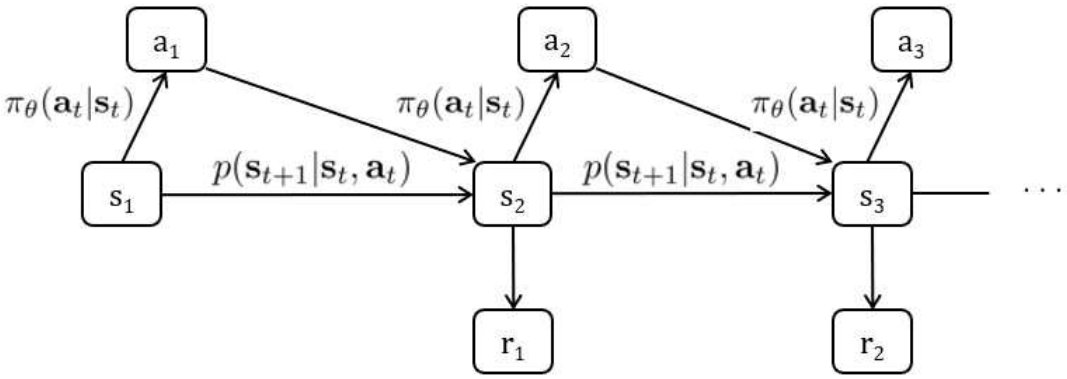


그림 3-1. 마르코프 결정 과정(Markov Decision Process, MDP)

제어 시점마다 얻은 보상의 총합을 반환(return)이라고 하며 식 (3-1)로 나타낸다. 여기서 $\gamma \in [0,1]$는 현재 시점의 즉각적인 보상과 미래에 얻게 될 보상의 가중치를 결정하는 역할을 하는 감가율(discount rate)이다. 또한, 제어 기간을 지정하지 않은 무한 에피소드에서 반환이 발산하지 않도록 하는 수학적인 역할을 한다.

$$G_t = r(s_t,a_t) + \gamma r(s_{t+1},a_{t+1}) + \gamma^2 r(s_{t+2},a_{t+2}) + \cdots + \gamma^{T-t} r(s_T,a_T)$$
$$= \sum_{k=t}^{T} \gamma^{k-t} r(s_k,a_k) \quad (3\text{-}1)$$

3.1.2 가치함수

강화학습의 목표는 반환(return)을 최대화하는 것으로 이를 확률적으로 추정하기 위해 가치함수(value function)를 도입한다. 어떤 상태에서 정책에 따라 행동을 결정했을 때, 반환의 기댓값을 상태 가치(state-value)라고 한다 (식 (3-2)). 여기서 τ는 어떤 상태 s_t에서 시작하여 정책 π에 따라 발생하는 궤적($a_t, s_{t+1}, a_{t+1}, s_{t+2}, a_{t+2}, \cdots, s_T, a_T$)이다. 한편, 어떤 상태에서 행동을 선택한 이후, 정책에 따라 생성되는 반환의 기댓값을 행동 가치(action-value)라고 한다. 이때, 가치함수의 궤적 τ는 ($s_{t+1}, a_{t+1}, s_{t+2}, a_{t+2}, \cdots, s_T, a_T$)이다 (식 (3-3)). 상태변수($s_t$)에서 선택할 수 있는 모든 행동($a_t$)에 대한 행동 가치($Q^{\pi}(s_t, a_t)$)의 평균값이 상태 가치 ($V^{\pi}(s_t)$)이며, 식 (3-4)의 관계를 갖는다.

$$V^{\pi}(s_t) = E_{\tau \sim p(\tau \mid s_t)}[\sum_{k=t}^{T} \gamma^{k-t} r(s_k, a_k) \mid s_t] \tag{3-2}$$

$$Q^{\pi}(s_t, a_t) = E_{\tau \sim p(\tau \mid s_t, a_t)}[\sum_{k=t}^{T} \gamma^{k-t} r(s_k, a_k) \mid s_t, a_t] \tag{3-3}$$

$$V^{\pi}(s_t) = E_{a_t \sim \pi(a_t \mid s_t)}[Q^{\pi}(s_k, a_k)] \tag{3-4}$$

현재 시점의 상태 가치 및 행동 가치와 다음 시점의 상태 가치 및 행동 가치는 각각 식 (3-5) 및 식 (3-6)의 관계를 따르며, 이를 벨만 방정식이라 부른다.

$$V^{\pi}(s_t) = E_{a_t \sim \pi(a_t \mid s_t)}[r(s_t, a_t) + E_{s_{t+1} \sim p(s_{t+1} \mid s_t, a_t)}[\gamma V^{\pi}(s_{t+1})] \tag{3-5}$$

$$Q^{\pi}(s_t, a_t) = r(s_t, a_t) + E_{s_{t+1} \sim p(s_{t+1} \mid s_t, a_t)}[E_{a_{t+1} \sim \pi(a_{t+1} \mid s_{t+1})}[\gamma Q^{\pi}(s_{t+1}, a_{t+1})]] \tag{3-6}$$

상태 가치를 최대화하는 정책을 최적 정책이라고 하며, 이를 적용했을 때의 가치함수를 최적 가치함수라 한다. 최적 상태 가치함수와 최적 행동 가치함수의 벨만 방정식은 각각 식 (3-7) 및 식 (3-8)로 정리되며, 벨만 최적 방정식이라고 한다. 벨만 방정식과 같이, 현재 시점의 최적 가치와 다음 시점의 최적 가치의 관계를 나타낸다. 최적 상태 가치함수(식 (3-7))와 최적 행동 가치함수(식 (3-8)) 사이에는 식 (3-9)의 관계가 있다. 한편, 최적 정책은 식 (3-10)과 같이 쓸 수 있다.

$$V^*(s_t) = \max_{a_t} [r(s_t, a_t) + E_{s_{t+1} \sim p(s_{t+1} \mid s_t, a_t)} [\gamma V^*(s_{t+1})]] \tag{3-7}$$

$$Q^*(s_t, a_t) = r(s_t, a_t) + E_{s_{t+1} \sim p(s_{t+1} \mid s_t, a_t)} [\gamma \max_{a_{t+1}} Q^*(s_{t+1}, a_{t+1})] \tag{3-8}$$

$$V^*(s_t) = \max_{a_t} Q^*(s_t, a_t) \tag{3-9}$$

식 (3-8)과 식 (3-9)를 식 (3-10)에 대입하여 식 (3-11)로 표현할 수 있다. 즉, 최적 정책은 어떤 상태에서 행동 가치가 가장 큰 행동을 취하는 것이다. 벨만 최적성 원리에 따라, 최적 정책으로 생성된 궤적은 반환을 최대화하는 최적 시퀀스가 된다.

$$\pi^*(a_t \mid s_t) = \arg\max_{a_t} [r(s_t, a_t) + E_{s_{t+1} \sim p(s_{t+1} \mid s_t, a_t)} [\gamma V^*(s_{t+1})]] \tag{3-10}$$

$$\pi^*(a_t \mid s_t) = \arg\max_{a_t} Q^*(s_t, u_t) \tag{3-11}$$

3.2 정책 그래디언트 강화학습

3.2.1 Reinforce 알고리즘

강화학습은 문제 해결을 위해, ① 다양한 행동을 수행하면서 데이터를 수집하고, ② 이를 기반으로 가치함수를 추정한 다음, ③ 정책을 업데이트한다. 이 과정을 반복하여 일정 수준의 반환에 수렴하였을 때의 정책을 최적 정책으로 간주한다. 가치함수를 기댓값의 형태로 추정하는 이유는 반환이 현재 상태에서는 알 수 없는 미래의 보상을 포함하기 때문이다. 학습 기간을 통해 다양한 행동을 수행하고 이에 따른 반환을 많이 관측할수록 더욱 정확한 가치함수 추정이 가능하다. 가치함수를 추정하여 최적 정책을 구하는 방법을 가치기반 강화학습이라고 한다. 각 상태에서 행동 가치가 가장 큰 행동을 선택하는 방법으로 DQN(Deep Q Network)이 대표적이다(안기언, 2018).

한편, 정책 함수를 파라미터화하고, 목적함수의 그래디언트를 이용하여 이를 추정하는 방법을 정책 그래디언트(Policy Gradient)라고 한다(Sutton et al., 1999). 파라미터가 θ인 정책을 π_θ라고 하며, 이에 따라 생성되는 궤적의 확률밀도함수를 $p_\theta(\tau)$라고 한다. π_θ로 얻는 반환의 기댓값을 식 (3-12)와 같이 기술하며, $J(\theta)$를 가장 크게 만드는 π_θ를 구하는 것이 정책 그래디언트의 목표이다. 정책은 인공신경망(Artificial Neural Network, ANN)의 형태로 구축되고, 그 파라미터(가중치, 편향)가 θ이다. 이를 정책 신경망(policy-network)이라고 한다. 학습을 통해 정책 신경망의 파라미터를 변경하면서 최적 정책을 탐색한다.

$$\begin{aligned} J(\theta) &= E_{\tau \sim p_\theta(\tau)}\left[\sum_{k=0}^{T} \gamma^k r(s_k, a_k)\right] \\ &= \int_\tau p_\theta(\tau)\left(\sum_{t=0}^{T} \gamma^t r(s_t, a_t)\right) d\tau \end{aligned} \tag{3-12}$$

정책 그래디언트의 목적함수인 $J(\theta)$의 최대값을 탐색하기 위해, $J(\theta)$를 θ로 미분하고, 이를 목적함수의 그래디언트 식이라 한다 (식 (3-13)). 그래디언트 식에서 반환과 관련된 항($\sum \gamma^k r(s_k, a_k)$)은 모든 시점에 대한 것으로, 현재 시점 t에서의 반환만을 고려해야 한다. 또한, 실용적인 목적에서 예정된 보상에만 감가율을 적용하기 위해 식 (3-14)와 같이 그래디언트 식을 변경하여 이용한다. 식 (3-15)와 같이 목적함수의 그래디언트에 따라 경사하강법으로 정책 파라미터 θ를 업데이트하여 최적 정책을 탐색할 수 있다.

$$\begin{aligned}\nabla_\theta J(\theta) &= E_{\tau \sim p_\theta(\tau)}[(\sum_{t=0}^{T} \nabla_\theta \log \pi_\theta(a_t \mid s_t))(\sum_{t=0}^{T} \gamma^t r(s_t, a_t))] \\ &= E_{\tau \sim p_\theta(\tau)}[\sum_{t=0}^{T}(\nabla_\theta \log \pi_\theta(a_t \mid s_t)(\sum_{k=0}^{T} \gamma^k r(s_k, a_k))]\end{aligned} \tag{3-13}$$

$$\nabla_\theta J(\theta) = E_{\tau \sim p_\theta(\tau)}[\sum_{t=0}^{T}(\nabla_\theta \log \pi_\theta(a_t \mid s_t)(\sum_{k=t}^{T} \gamma^{k-t} r(s_k, a_k))] \tag{3-14}$$

$$\theta \leftarrow \theta + \alpha \nabla_\theta J(\theta) \tag{3-15}$$

하지만, $E_{\tau \sim p_\theta(\tau)}$의 값은 수학적으로 계산할 수 없으므로 샘플링하여 이들의 평균으로 근사한다. 정책을 이용하여 궤적을 생성함으로 샘플을 얻을 수 있으며, 이를 Reinforce 알고리즘이라 한다(Williams, 1992). 에피소드가 끝날 때까지 업데이트가 불가능하며, 반환의 분산이 크다는 문제가 있지만, 여타 정책 그래디언트 알고리즘의 기초가 된다는 점에서 의의가 있다.

3.2.2 Actor-Critic 구조

현재 이용되고 있는 대부분의 정책 그래디언트 알고리즘은 Actor-Critic 구조를 가진다. 에피소드가 진행되면서 정책 신경망을 업데이트함과 동시에 가치함수도 함께 학습하는데, 이를 이용하여 정책 신경망을 평가한다. Actor-Critic 구조의 목적함수 그래디언트 식을 살펴보면, 기존의 반환 관련 항을 가치함수에 대한 항으로 정리하여 이용한다 (식 (3-16)).

$$\nabla_{\theta} J(\theta) = \sum_{t=0}^{T} (E_{s_t \sim d_{\theta}(s_t),\, a_t \sim \pi_{\theta}(a_t \mid s_t)} [\nabla_{\theta} \log \pi_{\theta}(a_t \mid s_t) Q^{\pi_0}(s_k, a_k)]) \qquad (3\text{-}16)$$

즉, 목적함수의 그래디언트 식의 반환값을 구하기 위해 에피소드 끝날 때까지 기다리지 않고, 현재 시점의 가치함수를 이용하면 된다. 식 (3-16)에 따르면, 가치함수를 잘 추정할수록 목적함수의 그래디언트가 정확해지며, 결과적으로 Actor-Critic 알고리즘의 성능이 좋아진다. 가치함수 역시 인공신경망(ANN)으로 파라미터화하며, 이를 크리틱 신경망(critic-network)이라고 한다. 한편, 정책 신경망은 행동을 결정하는 역할을 하기에 액터 신경망(actor-network)이라고도 부른다.

3.3 심층 확정적 정책 그래디언트(DDPG)

행동 변수의 형태에 따라 정책 신경망의 출력층이 결정된다. n개의 행동 변수가 이산적일 경우, 선택 가능한 m개의 행동에 따라 출력 노드의 수는 n×m개가 된다. 하지만 행동 변수가 연속일 경우 필요한 출력 노드의 개수는 무한대이다. 이 경우, 액터 신경망은 행동의 확률을 출력하는 대신 확률 밀도함수($\pi(a_t \mid s_t)$)를 특정 분포로 고정한 다음, 그 파라미터를 출력한다. 보편적으로 가우시안 분포를 이용하며, 각 변수의 평균과 분산을 출력한다(출력 노드는 2n개). 확정적 정책 그래디언트(Deterministic Policy Gradient)은 행동의 확률밀도함수, $\pi(a_t \mid s_t)$ 대신 행동, $a_t = \pi(s_t)$을 직접 출력하는 방법이다(Silver et al., 2014). 정책이 특정 확률분포의 형태를 따르지 않기 때문에 자유도가 높고, 액터 신경망의 출력 노드 수는 변수의 개수인 n개로 줄어든다. 액터 및 크리틱 신경망을 심층 신경망으로 파라미터화한 것이 심층 확정적 정책 그래디언트(Deep Deterministic Policy Gradient, DDPG)이다.

확정정 정책 그래디언트의 목적함수는 상태 가치함수를 이용하여 기술한다 (식 (3-17)). 이를 미분한 목적함수의 그래디언트는 식 (3-18)과 같다. 확정적 정책을 가정하면, 식 (3-4)에 따라 상태가치 함수와 행동가치 함수가 같아진다. 이를 이용하여 $s = s_0, s_1, s_2, \cdots$에서의 가치함수 그래디언트 구한 다음, 식 (3-18)에 대입하면 식 (3-19)를 얻을 수 있다.

$$J(\theta) = \int_{s_0} V^{\pi_0}(s_0)p(s_0)ds_0 = E_{s_0 \sim p(s_0)}[V^{\pi_0}(s_0)] \tag{3-17}$$

$$\nabla_\theta J(\theta) = \int_{s_0} \nabla_\theta V^{\pi_0}(s_0)p(s_0)ds_0 \tag{3-18}$$

$$\nabla_\theta J(\theta) = \sum_{t=0}^{\infty}(E_{\bar{\tau} \sim p(\bar{\tau})}[\gamma^t \nabla_\theta \pi_\theta(s_t)\nabla_{a_t} Q^{\pi_0}(s_t, a_t)]) \tag{3-19}$$

여기서 궤적 $\bar{\tau}=(s_0,s_1,s_2,\cdots,s_t)$는 과거의 정책으로 생성되었다 할지라도 상태전이함수 $p(s_{t+1} \mid s_t,a_t)$의 결과에 차이가 없다. 또한 감가율 γ^t은 에피소드 후반에 수집된 데이터의 이용을 방해하게 되는 단점이 있기에 DDPG에서는$\gamma=1$로 가정한다. 즉, 식 (3-19)를 다시 변형하여 식 (3-20)의 형태로 만들 수 있다. $p_{\theta OLD}(\bar{\tau})$는 이전 정책에 따라 생성된 궤적을 이용하여 현재 정책을 개선할 수 있다는 것을 의미한다. 즉 DDPG는 오프-폴리시 방법으로, 효율적인 학습이 가능하다.

$$\nabla_\theta J(\theta)=\sum_{t=0}^{\infty}(E_{\bar{\tau}\sim p_{\theta OLD}(\bar{\tau})}[\nabla_\theta \pi_\theta(s_t)\nabla_{a_t}Q^{\pi_0}(s_t,a_t)]) \tag{3-20}$$

DDPG의 목적함수 그래디언트를 구하기 위해서는 행동가치를 알아야 한다. 즉, ϕ를 파라미터로 하는 크리틱 신경망 $Q_\phi(s_t,a_t)$으로 행동 가치 함수 $Q^{\pi_0}(s_t,a_t)$를 추정한다. 크리틱 신경망은 행동가치에 대한 벨만 방정식(식 (3-6))을 이용하여 크리틱 신경망의 시간차 타깃을 계산할 수 있다 (식 (3-21)). 학습 목표는 행동 가치를 정확하게 추정하는 것으로 y_t와 $Q_\phi(s_t,a_t)$의 차이가 작을수록 우수하다. 이에 따라 식 (3-22)를 손실함수로 크리틱 신경망을 학습한다. 여기서 N은 리플레이 버퍼(replay-buffer)에서 샘플링된 데이터 개수이다.

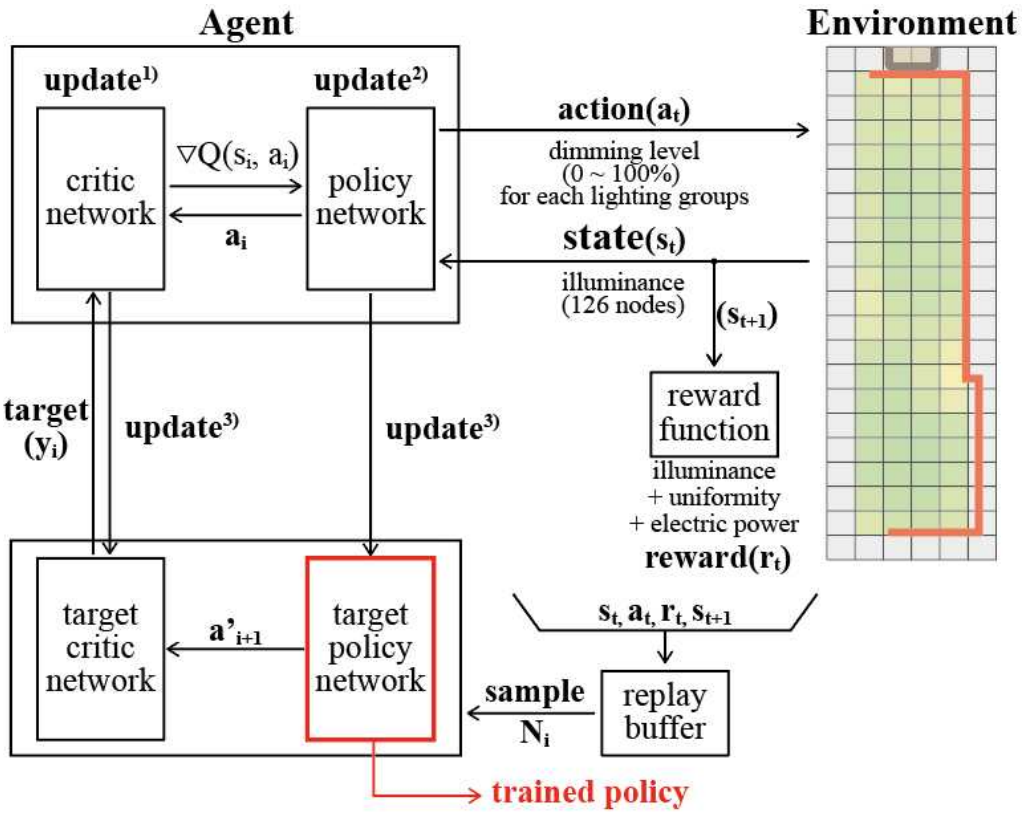
$$y_t=r(s_t,a_t)+\gamma Q_\phi(s_{t+1},\pi_\theta(s_{t+1})) \tag{3-21}$$

$$Loss=\frac{1}{2N}\sum_{t=0}^{N}(y_t-Q_\phi(s_t,a_t))^2 \tag{3-22}$$

DDPG를 이용하기 위해서는 리플레이 버퍼(replay-buffer)가 필요하다. 정책에 따라 생성된 궤적 데이터(s_t,a_t,r_t,s_{t+1})는 시간에 따른 상관관계가 있다. 딥러닝에 사용되는 확률적 경사하강법(Stochastic Gradient Descent,

SGD)은 모든 데이터가 독립이어야 하므로 데이터 간의 시간 상관성을 제거해야 한다. DDPG에서는 생성된 데이터를 일단 리플레이 버퍼에 저장한 다음, 충분한 데이터가 수집된 후에 샘플링하여 학습 데이터로 이용한다.

한편, 에이전트를 훈련할 때는, 환경을 효과적으로 탐색할 수 있어야 한다. 즉, 학습 데이터는 다양한 행동을 선택하여 환경을 탐색한 결과 이어야 한다. 확정적 정책 신경망은 확률분포함수와 달리 임의성이 없으므로 초기 행동 변수를 지속적으로 선택하게 된다. DDPG는 액터 신경망의 결과인 행동에 노이즈를 추가하여 이를 해결한다. 임의의 노이즈를 더함으로 초기값에서 벗어나 다양한 행동을 수행하면서 환경을 탐색할 수 있게 된다.



update1) : train critic network with loss function, $L = \frac{1}{2N}\sum_i (y_i - Q_\Phi(x_i, u_i))^2$
update2) : train policy network with gradient of objective function, $\nabla J_\theta(\theta)$
update3) : soft update both target networks

그림 3-2. Deep Deterministic Policy Gradient 학습 흐름도

DDPG 알고리즘의 구조는 그림 3-2와 같다. ① 현재 정책에 따라 현재 상태에 대한 행동을 도출하여 환경에 입력한다. ② 환경은 행동을 적용하여 보상과 다음 상태를 측정한다. ③ 생성된 궤적 데이터(s_t, a_t, r_t, s_{t+1})를 리플레이 버퍼에 저장한다. ④ 일정 수준의 샘플이 쌓이면 N개씩 샘플링하여 학습을 시작한다. ⑤ 시간차 타깃을 계산하고 손실함수를 이용하여 크리틱 신경망을 훈련시킨 다음 파라미터 ϕ를 업데이트한다(update[1)]). ⑥ 목적함수의 그래디언트 식으로 액터 신경망을 학습하고, 파라미터 θ를 업데이트 한다(update[2)]). ⑦ 점진적인 학습을 위해 분리했던 타겟 신경망을 각각의 파라미터로 업데이트 한다(update[3)])(Lillicrap et al., 2015).

3.4 소결

DDPG 강화학습 알고리즘은 행동 변수의 차원이 높은 경우, 확률적 정책 그래디언트 방법보다 우수함을 보여주었다(Silver et al., 2014). 건물 제어 분야에서도 최근 적극적으로 강화학습 관련 연구가 이어지고 있다. 하지만 대부분 가치기반 알고리즘을 이용한 연구이며, 제어변수가 연속인 일부 경우에 한정적으로 정책 그래디언트 알고리즘이 도입되었다(Wang & Hong, 2020)). 한 예로, Gao et al. (2019)는 설정 온도와 설정 습도를 연속 행동 변수로 선정하여 HVAC 제어를 수행하였으며, 여타 가치기반 알고리즘보다 높은 열 쾌적을 달성할 수 있었다.

본 연구에서는 실내 조명 시스템 디밍 제어에 DDPG 강화학습을 적용하여 목표 조도 환경을 만족하는 동시에 전기조명 에너지를 최소화하고자 한다. 다른 건물 제어와 달리, 조명 시스템 제어는 자연채광의 변화에 따라 실시간으로 수행되어야 한다. 최적제어 알고리즘은 제어 시점마다 현재 상태에 대한 최적화 과정을 매번 수행하여 제어변수를 탐색한다. 제어변수와 작업면이 많을수록 탐색에 많은 연산이 필요하다. 이에 비해 강화학습은 학습 기간을 통해 제어(행동) 변수를 탐색하고, 최적 정책을 도출한다. 실제로 제어를 수행할 때는 학습된 에이전트를 통해, 상태 변수에 따른 조명 시스템의 제어(행동) 변수를 즉시 결정할 수 있다. 조명 시스템의 디밍 제어변수는 연속된 값이다. 편의를 위해 단계별로 디밍 수준을 나누어 이산화하기도 하지만, 연속 공간에서 디밍 제어변수를 결정하면, 더욱 적극적인 에너지 절감이 가능하다. 본 연구에서는 작업 지점과 조명기구가 많은 경우(제어변수의 차원이 높은 경우)에 우수함이 증명된 정책 그래디언트 알고리즘인 DDPG를 도입하여 강화학습 제어를 구현한다.

제 4 장 조도 예측 모델 개발

4.1 대상 건물 및 조명 시스템

대상 건물은 인천광역시 동구에 있는 건물로, 중장비 기기의 엔진을 생산한다. 공장의 평면은 생산 공정에 따라 여러 구역으로 구분되어 있다. 본 연구에서는 공장 내, 일부 구역을 대상으로 하되, 인접 공간을 포함하여 모델을 제작하였다. 대상 구역은 가로 80m, 세로 22.5m의 개방된 공간이며, 가장 낮은 부분의 층고 (조명기기가 설치된 높이)는 7m이다 (그림 4-1. (b), 그림 4-2). 자연채광은 북측 천장에 설치된 채광창(skylight)을 통해 실내로 유입된다(그림 4-1. (a, b)).

(a) 공장 외부

(b) 공장 내부

그림 4-1. 대상 건물

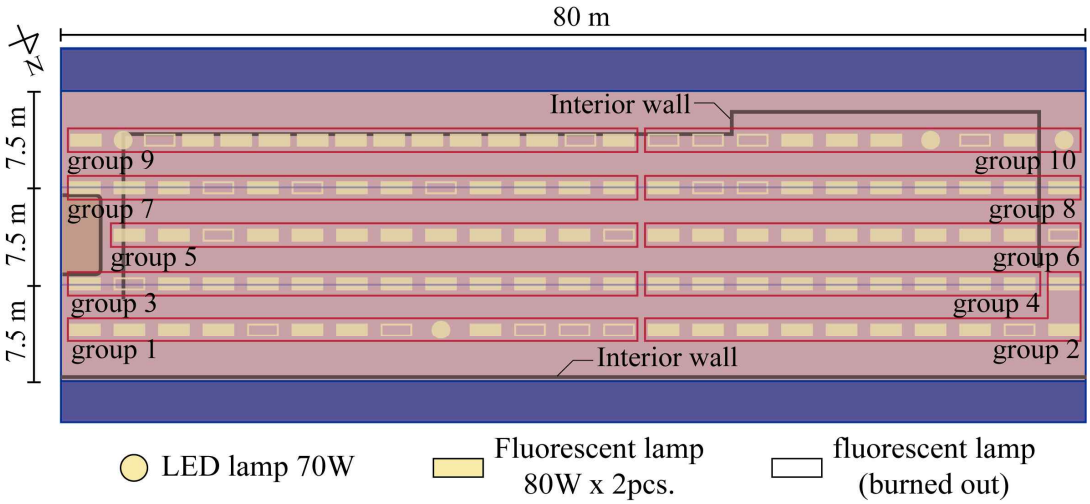


그림 4-2. 조명기기 그룹(#1-#10) 및 평면도

대상 구역에 설치된 조명기기의 수는 표 4-1과 같다. 조명기기 대부분은 형광등 조명기기(80W 형광램프 2개 매립)이며, 일부는 수명이 다하여 소등된 상태이다. 소등된 조명기기 중 일부는 70W LED 램프로 교체되었으며, 소등되지 않은 형광등과 함께 운용되고 있다. 각 조명기기는 5열로 설치되어 있으며, 동·서로 나누어 총 10개 그룹으로 구성된다(그림 4-2). 각 그룹은 원격 제어 시스템으로 연결되어 동시에 제어된다.

표 4-1. 그룹별 조명기기 개수 및 전력 사용량

Lighting group #	Number of LED lamp (70W)	Number of flourescent fixtures (80W×2 lamps, 167W)	Number of burned-out flourescent fixtures	Total wattage (W)
1	1	7	5	1,239
2	0	10	1	1,670
3	0	12	1	2,004
4	0	9	0	1,503
5	0	10	2	1,670
6	0	9	1	1,503
7	0	10	3	1,670
8	0	8	2	1,336
9	1	10	2	1,740
10	2	4	4	808
Total	4	89	21	15,143

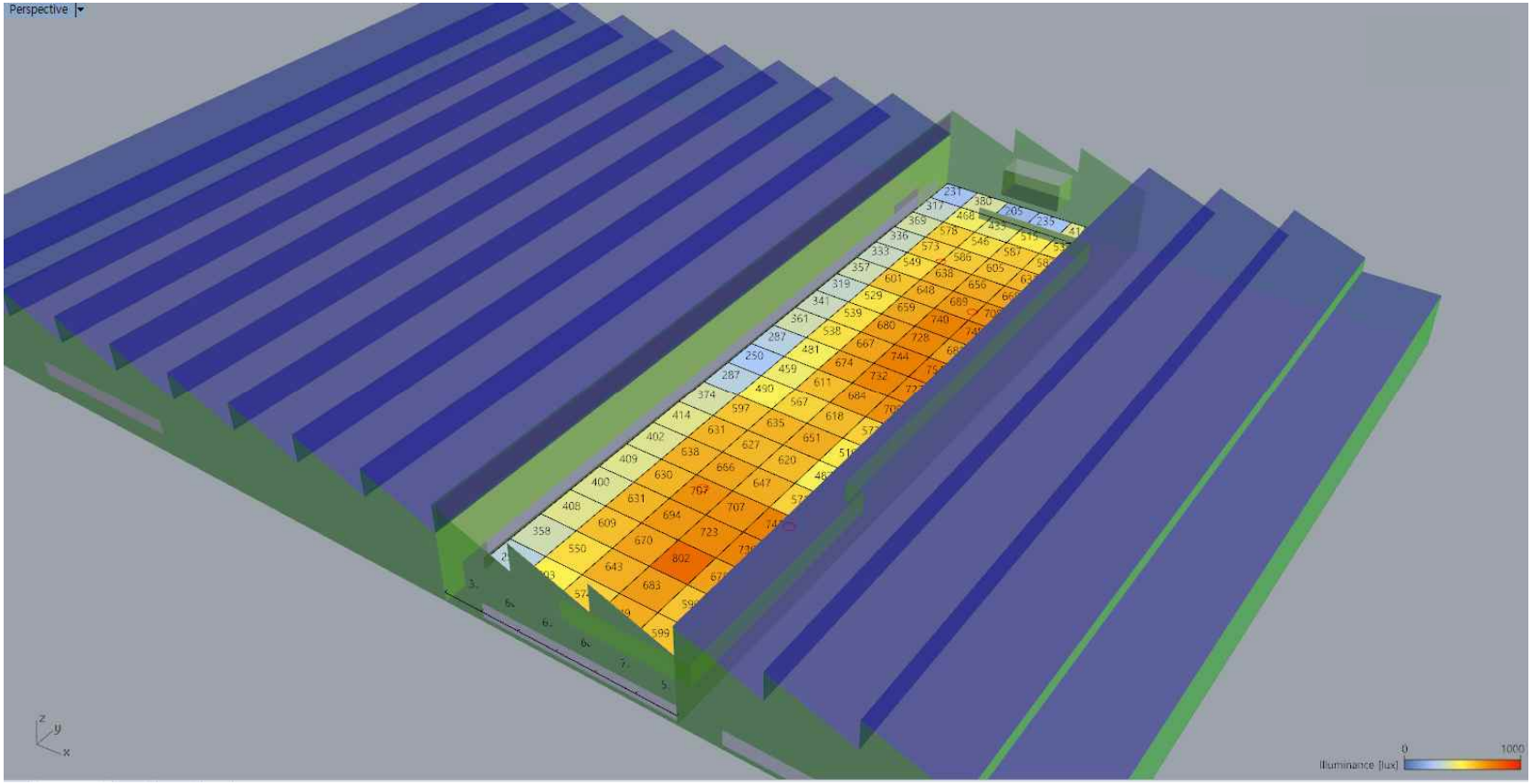
4.2 자연채광 조도 예측 모델 개발

4.2.1 자연채광 모델 작성

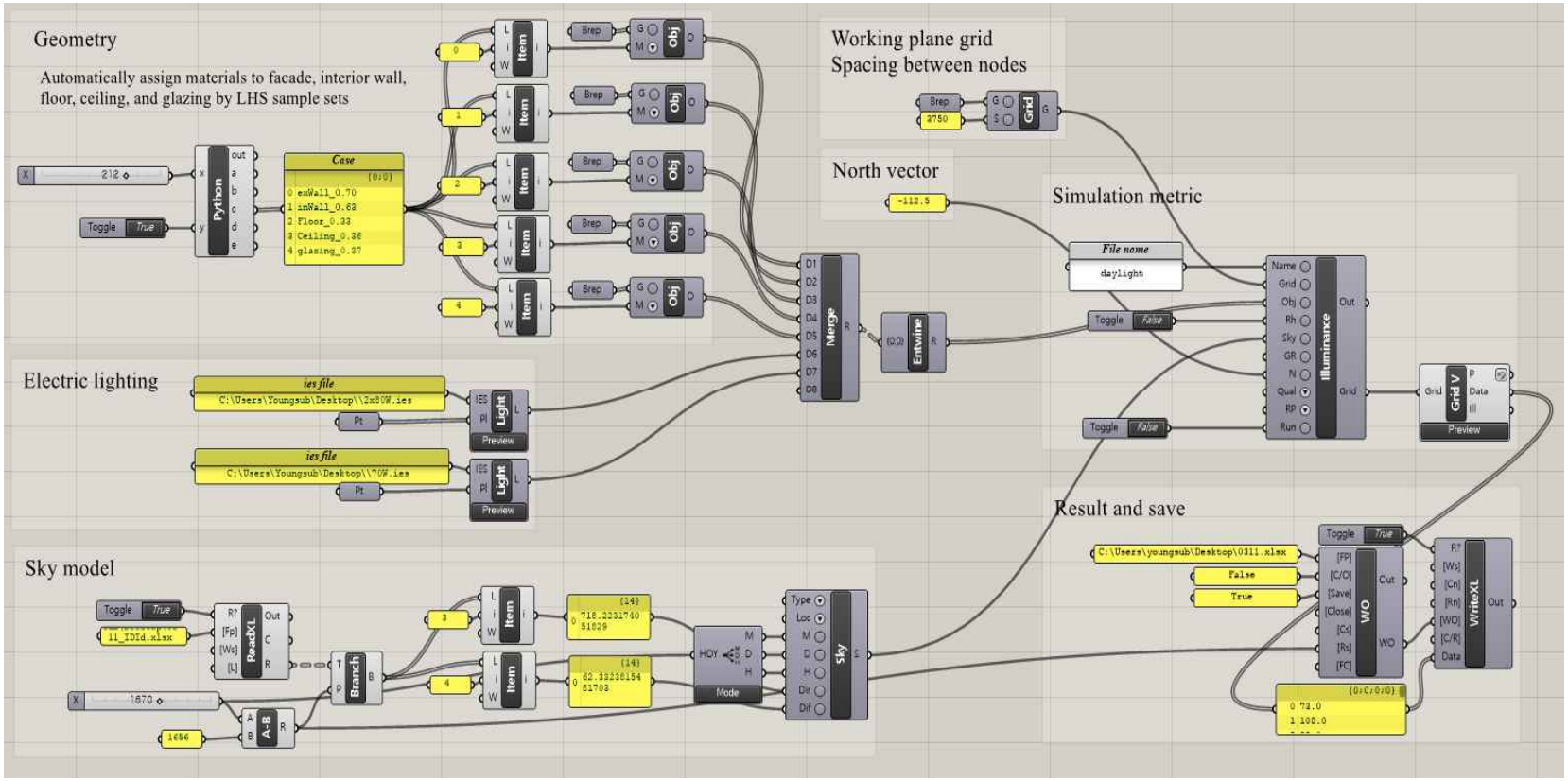
자연채광 모델과 전기조명 모델에 공통으로 필요한 3D 형상 정보는 Rhinoceros를 이용하여 작성하였다. 조도 시뮬레이션은 플러그인 소프트웨어인 DIVA(Solemma LLC, 2020; Jakubiec, & Reinhart, 2011)로 수행하였다. DIVA는 ray tracing 방법을 이용하는 Radiance(Ward, & Shakespeare, 1998)를 엔진으로 빛 환경 시뮬레이션을 수행할 수 있도록 사용자 인터페이스를 제공한다. DIVA의 조도 시뮬레이션은 1년, 8760시간 중 한 시점을 선택하여 그 순간의 조도를 예측한다.

제작된 모델의 신뢰성을 확보하기 위해서는 측정값에 기반하여 모델을 보정하고 검증하는 과정이 필요하다. 즉, 입력변수를 변경해가면서 반복 시뮬레이션을 수행할 수 있어야 한다. 이를 위해 Rhinoceros의 파라메트릭 프로그래밍 플랫폼인 Grasshopper를 이용하였다. 상술한 환경에서의 모델을 DIVA for Grasshopper 조도 예측 모델(그림 4-3, (b))이라 하며, 아래와 같은 절차로 제작하였다.

먼저 건축 도면을 바탕으로 실내 공간을 실측하여 Rhinoceros 3D 모델을 제작하였다(그림 4-3, (a)). 형상을 단순화하여 두께가 없는 면(surface)으로 제작하고 부위, 재질에 따라 레이어로 구분하였다. 조명기구의 위치는 점(point)으로 표기하였다. Brep 컴포넌트를 이용하면 Rhinoceros에서 제작된 요소(점, 선, 면, 입체)의 3차원 좌표 정보를 Grasshopper로 읽어올 수 있다. 하나의 brep 컴포넌트는 복수의 형상 요소를 포함할 수 있으므로, 내부 형상 요소(벽, 바닥, 천장, 유리창) 레이어마다 한 개씩 지정하면 된다. 다음으로, 각 brep 컴포넌트에 물성값을 지정해야 한다. 각 면의 반사율과 유리의 투과율은 가정하였고, 추후 측정 데이터를 이용하여 보정하였다 (4.2.2절 후술). 조도를 측정할 지점(작업면)을 지정하기 위해, 바닥에서 1m 높이에 가상의 수평면을 작성하였다. 센서의 간격을 3.75m로 지정하여 총 126개(21X6) 정사각형 그리드를 제작하였다.



(a) Rhinoceros 형상 모델



(b) DIVA for Grasshopper 시뮬레이션 모델

그림 4-3. DIVA 조도 예측 모델

DIVA는 일시, 시간에 따른 태양의 위치(방위각, 고도각)와 천공 상태에 따라 자연채광 조도를 계산한다. 이는 천공 모델의 형태로 DIVA에 입력되며, 본 연구에서는 Perez all weather 천공 모델(Perez, Seals, & Michalsky, 1993)을 선정하였다. 이 모델은 천구의 밝기계수와 청명계수를 토대로 sky patch의 분포를 계산하는데, 두 계수는 법선면 직달일사(DNI: direct normal irradiance)와 수평면 산란일사(DHI: diffuse horizontal irradiance)의 비율로 산정된다. 먼저 기상청의 수평면 총 일사량(GHI)를 수집하고(KMA, 2020), Watanabe 모델(Watanabe el al. 1983)을 적용하여 법선면 직달일사(DNI)와 수평면 산란일사(DHI)를 계산하였다. 정리하면, 자연채광을 설명하는 천공 모델의 입력변수로 일시, 시간, DNI, DHI가 있으며, sky-patch 별로 계산된 광량 분포를 DIVA에 입력하는 것이다. DIVA는 측정 지점 마다 역 방향 ray-tracing 방법으로 광선을 추적하여, 전체 126개 지점의 자연채광 조도를 예측한다 (그림 4-6의 Phase I).

4.2.2 자연채광 모델 보정 및 검증

4.2.1절의 모델 제작 과정에서 가정한 미지(味知) 변수(벽, 천장, 바닥의 반사율 및 유리의 투과율)를 보정하기 위해, 작업면의 자연채광 조도를 측정하였다. 측정 장비는 “TES1339R” 조도 센서로 1분마다 측정값을 저장한다. 작업자의 과업 위치가 수시로 변화하는 특성에 따라 현장 관리자와 상의하여, 작업을 방해하지 않도록 센서 위치를 선정하였다(그림 4-4). 1월 9일-10일 동안 1시간 간격으로 총 16 시점에 대해 5개 지점에서 측정된 자연채광 조도를 보정 작업에 이용하였다 (80개 관측점 = 5개×16시간).

미지 변수 보정은 측정 데이터와 시뮬레이션 예측값의 차이가 가장 작은 변수 조합을 구하는 과정이다 (식 (4-1)). 미지 변수 조합의 개수는 무한대이며, 이를 모두 검토하는 것은 불가능하다. 본 연구에서는 몬테카를로 방법의 하나인 라틴 하이퍼큐브 샘플링(Latin Hypercube Sampling, LHS)으로 300개의 미지변수 조합을 생성하여, 가장 우수한 조합을 구하였다.

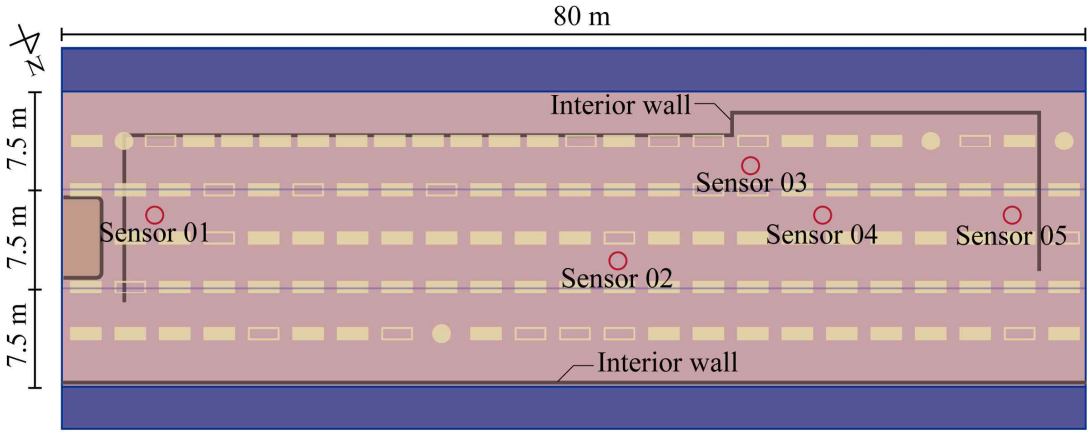


그림 4-4. 모델 보정 및 검증용 조도 센서 설치 위치

$$\arg\min_{i}\left(\sum_{t=1}^{16}\sum_{n=1}^{5}\left(E_{d,p}(i;k;t)-E_{d,m}(k;t)\right)\right) \quad (4-1)$$

여기서,

$E_{d,p}$: 예측된 자연채광 조도 [lx]

$E_{d,m}$: 측정된 자연채광 조도 [lx]

i: LHS를 통해 구한 i번째 변수 조합

k: 자연채광 및 전기조명 조도 측정점 $(k=5)$

t: 측정 시간

300회의 DIVA 시뮬레이션을 위해, 그래스호퍼에서 파이썬 코드를 실행하는 컴포넌트인 GhPython을 이용하였다. 파라메트릭하게 변경된 DIVA 미지변수에 대한 시뮬레이션이 수행되면, TT Toolbox를 통해 Excel 파일로 그 결과를 저장하고, 다음 미지변수 조합으로 변경하는 자동화 루프를 작성하였다. 변수별 탐색 범위는 대상 구역의 부위별 재질 및 유지관리 현황을 기반으로 설정하였다(그림 4-5).

(a) 벽체

(b) 바닥

(c) 유리창

그림 4-5. 대상 구역 부위별 재질 및 유리관리 현황

일반적인 경우, 벽체와 바닥, 천장의 반사율은 각각 0.5, 0.2, 0.7로 가정하므로, 이를 중간값으로 하여 탐색공간을 설정하였다 (Di Laura et al. 2011). 유리창의 투과율은 천장에 위치하여 유지관리가 어렵고, 부착된 기구(장애물)의 영향이 큰 점을 반영하여 일반적인 수준보다 낮은 0.5로 중간값을 설정하였다. 보정 결과로 모델링 결과와 실측값의 차이를 최소화하는 미지변수를 표 4-2와 같이 추정하였다.

표 4-2. 미지변수 추정 범위 및 보정값

Unknown Variable	Min	Median	Max	Calibrated value
Wall reflectance	0.3	0.5	0.7	0.54
Floor reflectance	0.1	0.2	0.4	0.21
Ceiling reflectance	0.5	0.7	0.9	0.61
Windows Visible Transmittance	0.3	0.5	0.7	0.34

표 4-2에서 추정된 파라미터 값을 시뮬레이션 모델에 대입한 결과, 측정값과 시뮬레이션 값 간의 RMSE(Root Mean Square Error)는 34.9 lux로, 시뮬레이션 모델이 충분히 정확함을 알 수 있다(그림 4-6의 Phase II). 또한, ASHRAE Guideline 14 (2002)에서 시간별 측정일 경우, 평균편향오차(Mean Bias Error, MBE)는 10% 미만, 변동계수(Coefficient of Variation of the Root Mean Square Error, CVRMSE)는 30% 미만을 권장한다(식 (4-2), (4-3)). 측정값과 시뮬레이션 값의 MBE는 4.9%, CVRMSE는 24%이다.

$$MBE = \frac{\sum_{i=1}^{n}(M_i - S_i)}{\sum_{i=1}^{n} M_i} \times 100 \qquad (4\text{-}2)$$

$$CVRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}(M_i - S_i)^2}{n}}}{\overline{M}} \times 100 \qquad (4\text{-}3)$$

여기서,

M_i : 실체 측정값

S_i : 실체 측정값

n : 데이터의 수

$\overline{M}$: M_i의 평균값

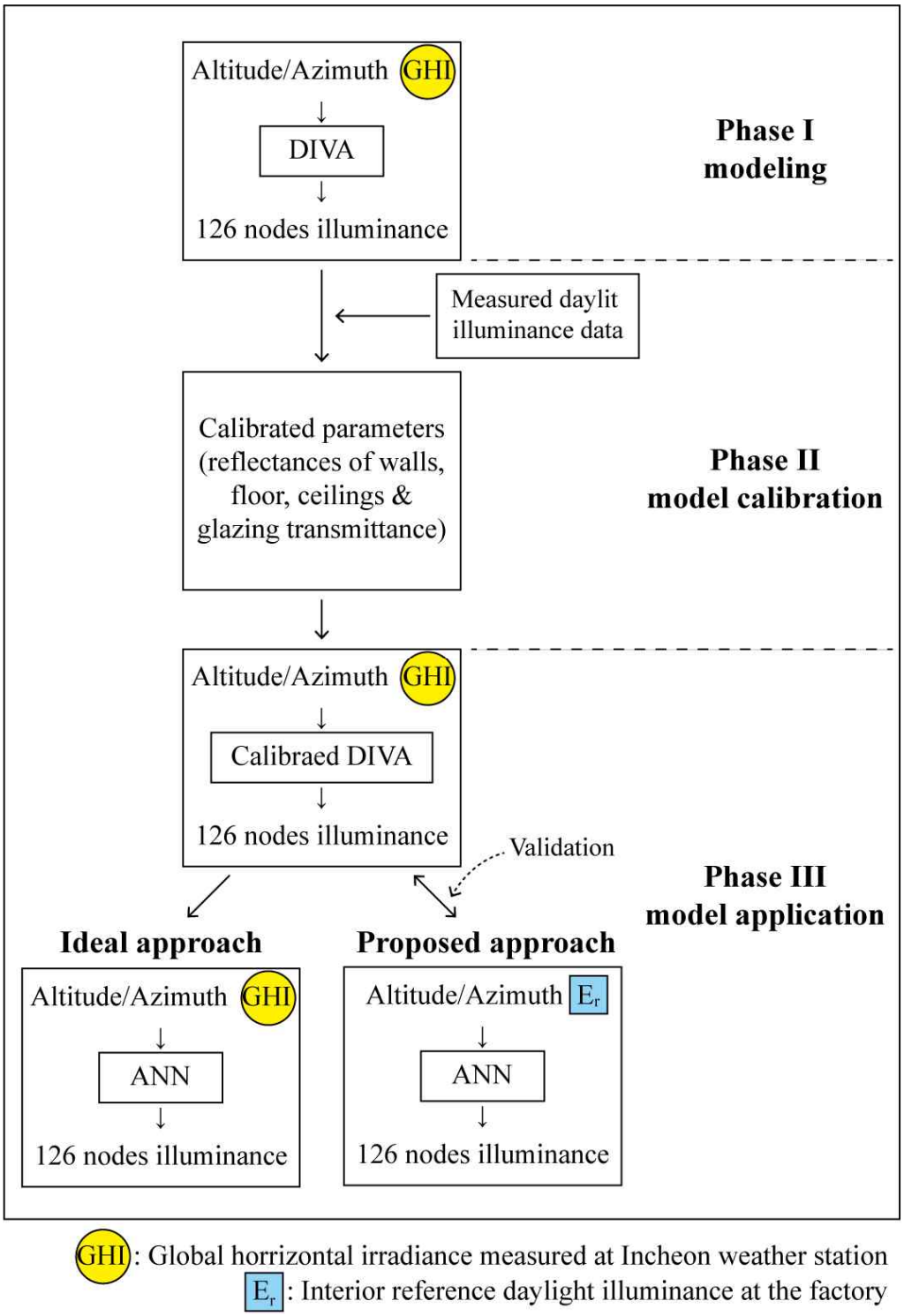


그림 4-6. 자연채광 조도 예측 모델 제작 방법
(이상적인 접근 vs. 제안하는 접근)

4.2.3 자연채광 대리모델 작성

조도 예측 모델을 이용하여 조명 시스템을 제어하기 위해서는 제어 시점마다 입력변수를 받고, 작업면의 조도를 실시간으로 예측할 수 있어야 한다. DIVA는 물리법칙 기반(ray tracing)의 몬테카를로 시뮬레이션 방법을 적용하므로, 많은 연산시간을 요구하여, 실시간 제어에 적용하기 어렵다. 따라서, 본 연구에서는 DIVA를 모사하는 대리모델(surrogate model)을 ANN (Artificial Neural Network)으로 개발하였다(김재민 외, 2019).

ANN은 입력층과 출력층 사이에 여러 노드 및 은닉층을 가진 퍼셉트론 구조이다 (Bishop, 2006). 각각의 노드별로 가중치(weight)와 편향(bias)을 조절하며 입력과 출력 사이의 관계를 학습한다(그림 4-7). 그림 4-6의 Phase III에서, 이상적인 접근(ideal approach)의 대리모델 제작은 일시와 시간 및 대상 건물의 위치, 그곳에서 측정된 수평면 총 일사량(GHI) 등의 변수를 입력하여, 자연채광 조도를 예측하는 것이다. 하지만, 인천 기상청에서 제공하는 GHI 정보는 실시간으로 제공되지 않으며, 대상 건물이 위치한 지역적인 특징을 반영하지 못한다. 대상 건물이 위치한 지점에 일사계를 설치하여 관련정보를 실시간으로 확보하는 것이 최선이지만, 현장 여건상, 센서를 설치할 수 없었다.

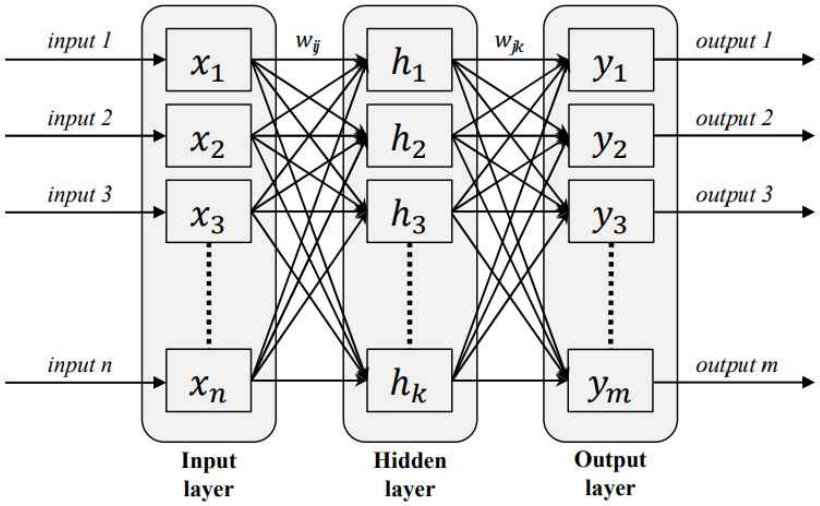
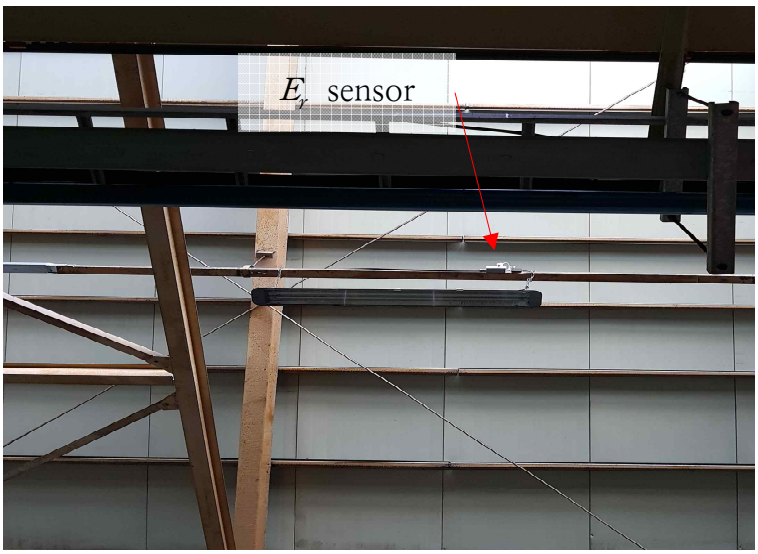


그림 4-7. ANN 모델의 구조(박성호, 2019)

대안으로, 실내 천장 부근 (채광창 근처)에 수평면 조도계를 설치하였다. GHI 값이 클수록 천장 근처의 실내 수평면 자연채광 조도 또한 커질 것으로 가정하였다. 센서는 조명기구보다 높은 위치에 설치되어 자연채광 조도만을 측정한다. 측정값은 실시간으로 무선 네트워크를 통해 원격 조명 제어 시스템으로 전송된다(지영민, 옥기수, 권동우, 2019) (그림 4-8).



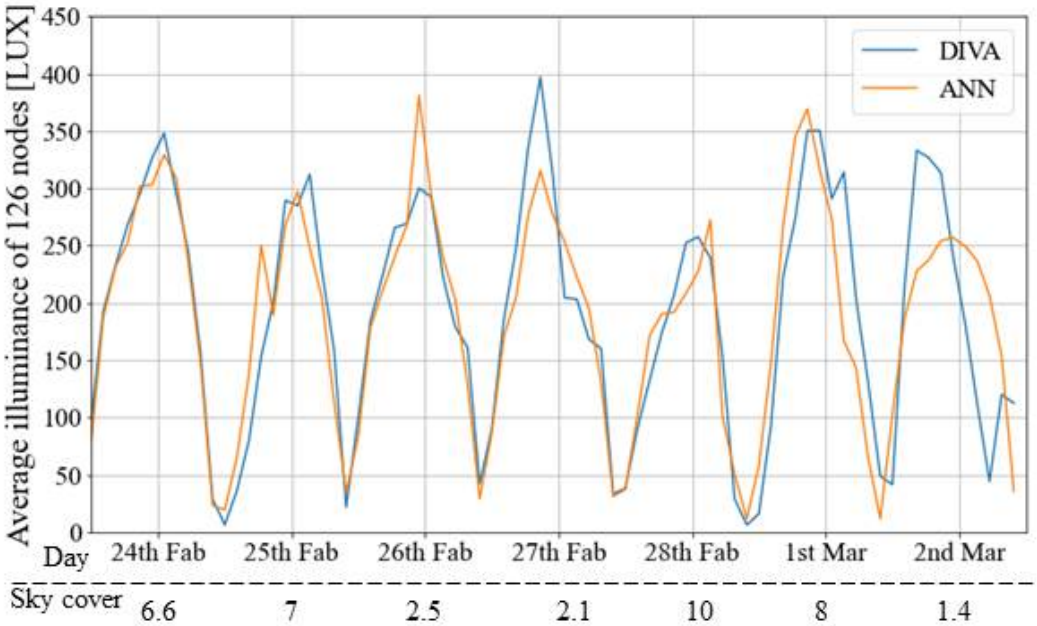
(a) E_r 센서 설치 위치

(b) E_r 센서

그림 4-8. 자연채광 조도 센서(E_r)

본 연구에서는 이상적인 접근(ideal approach)의 GHI를 센서에서 측정된 E_r로 대체하여 ANN 모델을 제작하였다 (그림 4-6, 제안된 접근(Proposed approach)). 먼저 시간별로 기상청에서 수집된 GHI와 태양의 위치 정보(방위각, 고도각)를 입력변수로 한 DIVA 모델로 시뮬레이션을 수행하여 작업면 조도 데이터(출력변수)를 구하였다. 모델의 입력변수 데이터에 출력변수를 병합하여 입/출력 데이터 세트를 작성한 다음, GHI를 동일한 시간에 측정된 E_r로 대체한 것이다. 최종 제작된 ANN 모델은 입력층 노드 3개(방위각, 고도각, E_r), 출력층 126개(4.2.1절의 그리드, E_{daylit})로 구성된다.

모델 학습 및 검증을 통해 GHI 값을 대체할 수 있는지 확인하였다. ANN 모델 학습에는 1월 17일 ~ 2월 23일 동안의 시간 단위 데이터를 사용하였으며, 검증은 2월 24일 - 3월 2일의 데이터로 수행하였다(그림 4-9). 대리모델의 MBE는 1.4%이고, CVRMSE는 25.6%이다 (그림 4-6의 Phase III, Validation에 해당).



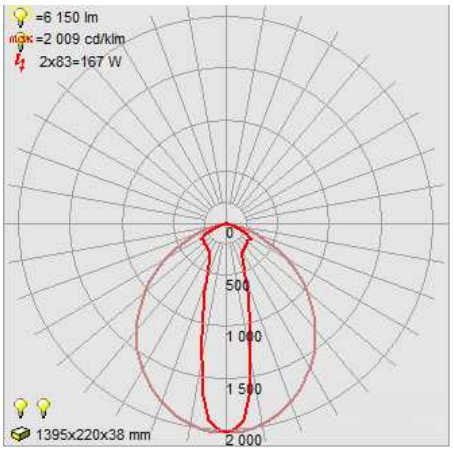
Sky cover = 0 : 청천공(Clear sky)

Sky cover = 10 : 담천공(Overcast sky)

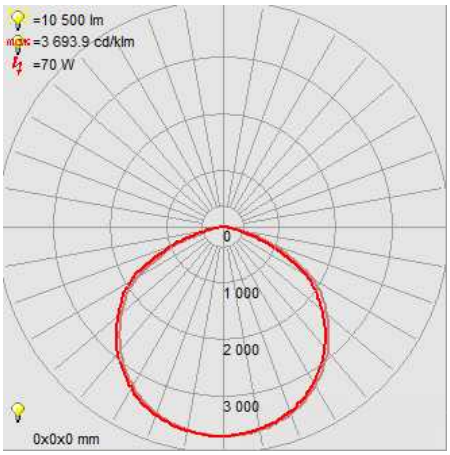
그림 4-9. DIVA 시뮬레이션 vs ANN 대리 모델 비교

4.3 전기조명 조도 예측 모델 개발

전기조명 모델을 DIVA에서 구현하기 위해 조명기구 제조사에서 제공하는 배광파일 (ies 파일)이 필요하다. 배광파일은 조명기구의 광도 분포 곡선 (Candlepower Distribution Curve, CDC) 및 조명 기구의 스펙 정보를 포함한다. 본 연구에서는 해당 조명기구 제조사 (그림 (4-2)의 LED 조명기구, 형광등 조명기구 제조사) 홈페이지에서 배광파일을 다운받아 사용하였다 (그림 4-10).



(a) fluorescent lamp



(b) LED lamp

그림 4-10. 조명기구별 배광 곡선(Candlepower Distribution Curve, CDC)

전기조명 예측을 위해서는 미지변수인 감광보상률(Light Loss Factor, LLF)을 결정해야 한다. LLF는 세부적으로 회복 가능한 LLF와 그렇지 못한 LLF로 구성된다. 회복이 불가능한 LLF는 조명 기기 자체적인 요인과 설치 환경적 요인으로 인해 일반적인 유지보수로는 회복할 수 없는 감광을 반영하는 계수이다. 반면, 회복 가능한 LLF의 경우, 내부 청결도, 유지관리 주기, 조명기기 연식 등에 따라 경험에 기반하여 제작된 식에 의해 계산할 수 있다 (Di Laura et al., 2011). 하지만, 이렇게 계산된 LLF라 할지라도 실제와 차이가 발생하므로, 편의성을 위해 미지변수로 가정하게 된다. 본 연구에서는 자연채광 모델의 미지변수를 모델 보정에 이용한 것과 같이, LED 및 형광등 조명기구 각각의 LLF를 모델 보정을 위한 추정 변수로 이용하였다. 먼저 라틴 하이퍼큐브 샘플링(LHS) 방법으로 300개의 LLF 후보 조합(LED LLF, 형광등 LLF)을 생성하였다. 다음으로 전기조명 시뮬레이션 모델의 예측값과 실측값의 차이를 최소화하는 미지변수 조합을 식 (4-3)으로 탐색하였다. 사용한 실측 데이터는 야간 시간 (저녁 8시), 모든 조명기구를 점등한 상태로 5개 지점에서 작업면 조도를 측정한 것이다.

$$\arg\min_{i} \left(\sum_{t=1}^{16} \sum_{n=1}^{5} \left(E_{d,p}(i;k;t) - E_{d,m}(k;t)\right)\right) \qquad (4-3)$$

여기서,

$E_{e,p}$: 예측된 전기조명 조도 [lx]

$E_{e,m}$: 측정된 전기조명 조도 [lx]

i: LHS를 통해 구한 i번째 변수 조합

k: 자연채광 및 전기조명 조도 측정점 ($k=5$)

LLF는 배광파일의 CDC 비율을 조정하는 방법으로 적용하였다. 보정 결과, LED LLF는 0.58, 형광등 LLF는 0.69이다. 보정 변수를 적용한 전기조명 모델의 시뮬레이션 결과값과 측정값의 RMSE는 41.7 lux, MBE는 3.7%, CVRMSE는 7.7%로, 모델의 정확성이 적절함을 판단하였다.

DIVA 조도예측모델은 전기 조명기기의 제어변수가 이산적인 경우(ON·OFF 제어)는 별도의 데이터 가공 없이 곧바로 시뮬레이션할 수 있다. 하지만, 디밍과 같이 연속적인 제어변수에 대한 시뮬레이션을 위해서는 제조사에서 제공한 배광파일을 가공하여 디밍 수준에 따라 모두 생성해야 한다. 한편, 제어변수와 작업면의 조도, 제어변수와 전력량의 관계를 각각 선형으로 근사할 수 있다. 따라서 본 연구에서는 조명기기 그룹별로 100% 디밍 수준일 때의 작업면 조도를 DIVA 모델로 구한 다음, 선형 근사를 적용하여 전기조명 시스템 모델을 작성하였다. 식 (4-4)는 제어변수와 작업면 조도 사이의 상관관계를 설명하는 매트릭스이다.

$$A = \begin{array}{c} \begin{bmatrix} 159 & 0 & 29 & 0 & 9 & 0 & 2 & 0 & 1 & 0 \\ 156 & 0 & 114 & 0 & 15 & 0 & 7 & 0 & 2 & 0 \\ 34 & 0 & 7 & 0 & 57 & 0 & 15 & 0 & 8 & 0 \\ 9 & 0 & 7 & 0 & 57 & 0 & 63 & 0 & 36 & 0 \\ 2 & 0 & 4 & 0 & 15 & 0 & 153 & 0 & 141 & 0 \\ 0 & 0 & 1 & 0 & 8 & 0 & 40 & 0 & 146 & 0 \\ & & & & & \vdots & & & & \\ 5 & 71 & 30 & 19 & 8 & 9 & 4 & 1 & 1 & 0 \\ 5 & 70 & 126 & 69 & 15 & 18 & 13 & 5 & 4 & 0 \\ 2 & 18 & 126 & 69 & 55 & 69 & 27 & 11 & 12 & 1 \\ 1 & 10 & 28 & 18 & 55 & 69 & 124 & 50 & 23 & 1 \\ 1 & 3 & 14 & 10 & 15 & 18 & 125 & 50 & 100 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 31 & 13 & 104 & 3 \\ & & & & & \vdots & & & & \\ 0 & 110 & 0 & 38 & 0 & 10 & 0 & 4 & 0 & 1 \\ 0 & 107 & 0 & 152 & 0 & 19 & 0 & 15 & 0 & 4 \\ 0 & 23 & 0 & 146 & 0 & 61 & 0 & 32 & 0 & 10 \\ 0 & 8 & 0 & 32 & 0 & 62 & 0 & 146 & 0 & 31 \\ 0 & 2 & 0 & 13 & 0 & 16 & 0 & 146 & 0 & 119 \\ 0 & 0 & 0 & 3 & 0 & 8 & 0 & 35 & 0 & 113 \end{bmatrix} \end{array} \begin{array}{l} node\,1 \\ node\,2 \\ node\,3 \\ node\,4 \\ node\,5 \\ node\,6 \\ \vdots \\ node\,67 \\ node\,68 \\ node\,69 \\ node\,70 \\ node\,71 \\ node\,72 \\ \vdots \\ node\,121 \\ node\,122 \\ node\,123 \\ node\,124 \\ node\,125 \\ node\,126 \end{array} \qquad (4-4)$$

여기서,

A: 그룹별 제어변수가 100%일 때와 작업면 조도의 관계 매트릭스

$a_{i,j}$: ‘j’ 조명기기 그룹에 의한 ‘i’ 지점의 조도[lx]

각 그룹의 디밍 제어는 [0~100]의 연속 변수로 수행되며, 식 (4-5)의 형태로 시스템 모델에 입력된다. 마지막으로 디밍 제어변수 조합에 따른 작업면의 조도는 식 (4-6)과 같이 계산된다.

$$D = \left[d_1\, d_2\, d_3\, d_4\, d_5\, d_6\, d_7\, d_8\, d_9\, d_{10} \right]^T \tag{4-5}$$

여기서,

D: 조명 시스템 모델의 입력 변수 벡터

d_j: 'j'그룹의 디밍 수준[%]

$$A \times D = E_{elec} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{124} \\ e_{125} \\ e_{126} \end{bmatrix} \tag{4-6}$$

여기서,

E_{elec}: 선형시스템 모델로 예측된 작업면 전기조명 조도 벡터

e_i: 'i'지점의 전기조명 조도 [lx]

4.4 모델 통합 및 검증

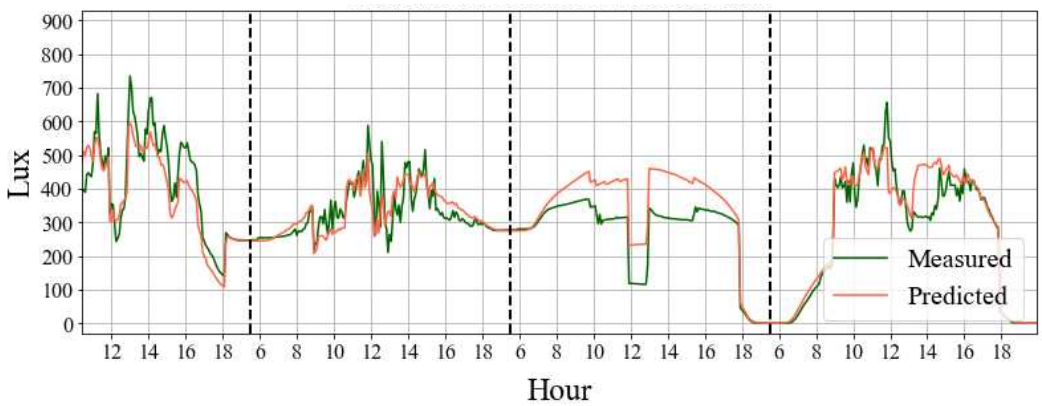
4.3장까지 개발한 자연채광 대리모델 및 전기조명 모델의 예측값을 합하면 작업면 총 조도(E_T: total of daylit and electric lighting illuminances)를 예측할 수 있다. 통합 모델을 검증하기 위해 3월 9일~12일의 기간 동안 작업면의 5개 지점(그림 4-4)에서 조도를 측정하였다. 측정된 조도는 1분 간격으로 저장하였다 (관측점: 25,500개(60분×85시간×5지점)).

통합 모델의 시뮬레이션 값(E_t)과 측정값의 차이는 그림 4-11과 같다. 3월 9일, 10일, 12일의 모델 예측 정확도는 비교적 높지만, 11일은 모델의 오차가 크다 (표 4-3의 CVRMSE 참조). 전기조명 조도의 경우, 시간에 따른 변화가 거의 없고, 모델 예측값과 실측값의 오차가 크지 않다. 따라서, 3월 11일 오차를 설명하기 위해, 자연채광 대리모델의 출력값과 자연채광 실측값을 비교하였다 (그림 4-12). 여기서 자연채광 실측값은 수평면 총 조도 측정값에서 전기조명 모델의 예측값을 뺀 것이다.

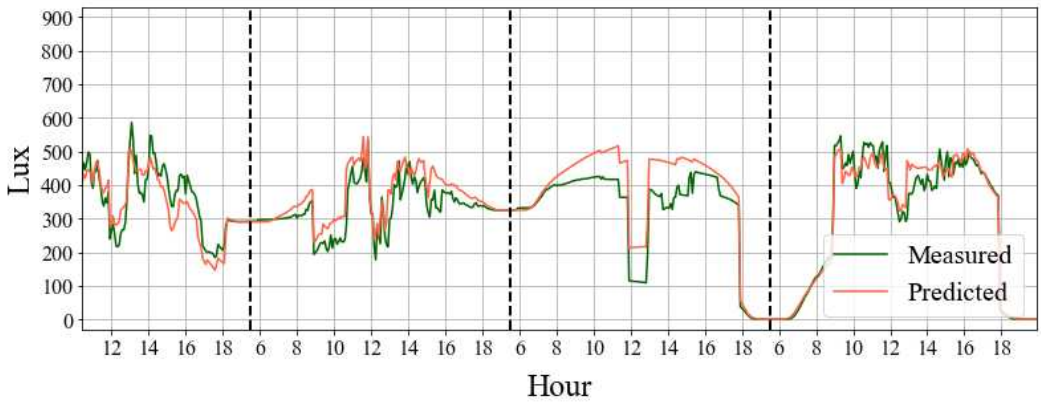
3월 11일의 경우, 평균 전운량이 0/10이고, 누적 일사량이 20.37MJ/m2인 청천공이었다 (표 4-3). 청천공인 날은 수평면 총 일사량(GHI) 및 법선면 직달 일사량(DNI) 값이 매우 크다. 하지만, 대상 건물의 채광창이 북향이라서, GHI 및 DNI가 실내 자연채광 조도에 미치는 영향은 청천공이 아닌 날에 비해 상대적으로 낮다. 기계학습 모델 제작 시, 청천공 상태에서의 데이터 부족으로 인해 예측 성능이 떨어진 것으로 판단하였고, 추후 보완이 필요하다.

표 4-3. 모델 검증 기간, 천공 상태 및 모델 정확도 (Mar 9-12, 2020)

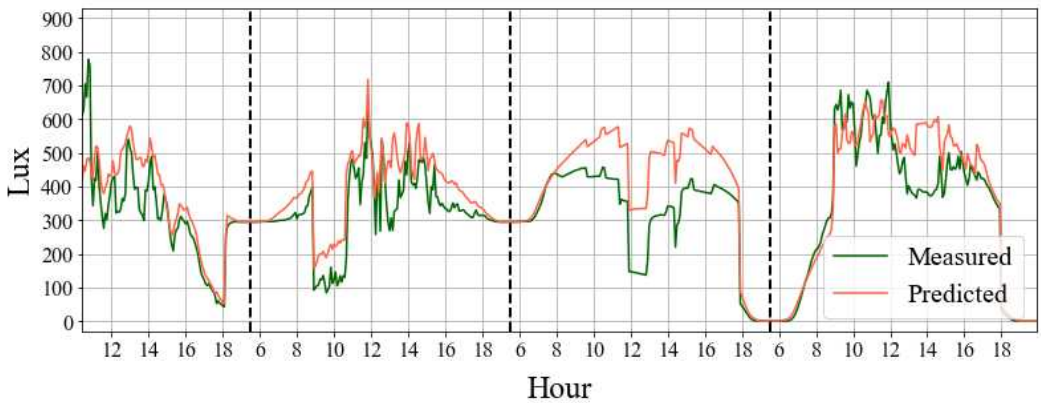
	Mar 9	Mar 10	Mar 11	Mar 12
Total cloud cover (1/10)	7.3	7.5	0	5.6
Total solar radiation quantity (MJ/m^2)	12.74	2.68	20.37	14.86
CVRMSE (%)	20	22	34	17



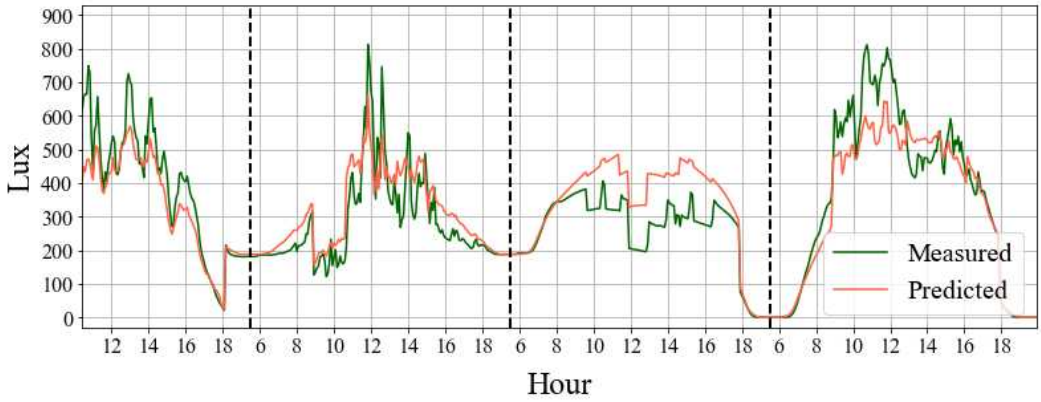
(a) Sensor #1



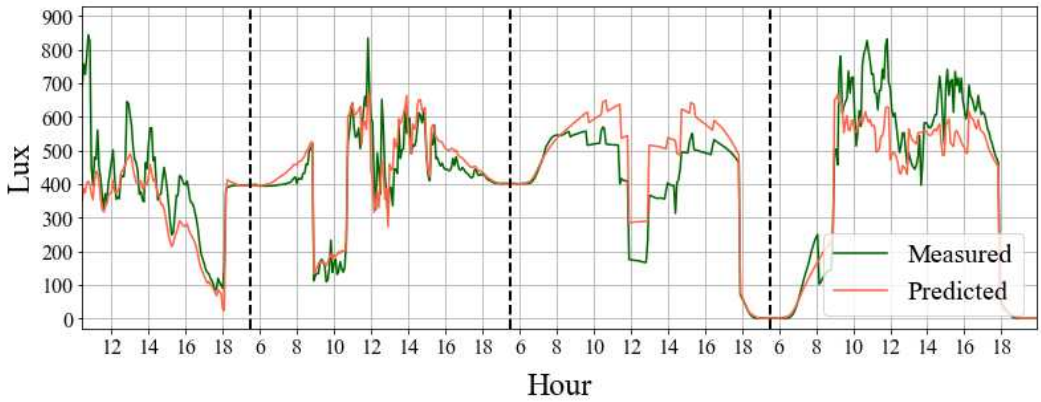
(b) Sensor #2



(c) Sensor #3

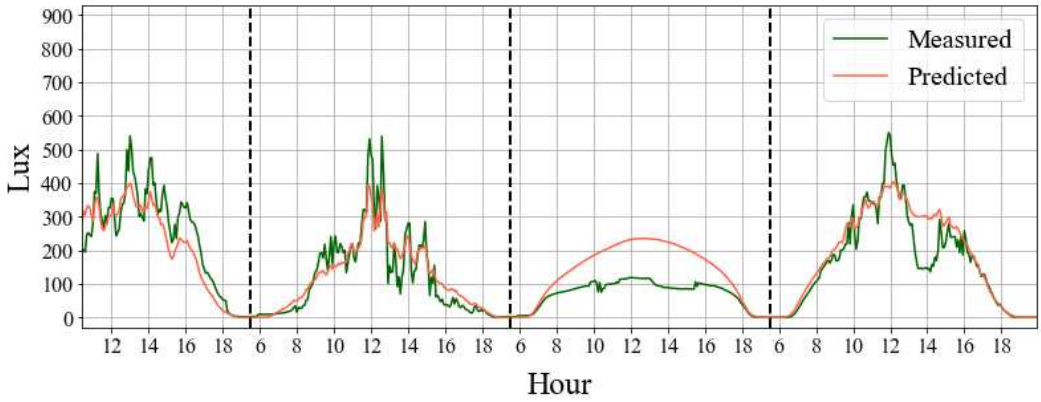


(d) Sensor #4

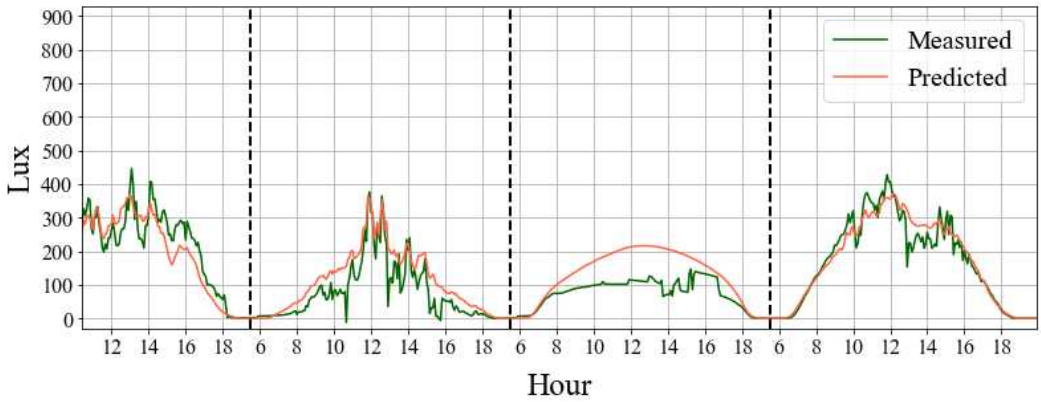


(e) Sensor #5

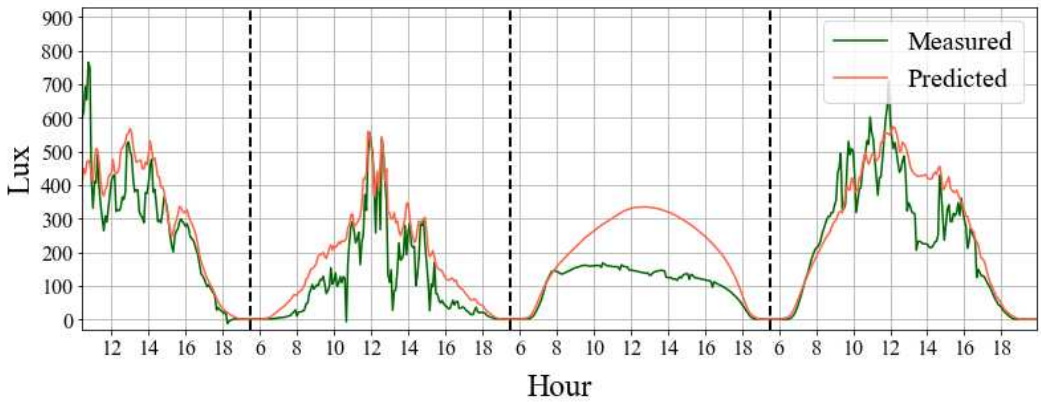
그림 4-11. 총 조도의 실측값($E_{t,measured}$) vs 예측값($E_{t,predicted}$) 비교 (Mar. 9-12, 2020)



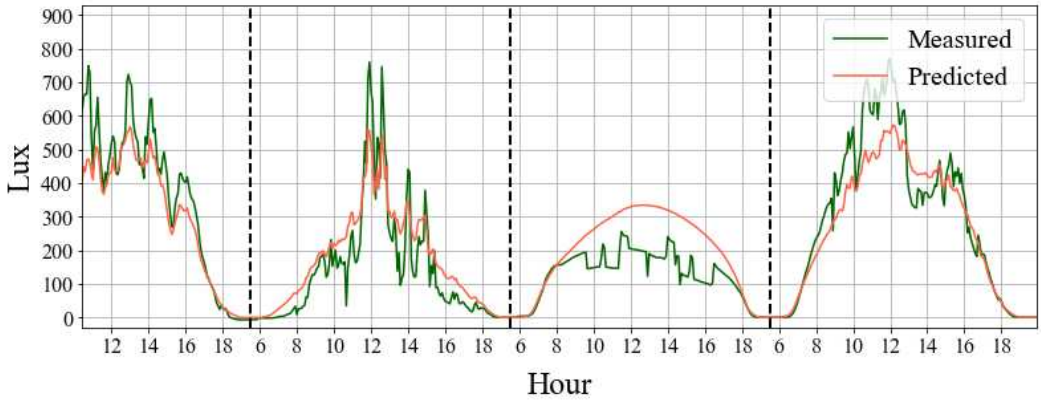
(a) Sensor #1



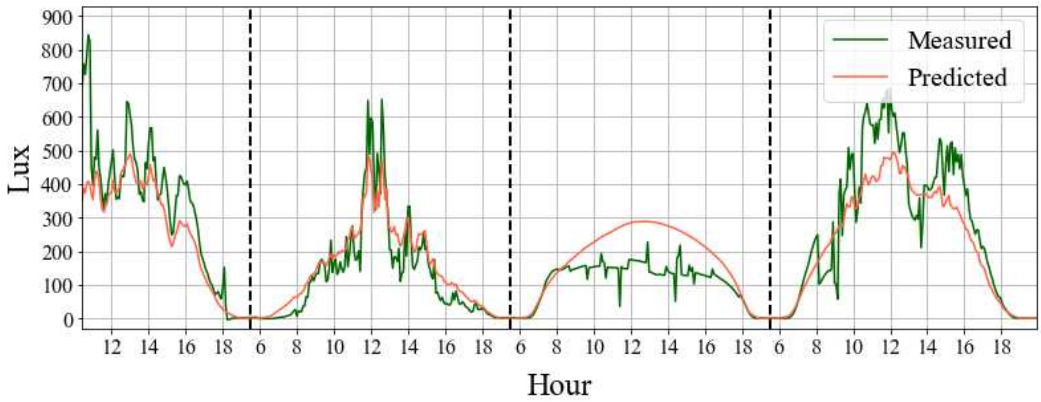
(b) Sensor #2



(c) Sensor #3



(d) Sensor #4



(e) Sensor #5

그림 4-12. 자연채광 조도 실측값($E_{d,measured}$) vs 예측값($E_{d,predicted}$) 비교 (Mar. 9-12, 2020)

제 5 장 DDPG 강화학습 조명 제어

5.1 강화학습 환경(environment) 제작

5.1.1 에피소드

강화학습 에이전트를 훈련하기 위해서는 시스템 환경이 필요하다. 훈련을 위한 시스템 환경 제작을 위해서는 시작 지점과 종료 조건 및 두 지점 사이를 일정 간격으로 나누어 제어를 수행할 시점을 지정해야 한다. 또한, 각 제어 시점에서는 현재 상태와 제어변수(행동)를 입력받은 후, 다음 상태와 보상을 출력해야 한다. 시작 지점부터 순차적으로 제어를 수행하다가 종료 조건에 도달하는 것을 하나의 에피소드(episode)라고 한다. 본 연구에서는 3월 9일~12일, 9시~19시를 15분 단위로 나누어 제어를 수행하는 것을 하나의 에피소드(총 160회)로 제작하였다. 시작 지점은 3월 9일 9시, 모든 조명기기가 꺼진 상태이며, 종료 조건은 3월 12일 18시에 도달하는 것이다.

상태는 작업면 그리드(126개 지점)의 작업면 총 조도, $E_{T,t}$이고, 행동은 조명기기 그룹별 디밍 제어변수, D이다. 각 제어 시점에 환경이 다음 상태(s_{t+1})로 출력하는 값은 다음 시점의 자연채광 조도 예측 모델 결과인 $E_{daylit,t+1}$와, $a_t = D$를 전기조명 모델에 입력하여 얻은 전기조명 조도인 $E_{elec,t}$의 합이 되어야 한다. 하지만, 조명 제어의 경우, 현재의 제어변수가 15분 뒤인 다음 시점의 조도 환경 변화에 미치는 시간 지연(time delay)이 거의 없다. 또한, 제어에 대한 보상 함수 계산도 미래 상태가 아니라 현재 상태에 대해 수행되어야 한다. 이와 같은 실시간 조명 제어의 특성을 반영하기 위해, 본 연구에서는 원래의 상태, $s_t = E_{T,t}$ 대신, $s_t = E_{daylit,t}$로 대체하였다. 환경이 출력하는 다음 상태 역시 다음 시점의 자연채광 조도인 $E_{daylit,t+1}$가 된다. 또한, 실시간 조명 제어에서는 미래에 받게 될 보상은 고

려하지 않아도 된다. 이에 따라 식 (3-1)에서 감가율(discount rate), γ를 0으로 설정하였다.

한편, 전기조명 모델은 보상 함수를 계산하는 데 이용된다. 먼저 현재 상태, $s_t = E_{daylit,t}$(이전 제어 시점의 다음 상태, s_{t+1})에 따라 정책으로 결정된 행동, $a_t = D$를 전기조명 모델에 입력하여 전기조명 조도, $E_{elec,t}$을 구한다. 다음으로, $E_{T,t} = E_{daylit,t} + E_{elec,t}$를 대상으로 보상 함수를 적용하여 r_t를 계산한다. 정리하면, 제어 시점마다 환경은 행동, $a_t = D$을 이용하여 다음 상태$s_{t+1} = E_{daylit,t+1}$와 보상, $r_t(E_{T,t})$을 출력한다. 다음 제어 시점에서 정책은 이전 시점의 상태($s_{t+1} = E_{daylit,t+1}$)를 현재 상태($s_t = E_{T,t}$)로 받아 행동($a_t = D$)을 환경에 전달한다 (그림 5-1).

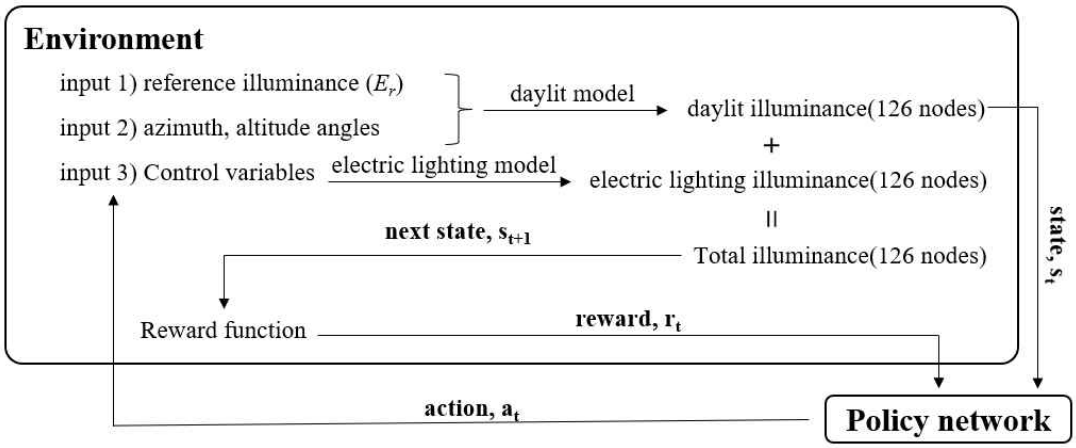


그림 5-1. 환경과 에이전트의 상호작용

5.1.2 보상함수

보상 함수는 최적화 알고리즘의 목적함수와 제약조건을 포괄하도록 작성하였다. 먼저, 표 (4-1)의 조명기기 그룹별 전력량(P)을 식 (2-3)에 대입하여 전력사용량을 평가하였으며, 이는 목적함수에 해당한다(식 (5-1)). 전력사용량은 낮을수록 유리하므로, 음수를 취한 다음, [0, 1] 사이의 값으로 정규화하였다(식 (5-2)의 소괄호).

$$SP_t = P^T \times D_t \tag{5-1}$$

여기서,

SP_t : 제어 시점 t의 전체 조명시스템 전력량.

$P = [1190, 1600, 1920, 1600, 1600, 1440, 1600, 1280, 1990, 780]^T$.

$D_t = [d_1, d_2, d_3, \cdots, d_{10}]^T$: 제어 시점 t의 조명기기 그룹별 디밍 수준.

$$elec\,power\,score = (1 - \frac{SP_t}{\sum_{i=1}^{10} P(i)}) \times 4 \tag{5-2}$$

또한, $E_{T,t}$를 이용하여 작업면 조도와 균제도를 평가하였으며, 이는 제약조건에 해당한다. 실제 작업면에 한하여 보상을 계산하기 위해, 작업면 그리드(6×21)에 따라 $E_{T,t}$를 재배열하고, 복도 영역을 배제하였다. 즉, 가장자리의 1열을 배제한 76개 지점(4×19)에 대해 계산하였다. 작업면 조도는 식

(5-3)과 같이 계산된다. 각 노드마다 450-550lx인 경우 100점을, 450lx 미만 혹은 550lx 이상인 경우는 0~100점을 획득하며, 획득한 점수의 총합을 [0, 1] 사이의 값으로 정규화하였다(그림 5-2). 균제도는, 작업면 그리드에서 가장 밝은 지점과 가장 어두운 지점의 비율을 의미한다. 두 지점의 조도가 같을 경우, 균제도는 최대값 '1'을 가지며, 모든 지점의 조도가 동일함을 의미한다(식 (5-4)).

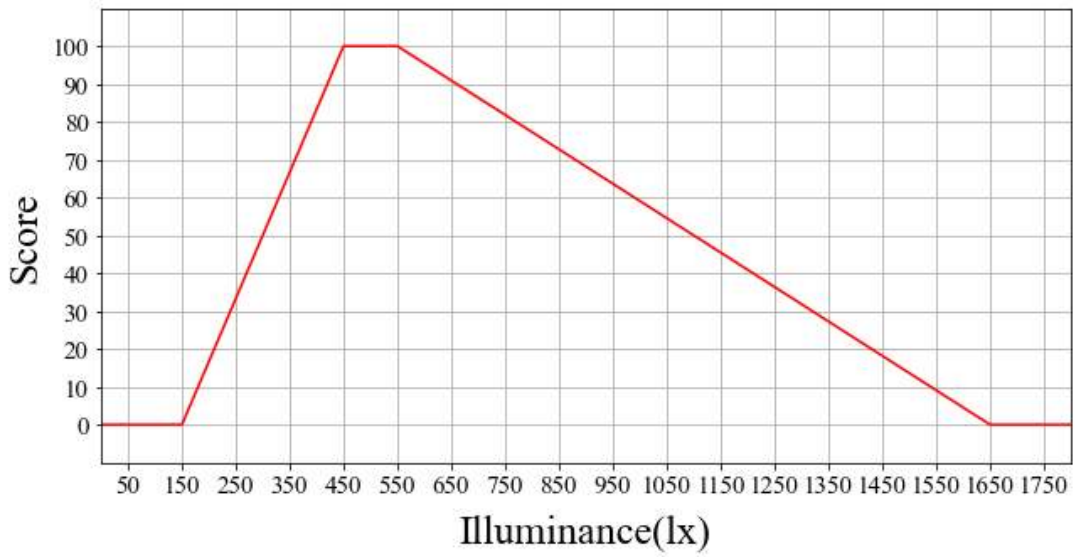


그림 5-2. 각 지점 조도에 따른 획득 점수

$$illuminance\ score = \frac{\sum_{i=2}^{5}\sum_{j=2}^{20} E_{T,t(i,j)}}{500 \times 76} \times 10 \qquad (5-3)$$

여기서, $E_{T,t_{(i,j)}} = (i,j)$번째 노드의 조도

$$uniformity\ score = \frac{\min(E_{T,t(i,j)})}{\max(E_{T,t(i,j)})} \times 2 \qquad (5-4)$$

식 (5-2), (5-3), (5-4)의 상수곱은 전체 보상에서 각 항목의 가중치를 의미한다. 이 상수는 제어 목적에 따라 변경할 수 있으며, 본 연구에서는 제약조건에 해당하는 항목에 더 높은 가중치를 부여하였다. 균제도 항목의 가중치(2)가 비교적 낮은 이유는 작업면 조도의 목표가 500lx로 모두 같기 때문이다. 즉, 작업면 조도 점수에 균제도가 일정부분 포함되었다고 할 수 있다. 각 항목을 모두 더하면, 제어 시점 t의 보상(r_t)이 된다. 각 시점, 보상의 최대값은 16, 최소값은 0이다. 이를 에피소드 1회(160개 제어 시점)에 대입하면, 보상의 총합인 반환은 [0, 2560] 사이의 값을 갖는다.

5.2 DDPG 훈련 및 검증

에피소드를 300회 반복하는 동안 DDPG 에이전트를 학습하였으며, 각 에피소드에서 받은 반환은 그림 5-3과 같다. 여기서 초기 5개 에피소드 까지는 리플레이 버퍼(Replay buffer)에 데이터를 수집하기 위한 것이다. 이때의 행동은 액터 네트워크를 이용하지 않고, 랜덤 변수 조합을 생성하여 행동으로 적용하였다. 리플레이 버퍼에 800개의 데이터 항목이 수집된 후. 제어 시점마다 256개 데이터 항목을 추출하여 액터 네트워크와 크리틱 네트워크에 대한 훈련을 수행하였다. 리플레이 버퍼에 수집된 데이터가 20,000항을 초과할 경우, 오래된 데이터 순으로 삭제하여 리플레이 버퍼 크기와 데이터의 질적 수준을 관리하였다. 에피소드를 140회 수행한 시점부터는 정책 네트워크의 성능이 개선되지 않고 반환값, 1,870 수준에 수렴하였다.

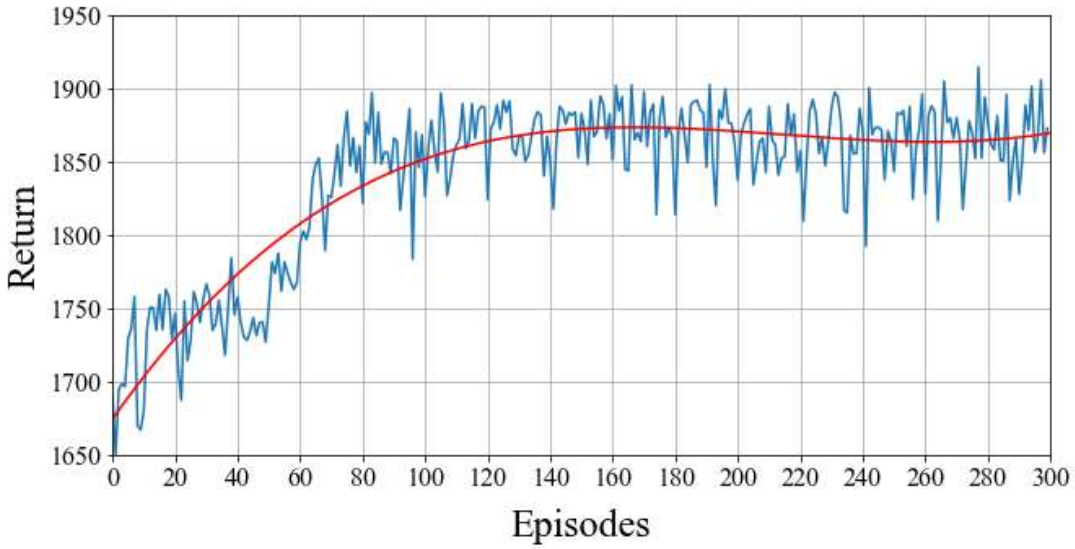


그림 5-3. DDPG 학습 곡선(에피소드별 반환)

안정적인 학습을 위해, 분리하였던 타겟 신경망이 충분히 업데이트될 수 있도록 에피소드는 수렴 지점을 초과하여 300회까지 모두 수행하였다. 학습된 정책 신경망을 에피소드에 적용하였을 때의 반환은 1,945이었다. 이때의 작업면 조도 환경은 평균 조도 478lx, 평균 균제도 1:1.66을 확보했으며, 평균 전력 사용량은 8,893W이다. 이는 전체 전력량 15,143W 대비, 약 59% 수준이다. 학습 기간에 포함되지 않은 3월 13일~16일을 검증 기간으로 설정한 다음, 정책 신경망으로 제어를 수행하였다. 이때의 반환은 1,993이었으며, 평균 조도 474lx 평균 균제도 1:1.60를 확보하였다. 평균 조명 전력량의 경우 7,931이며, 이는 전체 전력량 대비 52% 수준이다. 실제 제어를 수행하였을 때의 반환이 수렴 수준보다 높은 이유는 훈련과정에서 행동 변수에 더해지는 노이즈 때문이다. 훈련과정을 마치고 제어를 수행할 때는 더이상 환경을 탐색할 필요가 없다. 즉, 검증과정에서는 노이즈가 제거된 최적 정책에 따라 행동을 결정하므로, 훈련과정과 같은 에피소드라 할지라도 반환이 높아지는 것이다.

5.3 DDPG를 이용한 조명 시스템 제어 결과

5.3.1 DDPG 제어 결과

그림 5-4는 검증 기간 중 한 시점인 3월 13일 12시의 제어 결과를 보여준다. 이 시점의 작업면 자연채광은 240~470lx까지 분포한다. 건물의 형상(북측 천창, 3열)으로 인해, 북측의 조도가 낮고, 남측으로 갈수록 점차 높아진다(그림 5-4, (a)). 최적 정책으로 결정된 제어변수를 적용하여 확보된 전기조명 조도는 그림 5-4의 (b)와 같다. (제어변수는 그림 5-5의 (b)에서 확인할 수 있음) 자연채광 조도가 낮은 지점에 집중적으로 전기조명 조도가 공급된 것을 볼 수 있다. 자연채광과 전기조명 조도를 더한 작업면 총 조도(그림 5-4의 (c))를 살펴보면, 작업 지점 대부분이 권장 조도(500lx)를 확보하였으며 평균 조도는 495lx, 평균 균제도는 1:1.47이다. 전력사용량은 4,468W이다. 이를 근거로 정책 네트워크가 해당 시점에 실내 자연채광 상태에 따라 조명 시스템 제어를 합리적으로 수행하였다고 판단하였다.

그림 5-5 및 표 5-2~5를 통해, 제어변수의 시간별, 날짜별 변화를 살펴볼 수 있다. 조명기기 그룹 3의 경우, 모든 제어 시점에서 100%로 제어가 수행되었으며, 그룹 4 역시 모든 제어 시점에서 높은 디밍 수준으로 제어가 수행되었다. 그룹 3, 4는, 그림 5-4의 (a)에서 살펴본 바와 같이, 자연채광이 낮은 작업 지점을 담당한다. 다른 시점에서도 이러한 자연채광 특징은 유사하므로, 그룹 3의 '모든 시점, 100% 디밍'이라는 제어 결과는 타당하다고 판단하였다. 그룹 5~8의 구역은 자연채광이 비교적 풍부한 편이며, 시간에 따른 변화가 크다. 이 조명 그룹은 시간별 제어변수가 모두 다르며, 자연채광 변화에 따라 유동적으로 디밍 수준을 변경하였다. 한편, 그룹 1과 2 및 그룹 9와 10은 복도 구역과 작업 구역을 모두 담당하며, 많은 시점에서 낮은 수준으로 제어가 수행되었다. 그룹 1, 2는 그룹 3, 4를 보조한다는 점에서 같은 역할을 하지만, 그룹 1이 그룹 2보다 높은 수준으로 제어가 수행되었다. 이는 그룹 1에 소등된 전등이 많이 포함되어서 조도 확보를 위해 높은 수준의 디밍 제어가 필요하기 때문으로 판단된다 (표 5-2).

N

332	321	357	380	390	393	397	400	404	403	401	401	398	351	622	622	614	597	564	518	793
204	354	410	438	448	460	464	464	471	470	469	469	467	470	472	471	458	455	440	407	774
124	332	380	400	408	410	413	415	417	420	420	419	421	415	422	415	414	410	403	374	769
148	261	294	297	298	304	302	306	306	309	310	305	307	309	307	308	307	308	306	292	719
189	246	287	266	256	257	259	260	261	261	262	260	259	262	260	263	265	264	262	265	620
87	144	224	205	204	205	207	210	212	209	212	212	212	211	211	211	210	206	205	216	278

(a) 자연채광 조도

6	3	3	2	2	3	3	3	3	4	4	4	4	2	18	20	21	22	20	18	10
19	31	34	36	40	43	45	44	46	47	47	48	46	41	45	58	61	61	59	55	39
26	66	77	82	96	103	106	104	105	106	97	88	90	87	93	104	107	108	104	95	65
41	118	196	223	239	248	249	250	246	243	232	220	218	221	222	226	224	223	218	203	152
124	131	175	203	210	215	214	214	212	208	202	199	196	195	195	196	195	194	190	180	148
37	43	48	55	57	59	58	58	59	56	54	52	51	50	49	50	50	50	48	46	38

(b) 전기조명 조도

338	324	360	382	393	396	400	403	407	407	405	406	403	353	640	641	635	619	585	536	803
223	385	444	474	489	503	510	508	517	517	516	517	513	510	517	529	519	516	499	462	813
150	397	457	482	504	513	519	520	522	526	517	507	511	503	515	520	521	518	508	468	834
189	379	490	520	537	553	551	557	552	551	542	525	525	530	529	534	531	531	524	495	871
313	378	462	469	466	472	473	474	473	469	464	459	455	457	455	459	460	458	452	445	768
124	187	272	260	261	264	265	268	271	266	266	264	263	262	260	261	259	256	253	261	316

(c) 작업면 총 조도

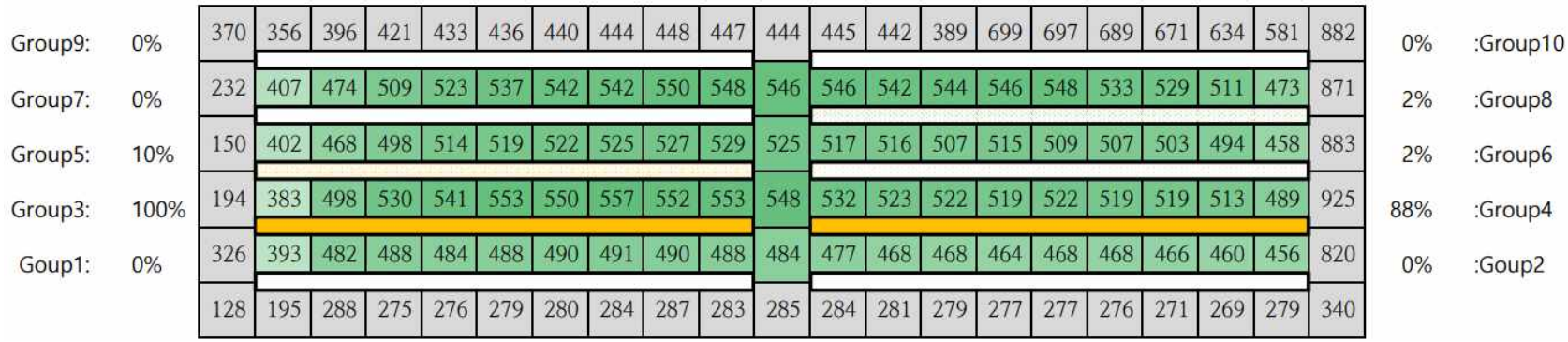
그림 5-4. 3월 13일 12시 광원별 작업면 조도

Group9:	1%	276	264	288	304	312	315	320	321	327	329	328	327	319	274	507	515	513	503	476	435	626	0%	:Group10
Group7:	49%	242	377	410	422	435	455	469	458	474	488	487	485	465	440	460	502	504	505	489	456	697	64%	:Group8
Group5:	45%	160	403	452	459	487	507	520	511	520	535	522	507	502	479	505	540	551	552	540	496	730	39%	:Group6
Group3:	100%	180	365	467	493	518	537	538	542	537	534	516	497	507	517	520	530	530	530	521	486	733	100%	:Group4
Goup1:	34%	319	383	455	456	454	467	458	459	459	439	421	415	416	420	419	424	426	424	418	408	638	0%	:Goup2
		154	212	275	257	255	265	255	256	261	238	225	219	218	218	217	218	217	215	212	217	256		

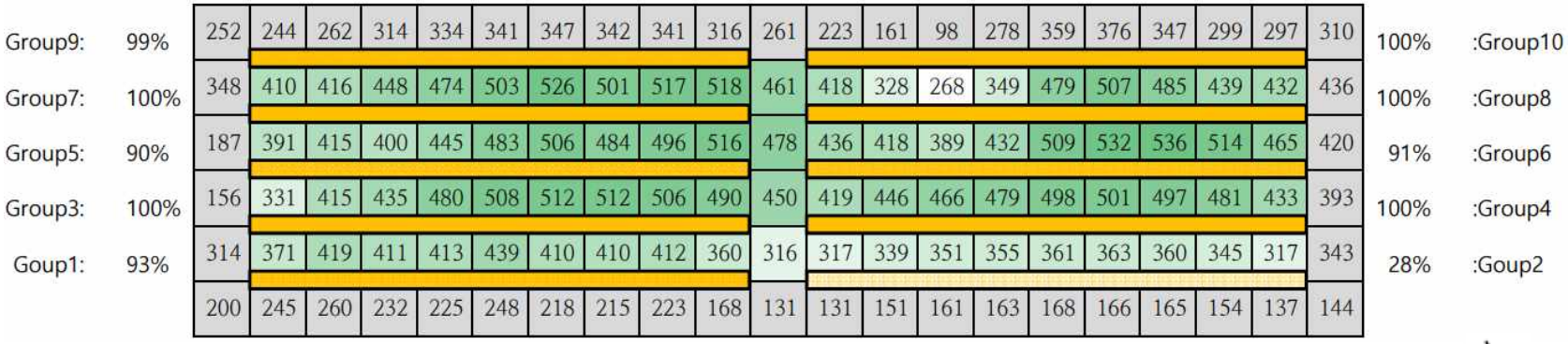
(a) 09시

Group9:	0%	338	324	360	382	393	396	400	403	407	407	405	406	403	353	640	641	635	619	585	536	803	0%	:Group10
Group7:	7%	223	385	444	474	489	503	510	508	517	517	516	517	513	510	517	529	519	516	499	462	813	17%	:Group8
Group5:	25%	150	397	457	482	504	513	519	520	522	526	517	507	511	503	515	520	521	518	508	468	834	17%	:Group6
Group3:	100%	189	379	490	520	537	553	551	557	552	551	542	525	525	530	529	534	531	531	524	495	871	93%	:Group4
Goup1:	4%	313	378	462	469	466	472	473	474	473	469	464	459	455	457	455	459	460	458	452	445	768	0%	:Goup2
		124	187	272	260	261	264	265	268	271	266	266	264	263	262	260	261	259	256	253	261	316		

(b) 12시



(c) 15시



(d) 18시

그림 5-5. 3월 13일, DDPG 최적 정책의 조명기기 그룹별 디밍 제어변수 및 작업면 조도

최적 정책이 일몰 이후(오후 6시 30분 이후), 자연채광이 없는 시점의 환경을 기반으로 결정한 제어변수 조합은 모든 조명 그룹을 100%로 제어하는 것이다. 전기조명 시스템만으로 확보된 환경의 평균 조도는 405lx, 균제도는 1.2.6이다. 즉, 자연채광 없이 대상 구역의 조명 시스템 출력만으로는 권장 조도 환경을 확보할 수 없었으며, 최대한 쾌적한 조도 환경 확보를 위해 제어변수를 결정하였음을 확인할 수 있다.

한편, 표 5-2~5를 서로 비교하면, 날짜별로 디밍 제어 패턴에 차이가 있다. 13일과 15일은 실내 자연채광이 비교적 풍부하여 낮은 수준의 디밍 제어가 수행되었다. 반대로 14일과 16일은 구름양이 적은 청천공인 날로, 합계 일사량은 많았지만, 북측 하늘의 일사량은 적어서 대상 건물의 실내 자연채광 조도가 낮았다(표 5-1). 이로 인해 표 5-3, 5-5와 같이 비교적 높은 수준의 디밍 제어가 필요했다.

표 5-1. 3월 13일-16일, 천공 상태

	Mar 13	Mar 14	Mar 15	Mar 16
Total cloud cover (1/10)	6.6	0.1	3.9	1
Total solar radiation quantity (MJ/m^2)	18.02	19.20	18.94	20.83

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
9:00	34%	0%	100%	100%	45%	39%	49%	64%	1%	0%
9:15	26%	0%	100%	100%	42%	36%	41%	56%	1%	0%
9:30	21%	0%	100%	99%	40%	33%	35%	50%	0%	0%
9:45	17%	0%	100%	98%	39%	31%	31%	45%	0%	0%
10:00	6%	0%	100%	94%	29%	23%	14%	23%	0%	0%
10:15	7%	0%	100%	95%	31%	25%	16%	27%	0%	0%
10:30	8%	0%	100%	96%	32%	26%	18%	29%	0%	0%
10:45	2%	0%	100%	92%	22%	12%	4%	12%	0%	0%
11:00	3%	0%	100%	93%	25%	16%	7%	16%	0%	0%
11:15	7%	0%	100%	95%	32%	25%	17%	27%	0%	0%
11:30	2%	0%	100%	92%	19%	9%	2%	9%	0%	0%
11:45	0%	0%	100%	81%	4%	0%	0%	0%	0%	0%
12:00	4%	0%	100%	93%	25%	17%	7%	17%	0%	0%
12:15	0%	0%	100%	86%	8%	2%	0%	1%	0%	0%
12:30	0%	0%	100%	79%	3%	0%	0%	0%	0%	0%
12:45	1%	0%	100%	91%	15%	6%	1%	6%	0%	0%
13:00	6%	0%	100%	94%	30%	24%	14%	24%	0%	0%
13:15	2%	0%	100%	92%	20%	11%	3%	11%	0%	0%
13:30	5%	0%	100%	94%	29%	23%	13%	23%	0%	0%
13:45	7%	0%	100%	95%	31%	24%	16%	26%	0%	0%
14:00	2%	0%	100%	92%	20%	10%	3%	10%	0%	0%
14:15	1%	0%	100%	90%	13%	4%	1%	4%	0%	0%
14:30	0%	0%	100%	86%	8%	2%	0%	2%	0%	0%
14:45	0%	0%	100%	82%	5%	1%	0%	1%	0%	0%
15:00	0%	0%	100%	88%	10%	2%	0%	2%	0%	0%
15:15	1%	0%	100%	91%	15%	6%	1%	6%	0%	0%
15:30	8%	0%	100%	95%	32%	25%	18%	29%	0%	0%
15:45	13%	0%	100%	97%	36%	28%	25%	38%	0%	0%
16:00	20%	0%	100%	99%	40%	33%	35%	49%	0%	0%
16:15	48%	0%	100%	100%	49%	45%	61%	74%	3%	0%
16:30	56%	0%	100%	100%	52%	48%	68%	79%	6%	0%
16:45	72%	2%	100%	100%	57%	55%	79%	87%	16%	1%
17:00	81%	6%	100%	100%	59%	56%	93%	95%	24%	15%
17:15	79%	15%	100%	100%	68%	66%	100%	100%	40%	41%
17:30	77%	21%	100%	100%	72%	73%	100%	100%	52%	56%
17:45	75%	28%	100%	100%	78%	82%	100%	100%	72%	80%
18:00	93%	28%	100%	100%	90%	91%	100%	100%	99%	100%
18:15	100%	95%	100%	100%	100%	93%	100%	100%	100%	100%
18:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

표 5-2. 3월 13일, 각 제어 시점별 강화학습 제어변수

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
9:00	81%	8%	100%	100%	62%	59%	99%	98%	29%	26%
9:15	82%	7%	100%	100%	60%	56%	97%	96%	25%	22%
9:30	80%	5%	100%	100%	59%	57%	90%	93%	24%	9%
9:45	77%	3%	100%	100%	59%	57%	83%	89%	22%	3%
10:00	71%	2%	100%	100%	57%	54%	79%	86%	15%	1%
10:15	63%	1%	100%	100%	54%	51%	73%	82%	9%	0%
10:30	53%	0%	100%	100%	51%	47%	65%	76%	5%	0%
10:45	45%	0%	100%	100%	49%	44%	59%	72%	3%	0%
11:00	38%	0%	100%	100%	46%	41%	53%	67%	2%	0%
11:15	32%	0%	100%	100%	44%	38%	47%	62%	1%	0%
11:30	28%	0%	100%	100%	43%	37%	43%	58%	1%	0%
11:45	23%	0%	100%	99%	41%	34%	37%	52%	1%	0%
12:00	19%	0%	100%	99%	39%	32%	32%	47%	0%	0%
12:15	19%	0%	100%	99%	39%	32%	33%	48%	0%	0%
12:30	19%	0%	100%	99%	39%	32%	32%	47%	0%	0%
12:45	18%	0%	100%	98%	39%	31%	31%	45%	0%	0%
13:00	11%	0%	100%	96%	34%	27%	21%	34%	0%	0%
13:15	17%	0%	100%	98%	38%	31%	30%	45%	0%	0%
13:30	17%	0%	100%	98%	38%	31%	30%	44%	0%	0%
13:45	20%	0%	100%	99%	40%	33%	34%	49%	0%	0%
14:00	17%	0%	100%	98%	38%	31%	30%	45%	0%	0%
14:15	10%	0%	100%	96%	34%	26%	20%	32%	0%	0%
14:30	12%	0%	100%	97%	36%	28%	24%	37%	0%	0%
14:45	7%	0%	100%	95%	31%	25%	16%	27%	0%	0%
15:00	9%	0%	100%	96%	33%	26%	19%	31%	0%	0%
15:15	11%	0%	100%	96%	34%	27%	21%	34%	0%	0%
15:30	15%	0%	100%	98%	37%	30%	27%	41%	0%	0%
15:45	43%	0%	100%	100%	48%	43%	57%	70%	3%	0%
16:00	26%	0%	100%	100%	42%	36%	41%	56%	1%	0%
16:15	34%	0%	100%	100%	45%	39%	49%	63%	1%	0%
16:30	63%	1%	100%	100%	54%	51%	73%	82%	9%	0%
16:45	67%	1%	100%	100%	55%	52%	76%	84%	11%	0%
17:00	80%	5%	100%	100%	59%	56%	91%	93%	24%	10%
17:15	79%	13%	100%	100%	66%	65%	100%	99%	37%	38%
17:30	76%	22%	100%	100%	74%	76%	100%	100%	57%	62%
17:45	75%	28%	100%	100%	79%	82%	100%	100%	72%	80%
18:00	93%	28%	100%	100%	90%	91%	100%	100%	99%	100%
18:15	100%	96%	100%	100%	100%	93%	100%	100%	100%	100%
18:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

표 5-3. 3월 14일, 각 제어 시점별 강화학습 제어변수

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
9:00	82%	7%	100%	100%	60%	56%	97%	96%	26%	22%
9:15	79%	4%	100%	100%	59%	57%	88%	92%	23%	7%
9:30	77%	3%	100%	100%	59%	57%	83%	89%	23%	3%
9:45	73%	2%	100%	100%	58%	55%	80%	87%	17%	1%
10:00	67%	1%	100%	100%	55%	52%	75%	84%	11%	0%
10:15	68%	1%	100%	100%	56%	53%	76%	84%	12%	0%
10:30	61%	1%	100%	100%	53%	50%	71%	81%	8%	0%
10:45	53%	0%	100%	100%	51%	47%	65%	77%	5%	0%
11:00	36%	0%	100%	100%	46%	40%	51%	65%	2%	0%
11:15	16%	0%	100%	98%	38%	31%	29%	43%	0%	0%
11:30	0%	0%	100%	89%	11%	3%	0%	3%	0%	0%
11:45	2%	0%	100%	92%	19%	10%	2%	9%	0%	0%
12:00	0%	0%	100%	85%	7%	1%	0%	1%	0%	0%
12:15	0%	0%	100%	77%	3%	0%	0%	0%	0%	0%
12:30	27%	0%	100%	100%	43%	36%	42%	57%	1%	0%
12:45	25%	0%	100%	99%	42%	35%	40%	55%	1%	0%
13:00	31%	0%	100%	100%	44%	38%	46%	61%	1%	0%
13:15	33%	0%	100%	100%	45%	39%	48%	63%	1%	0%
13:30	37%	0%	100%	100%	46%	40%	52%	66%	2%	0%
13:45	24%	0%	100%	99%	42%	35%	39%	54%	1%	0%
14:00	13%	0%	100%	97%	36%	29%	25%	38%	0%	0%
14:15	7%	0%	100%	95%	32%	25%	17%	28%	0%	0%
14:30	7%	0%	100%	95%	31%	24%	15%	25%	0%	0%
14:45	21%	0%	100%	99%	40%	33%	35%	50%	0%	0%
15:00	34%	0%	100%	100%	45%	39%	49%	64%	1%	0%
15:15	17%	0%	100%	98%	38%	31%	30%	45%	0%	0%
15:30	24%	0%	100%	99%	42%	35%	39%	54%	1%	0%
15:45	11%	0%	100%	96%	35%	27%	22%	34%	0%	0%
16:00	28%	0%	100%	100%	43%	37%	43%	58%	1%	0%
16:15	59%	1%	100%	100%	53%	49%	70%	80%	7%	0%
16:30	73%	2%	100%	100%	57%	55%	80%	87%	16%	1%
16:45	80%	5%	100%	100%	59%	57%	90%	93%	24%	9%
17:00	80%	9%	100%	100%	63%	61%	100%	99%	31%	29%
17:15	77%	3%	100%	100%	59%	57%	83%	89%	22%	3%
17:30	78%	16%	100%	100%	69%	67%	100%	100%	43%	44%
17:45	75%	25%	100%	100%	77%	80%	100%	100%	63%	70%
18:00	91%	28%	100%	100%	89%	90%	100%	100%	98%	99%
18:15	100%	91%	100%	100%	100%	93%	100%	100%	100%	100%
18:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

표 5-4. 3월 15일, 각 제어 시점별 강화학습 제어변수

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
9:00	81%	8%	100%	100%	62%	59%	99%	98%	29%	26%
9:15	82%	6%	100%	100%	60%	56%	96%	96%	25%	20%
9:30	80%	4%	100%	100%	59%	57%	89%	92%	23%	7%
9:45	77%	3%	100%	100%	59%	57%	82%	89%	21%	3%
10:00	70%	1%	100%	100%	56%	54%	77%	85%	13%	1%
10:15	66%	1%	100%	100%	55%	52%	75%	83%	10%	0%
10:30	60%	1%	100%	100%	53%	49%	70%	80%	7%	0%
10:45	53%	0%	100%	100%	51%	47%	65%	77%	5%	0%
11:00	48%	0%	100%	100%	50%	45%	62%	74%	4%	0%
11:15	44%	0%	100%	100%	48%	43%	58%	71%	3%	0%
11:30	40%	0%	100%	100%	47%	41%	54%	68%	2%	0%
11:45	36%	0%	100%	100%	46%	40%	51%	65%	2%	0%
12:00	32%	0%	100%	100%	45%	39%	47%	62%	1%	0%
12:15	31%	0%	100%	100%	44%	38%	46%	61%	1%	0%
12:30	31%	0%	100%	100%	44%	38%	46%	60%	1%	0%
12:45	31%	0%	100%	100%	44%	38%	46%	61%	1%	0%
13:00	31%	0%	100%	100%	44%	38%	46%	61%	1%	0%
13:15	32%	0%	100%	100%	44%	38%	47%	62%	1%	0%
13:30	35%	0%	100%	100%	45%	39%	49%	64%	1%	0%
13:45	37%	0%	100%	100%	46%	41%	52%	66%	2%	0%
14:00	39%	0%	100%	100%	47%	41%	54%	67%	2%	0%
14:15	42%	0%	100%	100%	48%	42%	56%	69%	2%	0%
14:30	45%	0%	100%	100%	49%	44%	59%	71%	3%	0%
14:45	49%	0%	100%	100%	50%	45%	62%	74%	4%	0%
15:00	52%	0%	100%	100%	51%	46%	64%	76%	4%	0%
15:15	53%	0%	100%	100%	51%	47%	65%	77%	5%	0%
15:30	58%	1%	100%	100%	53%	49%	69%	80%	7%	0%
15:45	64%	1%	100%	100%	55%	51%	74%	83%	10%	0%
16:00	70%	2%	100%	100%	56%	54%	78%	86%	14%	1%
16:15	76%	3%	100%	100%	59%	57%	82%	89%	21%	3%
16:30	80%	5%	100%	100%	59%	57%	90%	93%	24%	9%
16:45	81%	7%	100%	100%	60%	57%	98%	97%	27%	24%
17:00	80%	10%	100%	100%	64%	61%	100%	99%	32%	31%
17:15	79%	15%	100%	100%	68%	66%	100%	100%	41%	42%
17:30	77%	21%	100%	100%	73%	73%	100%	100%	52%	56%
17:45	76%	26%	100%	100%	77%	80%	100%	100%	65%	71%
18:00	87%	27%	100%	100%	86%	89%	100%	100%	94%	98%
18:15	100%	79%	100%	100%	100%	94%	100%	100%	100%	100%
18:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

표 5-5. 3월 16일, 각 제어 시점별 강화학습 제어변수

5.3.2 기존 제어 방법 및 DDPG 제어 성능 비교

DDPG 강화학습 알고리즘의 에너지 절감 성능을 평가하기 위해 현재 대상 건물에서 수행되고 있는 스케줄 기반 제어 및 센서의 피드백(closed loop)을 기반으로 한 선형계획법 최적화와 비교하였다(표 5-5). 평균 작업면 조도, 평균 균제도 및 전력사용량으로 비교하였으며, 에너지 절감률의 경우, 모든 제어 시점에 모든 조명기기 그룹을 100%로 제어하는 것을 기준으로 산정하였다.

먼저, 스케줄 제어는 점심시간(12시~1시, 1시간) 모든 전등을 소등하고 다른 작업 시간(09시~12시 & 13시~19시, 9시간)에는 모두 켜는 방법이다. 10시간 중 1시간 동안 전체 조명을 소등하였기에 에너지 절감률은 10%이다. 평균 균제도는 1:1.73으로 권장 수준(1:3)을 만족하였지만, 평균 조도는 656lx로 다소 높다.

표 5-6. 3월 13일-16일, 여러 가지 제어 방법의 성능 비교

	Schedule	Linear Programming	RL(DDPG)
Mean Illuminance (lx)	646	533	474
Mean Uniformity (min:max)	1:1.73	1:1.62	1:1.60
Mean electric energy (W)	13,629	10,488	7,931
Energy saving rate (%)	10%	31%	48%

자연채광을 고려한 전기조명 시스템 최적화를 수행하기 위해서는, Pandharipande, & Caicedo(2010)의 연구와 같이, 현재 시점의 제어변수를 적용한 전기조명 모델의 예측 결과와 작업 지점에 설치된 센서의 측정값이 차이를 이용하여 자연채광 조도를 파악해야 한다. 하지만, 대상 건물에는 작업면의 조도를 측정하는 센서가 설치되어 있지 않기 때문에, 4장에서 개발한 자연채광 모델의 시뮬레이션 결과를 자연채광 조도로 가정하였다. 이를 기반으로 선형계획법을 적용하여 최적화를 수행하였다. 적용한 제약조건은 작업 지점 (76지점 = 4×19, 복도 제외)에 한해, 450lx 이상을 확보하는 것이며 (제약조건의 기준 조도를 500lx로 할 경우, 이를 만족하는 해를 찾을 수 없었음), 목적함수는 전체 시스템의 전력사용량을 최소화하는 것이다. 제어 결과, 스케줄 제어와 비교하여, 균제도가 1:1.62로 개선되었으며, 평균 조도가 533lx로 권장 수준(500lx)에 가깝게 제어가 수행되었음을 알 수 있다. 전력사용량은 10,488W로 에너지 절감률은 31%이다.

강화학습(DDPG)의 제어 결과를 살펴보면, 균제도는 1.60이며 평균 조도는 474lx로 선형계획법과 비슷한 수준으로 제어가 수행되었다. 반면, 전력사용량은 세 가지 제어 방법 중 가장 낮았으며(7,931W), 선형계획법보다 17% 더 많은 에너지를 절감할 수 있었다.

그림 5-6은 그림 5-5와 같은 시점에 선형계획법으로 결정된 제어변수 및 작업면 조도이다. 두 그림을 비교하여 선형계획법과 강화학습 제어의 차이를 직관적으로 살펴볼 수 있다. 결정된 제어변수의 패턴은 두 방법이 비슷하다. 모든 제어 시점에서 그룹 3, 4를 높은 수준으로 디밍 제어하는 동시에, 기타 그룹은 실내 자연채광 상태에 따라 유동적으로 제어가 수행되었다. 하지만, 강화학습이 모든 제어 시점에서 그룹 3을 100%로 제어했던 것과 달리, 선형계획법은 이를 유동적으로 제어하는 동시에 그룹 1과 그룹 5를 더 높은 수준으로 제어하였다. 이를 통해 대부분의 제어 시점에서 제약조건(450lx 이상)을 모두 만족할 수 있었다(5-6의 (a)~(c)). 3월 13일 12시의 제어 결과를 DDPG 제어와 비교하면, 균제도는 1:1.44로 DDPG의 1:1.47보다 우수하며, 평균 조도 역시 528lx로 DDPG의 495lx보다 높다. 반변, 전력사용량은 6,025W로 DDPG의 4,468W보다 34% 많다.

두 방법의 제어 방법에 큰 차이가 있는 부분은 자연채광이 충분하지 않은 아침과 저녁 시간이다. 선형계획법의 제약조건이 각 작업 지점의 조도를 450lx 이상 확보하는 것이지만, 모든 제어 시점에서 이를 만족할 수는 없었다(그림 5-6의 (d)). 이는 최적 가능해가 존재하지 않거나, 탐색 실패로 인해 근사해를 도출하였기 때문이다. 근사해로 도출한 제어변수 조합 중에는, 제어변수 상한선인 '1'보다 큰 수가 포함되어 있었다. 각 제어변수가 1보다 클 경우, 1로 적용하였으며, 이를 전기조명 모델에 입력하여 작업면 총 조도를 구하였다. 표 5-7의 16:45분 이후 및 표 5-8의 10시 이전, 16시 30분 이후 등의 제어 시점이 여기에 해당하며, 다수의 조명 그룹이 100%로 제어가 수행되었다. 과도한 제어로 인해, 전력사용량 상승을 초래하였으며, 최적제어라고 판단하기 어렵다. 또한, 3월 13일, 16시 45분과 17시의 그룹 9, 10의 제어변수 변화를 살펴보면, 근사해로 결정된 제어변수의 경우, 일관성이 떨어지는 것도 확인할 수 있다. 제약조건을 500lx 이상으로 설정할 경우, 대부분의 제어 시점에서 가능해를 구할 수 없었으며, 이로 인해 적절한 제어변수를 결정할 수 없었다.

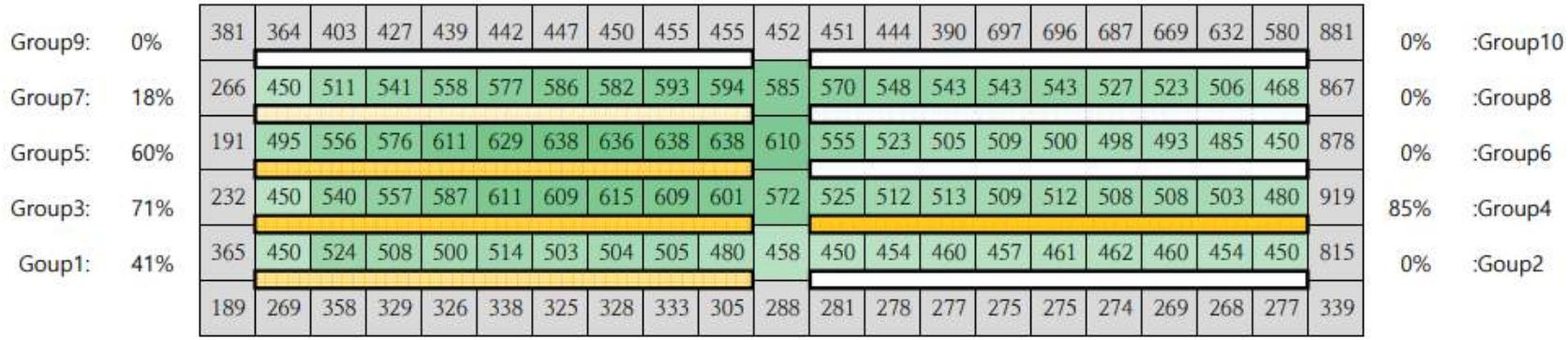
한편, 강화학습(DDPG)의 최적 정책 신경망은 모든 제어 시점에서 안정적으로 제어변수를 결정하였다. 또한, 시간의 흐름에 따라 연속적인 제어 패턴을 보인다(표 5-1~표 5-4). 앞서 비교한 바와 같이, 평균 조도 및 균제도는 선형계획법과 비교해 약간 낮은 수준이지만, 권장 조도 환경과 매우 유사하게 제어가 수행되었다. 전력사용량의 경우, 개별 제어 시점의 전력사용량과 평균 전력사용량 모두에서 강화학습의 성능이 여타 제어 방법보다 우수하다.

Group9:	0%	292	277	299	311	319	323	330	330	337	342	342	339	328	278	507	517	518	507	481	442	632	0%	:Group10
Group7:	84%	301	450	472	474	490	517	539	521	543	567	561	543	495	450	475	532	538	542	526	492	729	88%	:Group8
Group5:	93%	211	519	566	561	608	644	666	649	662	680	643	570	505	450	479	528	544	547	536	501	750	7%	:Group6
Group3:	99%	222	450	549	569	614	645	647	650	644	633	590	520	490	483	485	497	497	497	489	463	731	100%	:Group4
Goup1:	63%	371	450	521	513	510	531	513	515	515	480	450	451	475	489	491	497	498	495	476	450	681	42%	:Goup2
		203	272	333	306	301	319	299	300	307	268	245	253	280	294	295	298	296	291	277	262	300		

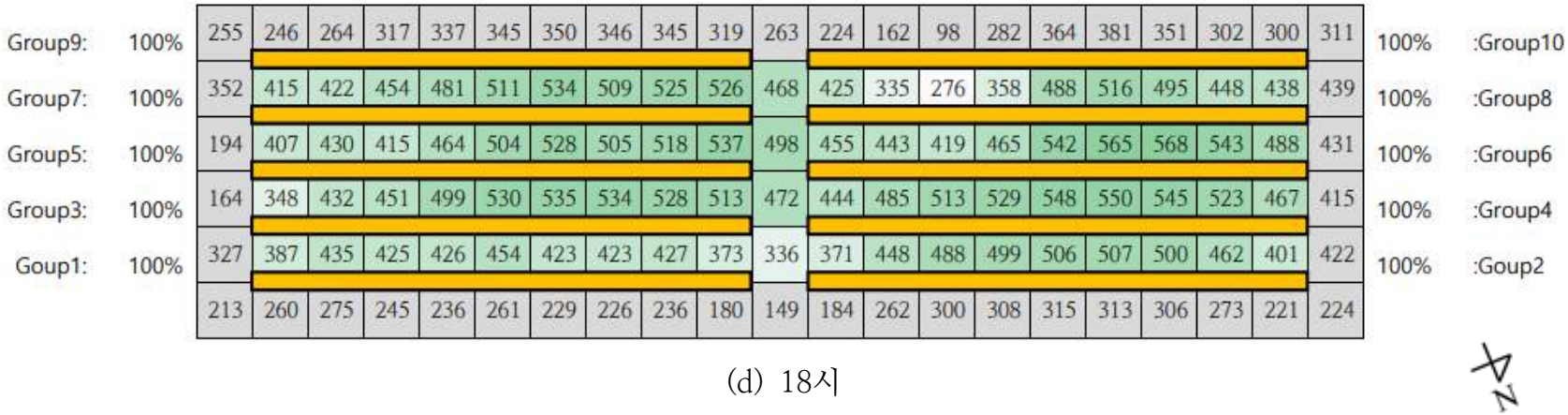
(a) 09시

Group9:	0%	354	337	371	392	402	406	411	414	419	421	418	416	407	355	637	638	632	616	582	534	803	0%	:Group10
Group7:	42%	277	450	499	520	537	559	572	563	578	585	577	558	526	510	515	527	516	514	497	461	814	23%	:Group8
Group5:	71%	197	501	558	571	610	633	645	639	644	650	618	555	510	481	489	493	494	492	483	450	828	0%	:Group6
Group3:	84%	228	450	545	564	598	624	623	628	622	613	580	528	512	510	508	512	509	510	504	481	869	100%	:Group4
Goup1:	48%	367	450	525	512	506	522	509	510	511	483	458	450	455	460	458	462	463	462	455	450	773	3%	:Goup2
		191	267	349	321	317	331	316	319	324	293	273	265	263	261	259	260	258	255	253	261	316		

(b) 12시



(c) 15시



(d) 18시

그림 5-6. 3월 13일, 선형계획법의 조명기기 그룹별 디밍 제어변수 및 작업면 조도

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
9:00	63%	42%	99%	100%	93%	7%	84%	88%	0%	0%
9:15	61%	38%	95%	100%	90%	9%	77%	61%	0%	0%
9:30	59%	33%	92%	100%	87%	6%	71%	50%	0%	0%
9:45	57%	29%	91%	100%	85%	3%	67%	46%	0%	0%
10:00	48%	3%	84%	100%	71%	0%	42%	23%	0%	0%
10:15	50%	8%	85%	100%	74%	0%	48%	29%	0%	0%
10:30	51%	12%	86%	100%	76%	0%	51%	32%	0%	0%
10:45	45%	0%	80%	96%	67%	0%	34%	16%	0%	0%
11:00	46%	0%	82%	98%	69%	0%	38%	19%	0%	0%
11:15	50%	9%	86%	100%	74%	0%	48%	29%	0%	0%
11:30	44%	0%	79%	94%	66%	0%	32%	13%	0%	0%
11:45	37%	0%	63%	74%	54%	0%	3%	0%	0%	0%
12:00	46%	0%	82%	99%	69%	0%	38%	19%	0%	0%
12:15	40%	0%	69%	82%	58%	0%	14%	0%	0%	0%
12:30	36%	0%	60%	70%	51%	0%	0%	0%	0%	0%
12:45	43%	0%	76%	91%	64%	0%	27%	9%	0%	0%
13:00	48%	4%	84%	100%	72%	0%	43%	24%	0%	0%
13:15	45%	0%	79%	95%	67%	0%	33%	14%	0%	0%
13:30	47%	2%	84%	100%	71%	0%	41%	22%	0%	0%
13:45	49%	7%	85%	100%	73%	0%	47%	28%	0%	0%
14:00	45%	0%	79%	95%	67%	0%	33%	14%	0%	0%
14:15	43%	0%	75%	89%	63%	0%	24%	6%	0%	0%
14:30	40%	0%	69%	83%	59%	0%	14%	0%	0%	0%
14:45	38%	0%	64%	75%	54%	0%	5%	0%	0%	0%
15:00	41%	0%	71%	85%	60%	0%	18%	0%	0%	0%
15:15	43%	0%	76%	91%	64%	0%	27%	9%	0%	0%
15:30	50%	11%	86%	100%	75%	0%	50%	31%	0%	0%
15:45	54%	22%	89%	100%	81%	0%	60%	41%	0%	0%
16:00	59%	32%	92%	100%	87%	6%	71%	50%	0%	0%
16:15	67%	65%	100%	100%	100%	45%	94%	100%	0%	0%
16:30	84%	79%	100%	100%	100%	78%	100%	100%	1%	0%
16:45	100%	100%	100%	100%	100%	100%	100%	100%	97%	36%
17:00	100%	100%	100%	100%	100%	100%	100%	100%	38%	100%
17:15	77%	100%	100%	100%	100%	94%	95%	100%	100%	100%
17:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
17:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:00	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:15	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

표 5-7. 3월 13일, 각 제어 시점별 선형계획법 제어변수

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
9:00	99%	100%	100%	100%	100%	100%	100%	100%	100%	73%
9:15	99%	100%	100%	100%	100%	100%	100%	100%	94%	100%
9:30	100%	100%	100%	100%	96%	100%	100%	100%	100%	100%
9:45	100%	100%	100%	100%	100%	100%	100%	100%	93%	100%
10:00	100%	100%	100%	100%	100%	100%	100%	100%	98%	30%
10:15	100%	95%	100%	100%	100%	100%	100%	100%	8%	9%
10:30	76%	73%	100%	100%	100%	64%	98%	100%	0%	0%
10:45	66%	57%	100%	100%	100%	34%	92%	100%	0%	0%
11:00	63%	44%	100%	100%	94%	6%	87%	100%	0%	0%
11:15	62%	41%	98%	100%	92%	7%	82%	81%	0%	0%
11:30	61%	39%	96%	100%	91%	8%	79%	68%	0%	0%
11:45	60%	35%	94%	100%	88%	8%	74%	53%	0%	0%
12:00	58%	30%	92%	100%	86%	4%	69%	48%	0%	0%
12:15	58%	31%	92%	100%	86%	5%	70%	49%	0%	0%
12:30	58%	30%	92%	100%	86%	4%	69%	48%	0%	0%
12:45	57%	29%	91%	100%	85%	4%	67%	46%	0%	0%
13:00	53%	17%	88%	100%	78%	0%	56%	37%	0%	0%
13:15	57%	29%	91%	100%	85%	3%	67%	46%	0%	0%
13:30	57%	28%	91%	100%	84%	3%	66%	46%	0%	0%
13:45	58%	32%	92%	100%	87%	5%	70%	49%	0%	0%
14:00	57%	28%	91%	100%	85%	3%	67%	46%	0%	0%
14:15	52%	15%	87%	100%	77%	0%	54%	35%	0%	0%
14:30	54%	21%	89%	100%	80%	0%	59%	40%	0%	0%
14:45	50%	8%	85%	100%	74%	0%	48%	29%	0%	0%
15:00	51%	13%	87%	100%	76%	0%	52%	33%	0%	0%
15:15	53%	17%	88%	100%	79%	0%	56%	37%	0%	0%
15:30	56%	25%	90%	100%	83%	0%	63%	45%	0%	0%
15:45	65%	52%	100%	100%	100%	27%	91%	100%	0%	0%
16:00	60%	37%	95%	100%	90%	9%	77%	59%	0%	0%
16:15	62%	42%	98%	100%	93%	7%	84%	87%	0%	0%
16:30	99%	94%	100%	100%	100%	100%	100%	100%	8%	8%
16:45	100%	100%	100%	100%	100%	100%	100%	100%	69%	9%
17:00	100%	100%	100%	100%	96%	100%	100%	100%	100%	100%
17:15	73%	100%	100%	100%	100%	100%	100%	100%	100%	100%
17:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	93%
17:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:00	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:15	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

표 5-8. 3월 14일, 각 제어 시점별 선형계획법 제어변수

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
9:00	92%	99%	100%	100%	100%	100%	100%	100%	94%	100%
9:15	100%	100%	100%	100%	100%	100%	100%	100%	100%	93%
9:30	100%	100%	100%	100%	100%	100%	100%	100%	90%	78%
9:45	100%	100%	100%	100%	100%	100%	100%	100%	96%	47%
10:00	100%	100%	100%	100%	100%	100%	100%	100%	63%	9%
10:15	100%	100%	100%	100%	100%	100%	100%	100%	82%	7%
10:30	93%	88%	100%	100%	100%	96%	100%	100%	6%	0%
10:45	76%	74%	100%	100%	100%	64%	98%	100%	0%	0%
11:00	63%	43%	99%	100%	94%	6%	86%	95%	0%	0%
11:15	57%	27%	91%	100%	84%	2%	65%	45%	0%	0%
11:30	42%	0%	73%	87%	61%	0%	21%	3%	0%	0%
11:45	45%	0%	79%	95%	66%	0%	32%	13%	0%	0%
12:00	40%	0%	68%	81%	57%	0%	12%	0%	0%	0%
12:15	35%	0%	58%	67%	49%	0%	0%	0%	0%	0%
12:30	61%	38%	96%	100%	90%	8%	78%	63%	0%	0%
12:45	60%	37%	95%	100%	89%	9%	76%	56%	0%	0%
13:00	62%	40%	97%	100%	92%	8%	81%	77%	0%	0%
13:15	62%	42%	98%	100%	93%	7%	83%	84%	0%	0%
13:30	63%	43%	100%	100%	94%	6%	86%	96%	0%	0%
13:45	60%	36%	94%	100%	89%	9%	75%	55%	0%	0%
14:00	55%	22%	89%	100%	81%	0%	61%	42%	0%	0%
14:15	50%	9%	86%	100%	74%	0%	48%	29%	0%	0%
14:30	49%	6%	85%	100%	73%	0%	45%	26%	0%	0%
14:45	59%	33%	92%	100%	87%	6%	71%	50%	0%	0%
15:00	62%	42%	99%	100%	93%	7%	84%	88%	0%	0%
15:15	57%	29%	91%	100%	85%	3%	67%	46%	0%	0%
15:30	60%	36%	94%	100%	89%	9%	75%	54%	0%	0%
15:45	53%	18%	88%	100%	79%	0%	56%	38%	0%	0%
16:00	61%	39%	96%	100%	91%	8%	79%	67%	0%	0%
16:15	90%	85%	100%	100%	100%	90%	100%	100%	4%	0%
16:30	100%	100%	100%	100%	100%	100%	100%	100%	99%	39%
16:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
17:00	76%	84%	100%	100%	100%	100%	100%	100%	100%	40%
17:15	100%	100%	100%	100%	100%	100%	100%	100%	91%	100%
17:30	100%	100%	100%	100%	100%	88%	100%	100%	100%	100%
17:45	100%	60%	100%	100%	100%	100%	100%	100%	100%	100%
18:00	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:15	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

표 5-9. 3월 15일, 각 제어 시점별 선형계획법 제어변수

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
9:00	100%	100%	100%	100%	100%	100%	100%	100%	100%	74%
9:15	100%	85%	100%	100%	100%	100%	100%	100%	100%	100%
9:30	100%	100%	100%	100%	100%	100%	100%	100%	69%	100%
9:45	100%	100%	100%	100%	100%	100%	100%	100%	84%	92%
10:00	100%	100%	100%	100%	100%	100%	100%	100%	100%	94%
10:15	100%	100%	100%	100%	100%	100%	100%	100%	36%	12%
10:30	91%	86%	100%	100%	100%	92%	100%	100%	5%	0%
10:45	77%	74%	100%	100%	100%	65%	98%	100%	0%	0%
11:00	68%	66%	100%	100%	100%	47%	95%	100%	0%	0%
11:15	65%	53%	100%	100%	100%	28%	91%	100%	0%	0%
11:30	64%	44%	100%	100%	98%	12%	88%	100%	0%	0%
11:45	63%	43%	99%	100%	94%	6%	86%	94%	0%	0%
12:00	62%	41%	98%	100%	92%	7%	82%	82%	0%	0%
12:15	62%	40%	97%	100%	92%	8%	81%	77%	0%	0%
12:30	62%	40%	97%	100%	92%	8%	81%	77%	0%	0%
12:45	62%	40%	97%	100%	92%	8%	81%	77%	0%	0%
13:00	62%	41%	97%	100%	92%	7%	82%	79%	0%	0%
13:15	62%	41%	98%	100%	92%	7%	82%	80%	0%	0%
13:30	63%	42%	99%	100%	93%	7%	84%	89%	0%	0%
13:45	63%	44%	100%	100%	94%	6%	86%	97%	0%	0%
14:00	63%	44%	100%	100%	97%	10%	88%	100%	0%	0%
14:15	64%	48%	100%	100%	100%	21%	89%	100%	0%	0%
14:30	66%	57%	100%	100%	100%	33%	92%	100%	0%	0%
14:45	69%	68%	100%	100%	100%	50%	95%	100%	0%	0%
15:00	74%	72%	100%	100%	100%	59%	97%	100%	0%	0%
15:15	77%	75%	100%	100%	100%	66%	99%	100%	0%	0%
15:30	88%	83%	100%	100%	100%	86%	100%	100%	3%	0%
15:45	100%	99%	100%	100%	100%	100%	100%	100%	14%	13%
16:00	100%	100%	100%	100%	100%	100%	100%	100%	100%	98%
16:15	100%	100%	100%	100%	100%	100%	100%	100%	82%	92%
16:30	100%	100%	100%	100%	97%	100%	100%	100%	100%	100%
16:45	93%	100%	100%	100%	100%	100%	100%	100%	100%	100%
17:00	100%	100%	100%	100%	100%	100%	100%	100%	87%	88%
17:15	100%	100%	100%	100%	100%	94%	100%	100%	100%	100%
17:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
17:45	100%	64%	100%	100%	100%	100%	100%	100%	100%	100%
18:00	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:15	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:30	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
18:45	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

표 5-10. 3월 16일, 각 제어 시점별 선형계획법 제어변수

제 6 장 결 론

실내 자연채광 현황에 맞추어 전기조명을 제어하는 방법을 제시한 많은 연구는 조도 센서의 피드백을 기반으로 한다. 조도 센서 기반의 조명 제어 방법은 물리적 한계로 인해, 사람의 행동과 센서 사이의 간섭, 다수의 조명 기기가 다수의 센서에 동시에 영향을 미치는 간섭 문제를 피할 수 없다. 이에 따라 '조도 센서의 설치 수량을 줄이고 최적 설치 위치를 찾는 방법', '작업면과 조도 센서의 설치 위치를 보정하는 방법', 및 '전체 작업 지점이 안정화 될 때까지 조명기기를 순차적으로 조절하여 최적화하는 방법' 등에 관한 연구를 수행해 왔다. 하지만, 조도 센서는 기본적으로 설치 및 지속적인 유지보수 비용을 유발한다. 정확한 작업 지점이 아닌 대체 지점에 설치된 조도 센서는 물리적인 한계를 완전히 극복할 수 없으며, 측정값의 정확도 하락으로 인해, 적절한 제어를 보장할 수 없다. 또한, 순차적 제어를 통해 조명 기기 사이의 간섭을 해결하는 최적화 과정은 실시간으로 변하는 자연채광 현황에 대응하기 어렵다.

한편, 자연채광 영향이 적거나 없는 실내 공간에서는 선형시스템 모델을 이용하여 전체 조명기구의 제어변수를 한 번에 결정하는 방법을 통해 센서 없는 조명 제어(Sensor free lighting control) 연구가 수행되었다. 모델을 통해 최적화 알고리즘을 적용하여 제어변수를 결정할 수 있게 된 것이다. 모델은 물리법칙에 따른 White box 모델 및 작업 지점에서 측정된 데이터를 수집하여 작성한 Black box 모델 등 다양한 형태로 제작되었다. 하지만, 조명기기 제어에 적용된 최적화 알고리즘은 대부분 선형계획법에 한정되어 있다. 선형계획법은 제약조건에 의해 해를 찾지 못할 경우, 안정성이 떨어지고, 성능이 보장되지 않는다. 제어 시점마다 최적화를 수행해야 하며, 실시간으로 제어변수를 결정하지 못할 수 있다.

전기조명 모델을 이용한 제어 최적화 방법을 자연채광이 풍부한 실내환경에 적용할 때는, 다시 작업면에 조도 센서를 설치하였다. 측정값과 전기조명 모델의 결과를 비교하여 그 차이만큼을 자연채광 조도로 간주한 것이

다. 자연채광으로 인한 작업면 조도를 실시간으로 예측하는 모델을 활용하여 전기조명 제어를 수행한 연구는 많지 않다. 대부분의 기상 데이터 기반 자연채광 모델링(Climate Based Daylight Modelling, CBDM) 연구는 연간 실내 자연채광의 수준을 평가하는 방법 및 이를 적용하여 에너지 절감량을 산출하기 위해 수행되었다(Brembilla, & Mardaljevic, 2019). 자연채광 조도 시뮬레이션을 위해서는 해당 시점의 법선면 직달 일사량(Direct Normal Irradiance, DNI)과 수평면 산란 일사량(Diffused Horrizontal Irradiance, DHI)이 필요하다. 기상청에서는 DNI와 DHI의 합인 수평면 총 일사량(Global Horrizontal Irradiance, GHI)만 제공하며 그마저도 실시간으로 이용하는 데 한계가 있다.

본 연구에서는 자연채광과 전기조명 조도를 예측하는 White box 모델을 각각 개발하고, 이를 실시간으로 이용할 수 있도록 확장하였다. 또한, 제작된 실시간 조도예측모델을 강화학습에 적용하여 조명 시스템을 제어하는 방법을 제안하였다. 대상 건물은 인천에 있는 공장 건물로, 북측 천장에 채광창이 설치되어 있다. 전기조명은 형광등 조명기구와 LED등이 함께 설치되어 있으며, 10개 그룹 나누어 제어된다.

자연채광 모델의 경우, 먼저 기상청에서 수집된 일사량을 입력으로 받는 물리법칙 기반의 모델을 제작했다. 모델은 Radiance 기반의 빛 환경 시뮬레이션 도구인 DIVA for Grasshopper로 제작되었다. 다음으로, 실시간성 확보를 위해 자연채광 조도만을 측정하는 센서를 설치하여 데이터를 수집했다. 앞서 제작한 모델에, 센서의 측정값과 같은 시점의 기상청 일사량으로 시뮬레이션을 수행하고, 입력변수(GHI)를 센서의 측정값으로 변경하여 기계학습 대리모델(Black box)로 확장하였다. 한편, 전기조명의 경우, 자연채광 모델과 같은 도구를 이용하여 조명기구 그룹별 시뮬레이션을 수행하였다. 그 결과로 디밍 제어변수에 대한 선형시스템 모델을 제작하였다. 작업면에 임시로 조도 센서를 설치하여 측정한 데이터로 모델의 미지변수를 보정하고, 모델의 정확도를 검증했다.

제작된 모델을 기반으로 강화학습을 적용하였다. 강화학습은 인공지능의 한 분야로 사람의 학습 방법과 흡사하다. 주어진 환경에서 제어변수를 결정

하고, 그 결과에 따라 보상을 받는다. 훈련 기간, 다양한 환경 조건에 대해, 다양한 제어변수를 적용하여 학습 데이터를 생성한다. 이를 이용하여 가장 많은 보상을 받을 수 있는 제어변수를 탐색하고, 결과로는 가장 높은 보상을 결정하는 방법인 최적 정책을 도출한다.

본 연구에서 적용한 심층 확정적 정책 그래디언트(DDPG) 강화학습 알고리즘은 제어변수를 확정적으로 결정하는 정책을 인공신경망(ANN)으로 표현하고, 그 파라미터를 변경하면서 최적 정책을 탐색하는 알고리즘이다. DDPG는 탐색의 자유도가 높고, 디밍 제어와 같은 연속 제어변수를 결정할 수 있다. 제어변수의 수가 많을수록 그 성능이 다른 정책 그래디언트 기반 알고리즘보다 우수함이 알려져있다. 훈련된 최적 정책 네트워크(Optimal Policy Network)로 제어를 수행한 결과, 평균 조도는 474lx, 균제도는 1:1.6으로 확보하였으며 이는 대상 건물의 권장수준을 만족하는 수준이다. 제어 기간에 사용한 전기에너지는 7,931W로 기존의 스케줄 제어보다 38% 더 많이 절감할 수 있었다.

강화학습의 제어 성능을 검증하고자 대상 건물에 선형계획법 최적화를 구현하였다. 두 방법 모두 자연채광이 부족한 부분에 집중하여 전기조명 제어가 수행되었으며, 권장 조도 환경을 만족하였다. 하지만 에너지 절감량의 경우, 강화학습이 선형계획법보다 17% 더 많이 절감할 수 있었다. 선형계획법은, 실내 자연채광이 부족하여 조명 시스템 제어를 통해 제약조건을 만족할 수 없는 경우, 제어변수 결정에 불안정성을 보였으며, 각 지점의 조도를 최대한 확보하기 위해 보수적으로 제어가 수행되었다. 한편, 강화학습은 시간에 따른 자연채광 변화에 따라 유연하고 안정적인 제어를 수행하였으며, 최적화 성능이 유지되었다.

본 연구는 작업면에 센서를 설치하지 않고도 실시간 전기조명 제어가 가능함을 보였으며, 강화학습의 제어 성능 및 안정성을 확인하였다. 추후 연구에서는 자연채광 모델 개선 및 확장된 대리모델의 보완이 필요하다. 제작된 모델은 1월~3월 데이터를 기반으로 보정 및 검증되었기에 다른 계절의 정확도는 하락할 수 있다. 또한, 대리모델 제작 과정에서 다양한 자연채광 및 천공 상태 데이터가 부족하여 예측 정확도가 하락하였으며, 데이터 추가

수집이 필요하다.

한편, 강화학습 제작 과정에서는, 연구자의 경험에 의존해야 하는 한계가 있다. 강화학습 파라미터 설정 및 보상함수 제작은 성능에 학습 및 제어 성능에 큰 영향을 미치는 요소이다. 하지만, 이를 정량적으로 결정할 방법이나 지침이 없다. 본 연구에서는 수많은 시도와 오류수정(Try and Error)을 통해 연구자의 직관으로 판단하였으며 이에 관한 후속 연구가 필요하다. 또한, 강화학습 제어 결과를 분석할 때, 결정된 제어변수와 작업면의 자연채광 현황을 비교하여 인과성을 검토하였다. 추후, 각 제어 시점별로 획득한 보상을 항목별로 구분하면, 강화학습이 결정한 디밍 수준에 대한 세부적인 분석이 가능할 것이다.

참 고 문 헌

1. Afram, A., Janabi-Sharifi, F., Fung, A. S., & Raahemifar, K. (2017). 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, 141, 96-113.
2. American Society of Heating, Ventilating, and Air Conditioning Engineers (ASHRAE). (2002). Guideline 14-2002 Measurement of Energy and Demand Savings, Technical Report. American Society of Heating, Ventilating, and Air Conditioning Engineers. Atlanta, GA, USA
3. Ahn, K.U. and Park, C.S. (2020), Application of deep Q-networks for model-free optimal control balancing between different HVAC systems, ASHRAE Science and Technology for the Built Environment, 26(1), 61-74.
4. Brembilla, E., & Mardaljevic, J. (2019). Climate-Based Daylight Modelling for compliance verification: Benchmarking multiple state-of-the-art methods. Building and Environment, 158, 151-164.
5. Caicedo, D., Pandharipande, A., & Willems, F. M. J. (2014). Daylight-adaptive lighting control using light sensor calibration prior-information. Energy and Buildings, 73, 105-114.
6. Bishop, C. (2006). Pattern Recognition and Machine Learning. New York. Springer-Verlag New York. 225-256
7. Di Laura, D. L., Houser, K. W., Mistrick, R. G., & Steffy, G. R. (2011). The lighting handbook: reference and application, New York, Illuminating Engineering Society of North America.
8. Di Louie, C. (2008). Light control handbook. New York. Illuminating Engineering Society of North America. 161-204.

9. Haq, M., Hassan, M., Abdullah, H., Rahman, H., Abdullah, M., Hussin, F., & Said, D. (2014). A review on lighting control technologies in commercial buildings. Renewable and Sustainable Energy Reviews, 33, 268－279.
10. Jakubiec, J A., & Reinhart, C. F. (2011). DIVA 2.0: Integrating daylight and thermal simulations using Rhinoceros 3D, Daysim and EnergyPlus, Proceedings of building simulation, 20(11), 2202-2209.
11. Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., Behar, J. V., Hern, S. C., & Engelmann, W. H. (2001). The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. Journal of Exposure Science & Environmental Epidemiology, 11(3), 231-252.
12. Kim, Y. J., and Park, C. S. (2016). Stepwise deterministic and stochastic calibration of an energy simulation model for an existing building. Energy and Buildings, 133, 455-468.
13. Koroglu, M. T., & Passino., K. M. (2014). Illumination Balancing Algorithm for Smart Lights. IEEE Transactions on Control Systems Technology, 22(2), 557-567.
14. Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D. (2015). Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971.
15. Mahdavi, A., Mathew, P., Kumar, S., Hartkopf, V., & Loftness, V. (1995). Effects of lighting, zoning, and control strategies on energy use in commercial buildings. Journal of the Illuminating Engineering Society, 24, 25-35.
16. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I.,

King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. nature, 518(7540), 529-533.

17. Ochoa, C. E., Aries, M. B., & Hensen, J. L. (2010). Current perspectives on lighting simulation for building science. IBPSA-NVL 2010 Event.
18. Pandharipande, A., & Caicedo, D. (2011). Daylight integrated illumination control of LED systems based on enhanced presence sensing. Energy and Buildings, 43(4), 944-950.
19. Perez, R., Seals, R., & Michalsky, J. (1993). All-weather model for sky luminance distribution—Preliminary configuration and validation. Solar Energy, 50(3), 235-245.
20. Seyedolhosseini, A., Masoumi, N., Modarressi, M., & Karimian, N. (2020). Daylight adaptive smart indoor lighting control method using artificial neural networks. Journal of Building Engineering, 29, 101141.
21. Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., & Riedmiller, M. (2014, June). Deterministic policy gradient algorithms.
22. Solemma LLC. (2020). DIVA. Retrieved May 11, 2020, from https://solemma.com/Diva.html
23. Sutton, R. S., & Barto, A. G. (1998). Introduction to reinforcement learning (Vol. 135). Cambridge: MIT press.
24. Sutton, R. S., McAllester, D. A., Singh, S. P., & Mansour, Y. (1999). Policy gradient methods for reinforcement learning with function approximation. In Neural Information Processing Systems, 12, 1057-1063.
25. Tran, D., & Tan, Y. K. (2014). Sensorless illumination control of a networked LED-lighting system using feedforward neural network.

IEEE transactions on industrial electronics, 61(4), 2113-2121
26. Ward, G., & Shakespeare, R. (1998). Rendering with RADIANCE, The Art and Science of Lighting Visualization. San Francisco, CA, USA. Morgan Kaufmann Publishers.
27. Watanabe, T., Urano, Y., & Hayashi, T. (1983). Procedures for separating direct and difuse insolation on a horizontal surface and prediction of insolation on tilted surfaces. Transactions of the Architectural Institute of Japan, 330, 96-108.
28. Wen, Y. J., & Agogino, A. M. (2008). Wireless networked lighting systems for optimizing energy savings and user satisfaction. 2008 IEEE Wireless Hive Networks Conference, Austin, TX, 1-7.
29. Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine learning, 8(3-4), 229-256.
30. 김재민, 김영섭, 신한솔, 박철수. (2019). 인공신경망을 이용한 자연채광 조도예측모델 개발. 대한건축학회 추계학술발표대회논문집, 39(2), 581-582.
31. 김영섭, 김재민, 신한솔, 박철수 (2019), 조도 예측 모델 기반의 조명 최적 제어, 대한건축학회 추계학술발표대회논문집, 39(2), 537-538.
32. 김영섭, 김재민, 신한솔, 박철수 (2020), 시뮬레이션 모델을 이용한 공장 건물의 조명 최적 제어, 대한건축학회 논문집 36(8), 127-135.
33. 기상청 국가기후데이터센터. (2020). 종관기상관측, 인천기상대, https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36
34. 박성호 (2019) 하이브리드 기계학습 모델 기반 건물 설비 시스템 최적 제어. 성균관대학교 일반대학원, 서울.
35. 박철수. (2006). 규범적 건물성능평가방법. 대한건축학회논문집 22(11), 337-344.
36. 안기언. (2019) Deep Q-learning의 강화학습 기반 건물 에너지 시스템의 model-free 최적제어. 성균관대학교 일반대학원, 서울.

37. 지영민, 옥기수, 권동우. (2019). IoT 기반 지능형 건물을 위한 환경 모니터링 시스템. 한국정보과학회 학술발표논문집, 345-346.
38. 온실가스종합정보센터. (2019). 2019 국가 온실가스 인벤토리 보고서.
39. 이하영, 윤성환, 박철수. (2015). 직달 산란 분리모델이 건물 에너지 예측에 미치는 영향. 대한건축학회 논문집 31(11), 221-229.
40. 한국에너지공단. (2020). 2019년도 에너지사용량 통계(에너지사용량 신고업체).

Abstract

Optimal Control of Lighting System based on Illuminance Prediction Model with Deep Deterministic Policy Gradient(DDPG)

Youngsub Kim
Department of Architecture and Architectural Engineering
The Graduate School
Seoul National University

In indoor space, an appropriate level of illumination should be secured for the efficiency and safety of the human activity. If daylight is introduced through the skylights, lighting energy can be saved by supplying only as much illuminance as needed with electric lighting. Daylight varies dynamically over time depending on the position of the sun and sky conditions. In order to determine the dimming control variables, the dynamic behavior of workplane illuminance by daylighting should be independently identified at each point of control. In addition, to meet a specific amount of electric lighting illuminance, the luminous flux from different lighting fixtures should be known, depending on the

control variable. With a system model, the control optimization can be conducted that satisfy the recommended lighting environment while minimizing electric power.

The conventional closed-loop lighting control demands illuminance sensors mounted on a horizontal workplane. However, if the control target space is large as well as shaded by complex indoor surroundings, the corresponding number of photo sensors would increase. In addition, workplane illuminance at a certain point is contributed by multiple neighboring lighting fixtures and daylighting from multiple skylights. So that, the accurate dimming control for each lighting fixture is difficult.

Simulation model prediction can replace the mounted illuminance sensors. In this study, the author developed daylighting and electric simulation models that predict the illuminance level at any points of interest. Radiance, one of the most sophisticated lighting simulation tools, was employed for this purpose. Then, a surrogate model was developed for fast computation and optimal control of a target building.

For optimal dimming control, Deep Deterministic Policy Gradient (DDPG) is used. The DDPG makes an optimal policy network and gets feedback from the model, and finds continuous dimming optimal control variables as the output. In the paper, the development process of the lighting simulation model and the DDPG, as well as an application of the both to a real-life case would be addressed.

keywords : Illuminance prediction model, Radience simulation, Real time lighting control, Reinforcement learning, Deep Deterministic Policy Gradient(DDPG)

Student Number : 2019-21516

감사의 글

서른, 다시 공부를 시작했습니다. 새로운 아이디를 발급받아 접속한 이곳은 알던 곳이 아니었습니다. 건축 환경과 건물 에너지는 지금까지 접했던 설계 및 컨설팅 업무와 맥락은 닿았지만, 그 결이 달랐습니다. 학기 시작과 함께, 연구원으로 과제에 참여는 했지만, 무엇을 해야 할지 몰라 막막했습니다. 건물 에너지 분야의 도메인 지식이라 는 것이 내게는 거의 없었기에, 기억 저편에 있던 수학과 물리를 끄집어내면서 부지런히 공부했습니다. 동시에 최적화 방법과 베이지안 기법들, 그리고 인공지능 기술 등, 이름만 들어도 어려운 타 학문의 지식도 흡수했습니다. 필요한 공부와 순서를 알려주시고, 어려움과 부족함이 있을 때마다 필요를 채워준 모든 손길에 고마운 마음을 전합니다.

먼저 본 논문을 작성할 수 있도록 지도, 심사해주신 세 분 교수님께 감사드립니다. 제자로 받아주시고, 학업의 길을 열어주신 박철수 교수님, 감사합니다. 연구의 방향을 잡아주시고, 다른 시각으로 문제를 보는 가치관을 키워주셨습니다. 부족한 결과를 내밀 때도 '훌륭'이라는 한마디로 용기를 주시고, 잘못된 부분은 다시 생각할 기회를 주신 덕분에 포기하지 않고 논문을 마칠 수 있었습니다. 여명석 교수님께 감사드립니다. 첫 학기, 교수님의 강의를 통해, 건물 설비 시스템의 맥락을 잡을 수 있었습니다. 논문 심사에서는 세밀한 지도 덕분에 기본을 잡고 객관성을 확보할 수 있었습니다. 흔쾌히 논문 심사를 맡아주신 김선숙 교수님께 감사드립니다. 연구의 목적과 필요성, 전체적인 흐름을 잡아주시고, 과도한 내용을 덜어주신 덕분에 논문의 완성도를 높일 수 있었습니다.

학위 과정 동안, 함께했던 건축에너지연구실 선배와 구성원께 감사드립니다. 다른 연구주제이지만, 끊임없는 관심을 기울여주고, 다른 연구 내용도 알려주어 지적 호기심을 키워준 동혁이형. 기존 연구자료 및 내용을 비롯한 여러 도움이 필요할 때마다 가장 먼저 챙겨준 한경이형. 그동안, 연구는 물론, 출장과 회의, 제안서, 보고서 작성 등 가장 많은 활동을 함께한 한솔.

코딩하는 방법과 요령, 코드 한 줄 한 줄 모두 설명해주었던 재민. 건물 에너지 분야의 다양한 연구주제와 동향을 발 빠르게 전달해준 윤담, 앞서 진행했던 한국전력사옥 과제에서 주요 정보와 내용을 정리하여 전달해준 성호, 유쾌한 에너지를 채워주고, 맡은 역할을 성실히 해내준 영서, 힘겹고 어려운 학업 과제들을 모두 함께하며, 놓쳤던 부분을 검토하고 해결책을 찾아준 성권, 짖궂은 장난에도 화 한 번 내지 않고, 옆자리를 지켜주었던 영식, 부담스럽도록 많은 조명 연구주제를 함께 열정적으로 감당해준 형곤, 그래스호퍼 최적화라는 공감대와 도움을 요청할 때마다 자기 일처럼 나서준 철홍, 가장 적절할 시기에 PRISM 과제에 참여하여 부족한 자리를 완벽하게 메워준 선중. 모두에게 다시 한번 고마운 마음을 전합니다. 또한, STASIO를 통해 연구실의 거의 모든 것을 압축적으로 보여준 기언이형, 아주 짧은 시간이었지만, 연구실 생활의 노하우를 전수해 주었던 성호형께도 감사를 전합니다.

같은 공간에서 함께 연구한 AEPRL의 연구원들께도 감사합니다. 덕분에 연구주제에 대한 지평을 넓힐 수 있었습니다. 지속할 수 있는 건축과 환경에 대한 개념을 알려주고, 학위 과정의 동기를 심어준 정호건 대표님을 비롯한 코다의 임직원께도 감사 인사를 남깁니다.

마지막으로 버팀목이 되어준 사랑하는 가족에게 고마움을 전합니다. 힘겨웠던 모든 날을 함께 이겨낸 아빠, 엄마, 그리고 지영이에게 작은 위로가 될 수 있기를 바랍니다. 부족한 사위를 지지하고 응원해주신 부모님께도 감사를 전합니다. 서로를 돕고 배우며, 즐겁고 행복한 시간을 함께하는 평생의 배필, 사랑하는 아내 혜민에게 큰 고마움을 표합니다.

한 걸음 앞에서 길을 밝히시고 소망이 되어주신 하나님께 감사합니다.

2021년 1월 14일

김 영 섭