Automated Multi-layer Optical Design via Deep Reinforcement Learning

Haozhu Wang
Department of EECS
University of Michigan
Ann Arbor, MI 48105
`hzwang@umich.edu`

Zeyu Zheng
Department of EECS
University of Michigan
Ann Arbor, MI 48105
`hzwang@umich.edu`

Chengang Ji
Department of EECS
Inlight Technology Co., Ltd
Ningbo, Zhejiang, China 315500
`hzwang@umich.edu`

L. Jay Guo
Department of EECS
University of Michigan
Ann Arbor, MI 48105
`hzwang@umich.edu`

Abstract

Optical multi-layer thin films are widely used in optical and energy applications requiring photonic designs. Engineers often design such structures based on their physical intuition. However, solely relying on human experts can be time-consuming and may lead to sub-optimal designs, especially when the design space is large. In this work, we frame the multi-layer optical design task as a sequence generation problem. A deep sequence generation network is proposed for efficiently generating optical layer sequences. We train the deep sequence generation network with proximal policy optimization to generate multi-layer structures with desired properties. The proposed method is applied to two energy applications. Our algorithm successfully discovered high-performance designs, outperforming structures designed by human experts in task 1, and a state-of-the-art memetic algorithm in task 2.

1 Introduction

Optical multi-layer films have been widely used in many applications, such as broadband filtering [1], photovoltaics [2], radiative cooling [3], and structural colors [4]. The design of optical multi-layer films is a combinatorial optimization problem that requires one to choose the best combination of candidate materials and layer thicknesses to form a multi-layer structure. Researchers and engineers often make such designs based on their physical intuition. However, a completely human-based design process is slow and often leads to sub-optimal designs, especially when the design space is enormous. Thus, computational methods for designing optical multi-layer structures, including evolutionary algorithms [5, 6, 7], needle optimization [8], and particle swarm optimization [9], have been proposed to tackle this problem. All of these previous methods frame the optical design task as an optimization problem and aim to synthesize a structure that meets user-specified design criteria. However, these methods for optical design are based entirely on hand-crafts heuristics, i.e., they do not learn a model to solve the design problems. When the heuristics are sub-optimal for a task, the search process may fail to identify a high-performance design.

In contrast, deep reinforcement learning (DRL) is a learning framework that learns to solve complex tasks through a trial-and-error process. It is proven to be highly scalable for solving large-scale and complicated tasks [10, 11]. Researchers have successfully applied DRL to various combinatorial
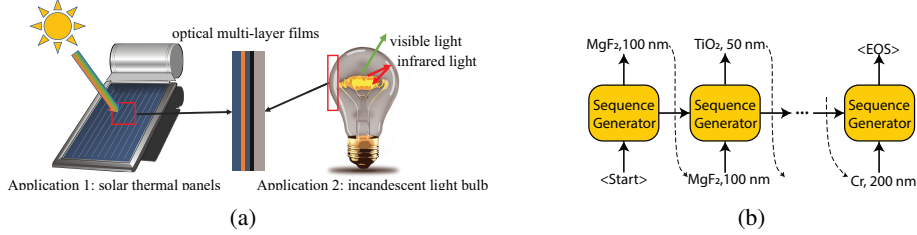
Figure 1: Illustration of two energy applications of optical multi-layer films (a) and sequence generation proces for optical multi-layer thin film designs. (a) For solar thermal panels, we can use multi-layer films as ultra-wideband absorbers to enhance light absorption efficiency. For incandescent light bulbs, we can coat multi-layer films on them to improve luminous efficiency by reflecting infrared light while transmitting visible light. (b) The generation process will stop when either the EOS token is sampled, or the length of the sequence reaches the maximum allowed length L.

optimization problems [12, 13, 14, 15]. Unlike heuristic-based search, reinforcement learning methods learn a model using the reward signal [16] and do not depend on hand-crafted heuristics. On some combinatorial optimization tasks, DRL has been shown to outperform classic heuristic search methods [17]. Recently, researchers applied DRL on designing optical devices with a structure template [18, 19], where the number of layers is fixed. However, when designing the optical multi-layer films, we often do not know the optimal structure template. Thus, the previous DRL approaches are not suitable for multi-layer designs.

Noticing that the multi-layer optical design task is equivalent to a sequence generation problem, we propose a DRL method called Optical Multi-layer Proximal Policy Optimization (OML-PPO) that can generate near-optimal multi-layer structures. We applied the proposed method to two optical design tasks that are relevant to energy applications (Figure 1a): 1) ultra-wideband absorbers that can enhance light-harvesting efficiency, e.g. for thermal photovoltaics and photothermal energy conversion 2) incandescent light bulb filters that can improve light bulb efficiency in emitting visible light. On these two tasks, we show that OML-PPO outperforms existing human expert designs and a state-of-the-art memetic algorithm.

2 Proposed Approach

Multi-layer films can be represented as sequences. Each layer is written as $s_l = (m_l, d_l)$. Thus, an N-layer structure is $\mathcal{S} = \{(m_0, d_0), (m_1, d_1), (m_2, d_2), \cdots, (m_{N-1}, d_{N-1})\}$, where m_l and d_l denote the material and the thickness of the l-th layer (counting from the top), respectively. When designing optical multi-layer films, we hope to synthesize a sequence that has the desired target spectral response $\tilde{T}$. Towards this end, we train a sequence generation network based on a GRU [20] with reinforcement learning[16]. The process for generating an optical layer sequence is illustrated in Figure. 1b.

Non-repetitive gating To explore the design space more efficiently, we introduce a non-repetitive gating function that removes the logit element corresponding to the most recently sampled material to prevent the sequence generator from generating the same materials in a row. This gating function is a matrix $I_{NR} \in \mathbb{R}^{|\mathcal{M}| \times (|\mathcal{M}|+1)}$ formed by removing the row corresponding to the most recently sampled material from an identity matrix. When multiplied with the logits vector σ_{m_l}, the element corresponding to that material will be removed, i.e., $\sigma'_{m_l} = I_{NR} \cdot \sigma_{m_l} \in \mathbb{R}^{|\mathcal{M}|}$. Then, we pass the transformed logit vector σ'_{m_l} to the softmax layer to obtain the sampling probability. By doing so, we set the sampling probability for the recurring material to 0. With the non-repetitive gating, the generated material sequence is guaranteed to have different materials for adjacent layers.

Auto-regressive generation of material and thickness Because the proper thickness of a layer should depend on the material, we input the sampled material m_l to the thickness MLP in addition to the hidden state h_l. A similar approach has been applied in RL problems where the actions are dependent on each other [11]. Instead of using a one-hot vector to represent the material, we train a material embedding matrix $emb \in \mathbb{R}^{|\mathcal{M}| \times d}$ together with the sequence generator network.
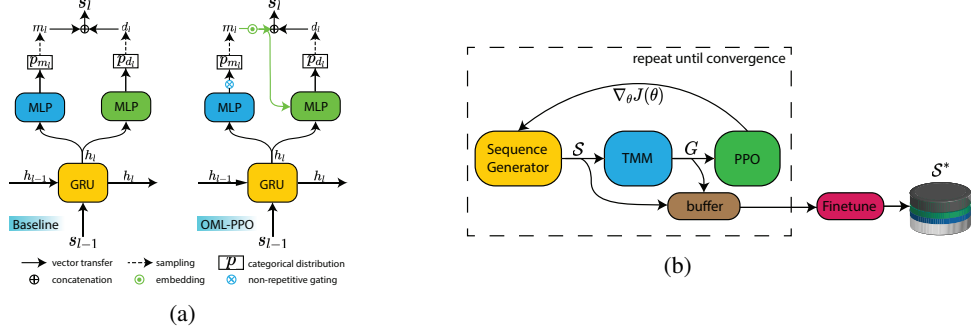
2

Figure 2: Neural network architectures for generating optical multi-layer films (a) and design generation pipeline (b). (a) Built-upon the baseline architecture (left), our proposed model (right) adds a non-repetitive gating function and auto-regressive connection between the sampled material and the thickness MLP. (b) Pipeline of the sequence generator training process.

Reinforcement learning training We train the sequence generation network with reinforcement learning. We set the reward to be 0 for all generation steps except the final step. At the final step (i.e., the structure $\mathcal{S}$ has been completely generated), we compute the spectrum of the generated structure with an optical spectrum calculation package TMM [21] and assign the final reward based on how well the structure spectrum matches with the target spectrum, i.e.,

$$G(\mathcal{S}) = 1 - \frac{1}{K} \sum_{k=0} \frac{1}{J} \sum_{j=0}^{J-1} |T^{\mathcal{S}}(\lambda_j, \delta_k) - \tilde{T}(\lambda_j, \delta_k)|, \tag{1}$$

where $T^{\mathcal{S}}(\lambda_j, \delta_k)$ is the spectrum of the generated structure $\mathcal{S}$ at wavelength λ_j under incidence angle δ_k. Because $T \in [0, 1]$, the cumulative reward is always non-negative. The reward value will become higher as the spectrum $T^{\mathcal{S}}$ gets closer to the target spectrum $\tilde{T}$ until it reaches 1 when the structure spectrum perfectly matches with the target spectrum. We use a state-of-the-art policy gradient algorithm PPO [22] to train the sequence generation network. Similar to the *active search* approach in Bello et al. [12], we output the best structure discovered throughout the entire training process (Figure 2b). Our model is implemented using PyTorch [23] and Spinning Up [24]. The data used in this study and our code will be publicly available.

3 Experiment

We applied the proposed method to two optical design tasks that are relevant to energy applications, i.e., 1) designing ultra-wideband absorbers and 2) designing incandescent light bulb filters. In task 1 ultra-wideband absorber design, we measure the quality of the designed structure by *average absorption*. In task 2 incandescent light bulb filter, we calculate the visible light *enhancement factor* to measure the performance of designed structures. We also did an ablation study to understand the effect of non-repetitive gating and auto-regressive materials/thickness sampling.

Task 1: ultra-wideband absorber Firstly, we apply our algorithm to the task of designing an ultra-wideband absorber for the wavelength range [400, 2000] nm. We choose the target spectrum as a constant 100% absorption under normal light incidence angle (i.e., the light is shining at the absorber at a right angle) to represent an ideal broadband absorber. This task has been previously studied by Yang et al. [1] based on physical models, where the broadband absorption is achieved by overlapping multiple absorption resonances and with an overall graded-index structure to minimize reflection. The authors designed a 5-layer structure using MgF_2, TiO_2, Si, Ge, and Cr. The simulated average absorption of their structure over the wavelength range is 95.37% under normal incidence. If not specified otherwise, we assume normal incidence when reporting average absorption.

We hypothesize that, when choosing from a larger set of materials than used in the previous work [1], it is possible to design a structure with higher average absorption than the human-designed structure. Thus, we expanded the original material set [1] to allow 11 more materials, including Ag, Al, Al_2O_3, Fe_2O_3, HfO_2, Ni, SiO_2, Ti, ZnO, ZnS, ZnSe. We set the available discrete thicknesses $\mathcal{D}$ to be

$\{15, 20, 25, \ldots, 200\}$ nm with a total of 38 different values. When training the sequence generator, we set the learning rate to 5×10^{-5} and the maximum length to $L = 6$. The material embedding size d is set to 5, i.e., $emb_m \in \mathbb{R}^5$. The generator is trained for a total of $3,000$ epochs with the batch size set to be $1,000$ generation steps. We repeat the training for 10 runs with different random seeds. The best structure discovered by the algorithm, exhibiting an average absorption of 97.64%, is {(SiO2, 115 nm), (Fe2O3, 70 nm), (Ti, 15 nm), (MgF2, 124 nm), (Ti, 148 nm)}. Its average absorption is 2.27% higher than expert designed structure. The spectrum under normal incidence are plotted in Figure 3a.
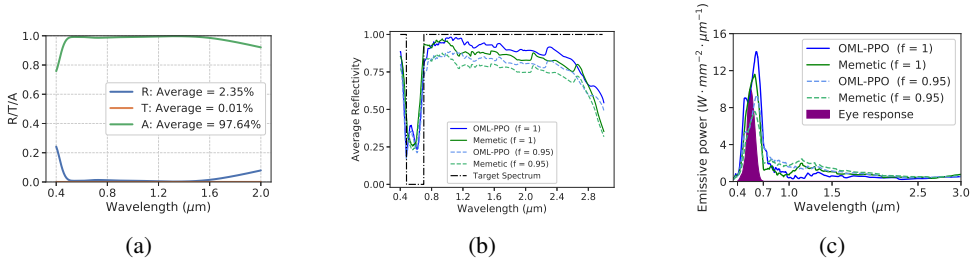

Figure 3: Spectrum of designed perfect absorber (a), incandescent light bulb filter (b), and emissive power spectrum of the light bulb filter (c). R: reflection, T: transimission, A: absorption. (a) The designed structure achieves almost perfect absorption between 500 to 1,600 nm. (b) The structure designed by our proposed algorithm achieves higher reflectivity in the infrared range than the structure designed by a memetic algorithm. We compute the spectrum under two view factors $f = 1$ and $f = 0.95$ (the ratio of light emitted by the light bulb filament that can reach the wall of the light bulb). (c) The structure designed by OML-PPO achieves higher emissive power than the previously reported structure.

Task 2: incandescent light bulb filter We further applied the proposed method to design a filter that can enhance the luminous efficiency of incandescent light bulbs [25, 26]. The goal is to reflect the infrared light emitted by the light bulb filament so that its energy can be recycled. To this end, we set the target reflectivity to be 0% in the range [480, 700] nm, and 100% outside this range (Figure. 3b). In this way, the infrared light, which cannot contribute to lighting, will be reflected back to heat up the emitter. A similar design has been previously studied [26, 6]. We choose the same seven dielectric materials as the available materials: Al_2O_3, HfO_2, MgF_2, SiC, SiO_2, and TiO_2 [6]. In Figure 3b, we compare the average reflectivity normalized over all incidence angles (0 - 90 degree) of the 42-layer structure designed with our algorithm and the 41-layer structure designed by a memetic algorithm [6]. Our structure has a higher average reflectivity in the infrared range (> 780 nm) than the 41-layer structure. Following a previous work[6], we calculated the enhancement factor for visible range (400 - 780 nm) under a fixed operating power. Our designed structure achieved an enhancement factor of **16.60** compared with **15.30** of a previously reported structure[6].

Ablation study On the ultra-wideband absorber design task, we conducted an ablation study to understand the effect of non-repetitive gating and auto-regressive generation of materials and thicknesses. We trained four different models: 1) OML-PPO with both non-repetitive gating and auto-regressive generation, 2) non-repetitive gating only, 3) auto-regressive generation only, 4) neither non-repetitive gating nor the auto-regressive generation. For each model, we repeated the training for ten times. The maximum absorption values discovered by each model before finetuning are reported in Table 1. Both non-repetitive gating and the auto-regressive material/thickness generation improve the performance of the baseline model.

Table 1: Highest absorption values discovered by each algorithm across 10 runs. The mean average absorption values and standard deviations of the 10 runs are reported.

Model	OML-PPO	Only Gating	Only auto-regressive	None (baseline)
Average Absorption	**94.98%(0.99%)**	94.05%(1.39%)	91.55%(1.14%)	91.03%(0.87%)

4

4 Conclusion

We introduced a novel sequence generation architecture and a deep reinforcement learning pipeline to automatically design optical multi-layer films. To the best of our knowledge, our work is the first to apply deep reinforcement learning to design multi-layer optical structures with the number of layers unfixed beforehand. On two real-life applications, we show that our proposed algorithm achieves better performance than human experts and a state-of-the-art memetic algorithm. In the future, we plan to apply the proposed algorithm for more complicated designs that involve micro-nano structures [27].

Broader Impact

Optical multi-layer thin films are widely used in imaging and energy applications. With our proposed algorithm, researchers can design multi-layer thin films automatically with higher performance than existing approaches. In addition, we believe that our algorithm can be applied to many other optical design tasks based on multi-layer structures.

References

[1] Chenying Yang, Chengang Ji, Weidong Shen, Kyu-Tae Lee, Yueguang Zhang, Xu Liu, and L Jay Guo. Compact multilayer film structures for ultrabroadband, omnidirectional, and efficient absorption. *Acs Photonics*, 3(4):590–596, 2016.

[2] Mukul Agrawal and Peter Peumans. Broadband optical absorption enhancement through coherent light trapping in thin-film photovoltaic cells. *Optics express*, 16(8):5385–5396, 2008.

[3] Aaswath P Raman, Marc Abou Anoma, Linxiao Zhu, Eden Rephaeli, and Shanhui Fan. Passive radiative cooling below ambient air temperature under direct sunlight. *Nature*, 515(7528):540–544, 2014.

[4] Wei Li, Yu Shi, Zhen Chen, and Shanhui Fan. Photonic thermal management of coloured objects. *Nature communications*, 9(1):1–8, 2018.

[5] Martin F Schubert, Frank W Mont, Sameer Chhajed, David J Poxson, Jong Kyu Kim, and E Fred Schubert. Design of multilayer antireflection coatings made from co-sputtered and low-refractive-index materials by genetic algorithm. *Optics express*, 16(8):5290–5298, 2008.

[6] Yu Shi, Wei Li, Aaswath Raman, and Shanhui Fan. Optimization of multilayer optical films with a memetic algorithm and mixed integer programming. *ACS Photonics*, 5(3):684–691, 2017.

[7] Chenglong You, Corey T Matyas, Yin Huang, Jonathon Dowling, and Georgios Veronis. Optimized multilayer structures with ultrabroadband near-perfect absorption. *IEEE Photonics Journal*, 2020.

[8] Alexander V Tikhonravov, Michael K Trubetskov, and Gary W DeBell. Application of the needle optimization technique to the design of optical coatings. *Applied optics*, 35(28):5493–5508, 1996.

[9] Rabi I Rabady and Almahdi Ababneh. Global optimal design of optical multilayer thin-film filters using particle swarm optimization. *Optik*, 125(1):548–553, 2014.

[10] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.

[11] Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 2019.

[12] Irwan Bello, Hieu Pham, Quoc V Le, Mohammad Norouzi, and Samy Bengio. Neural combinatorial optimization with reinforcement learning. *arXiv preprint arXiv:1611.09940*, 2016.

[13] Elias Khalil, Hanjun Dai, Yuyu Zhang, Bistra Dilkina, and Le Song. Learning combinatorial optimization algorithms over graphs. In *Advances in Neural Information Processing Systems*, pages 6348–6358, 2017.

[14] Azalia Mirhoseini, Hieu Pham, Quoc V Le, Benoit Steiner, Rasmus Larsen, Yuefeng Zhou, Naveen Kumar, Mohammad Norouzi, Samy Bengio, and Jeff Dean. Device placement optimization with reinforcement learning. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 2430–2439. JMLR. org, 2017.

[15] Azalia Mirhoseini, Anna Goldie, Mustafa Yazgan, Joe Jiang, Ebrahim Songhori, Shen Wang, Young-Joon Lee, Eric Johnson, Omkar Pathak, Sungmin Bae, et al. Chip placement with deep reinforcement learning. *arXiv preprint arXiv:2004.10746*, 2020.

[16] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[17] Hao Lu, Xingwen Zhang, and Shuang Yang. A learning-based iterative method for solving vehicle routing problems. In *International Conference on Learning Representations*, 2020.

[18] Iman Sajedian, Trevon Badloe, and Junsuk Rho. Optimisation of colour generation from dielectric nanostructures using reinforcement learning. *Optics express*, 27(4):5874–5883, 2019.

[19] Iman Sajedian, Heon Lee, and Junsuk Rho. Double-deep q-learning to increase the efficiency of metasurface holograms. *Scientific reports*, 9(1):1–8, 2019.

[20] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.

[21] Steven J Byrnes. Multilayer optical calculations. *arXiv preprint arXiv:1603.02720*, 2016.

[22] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[23] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, pages 8024–8035, 2019.

[24] Joshua Achiam. *Spinning Up in Deep Reinforcement Learning*. 2018.

[25] Jing Zhou, Xi Chen, and L Jay Guo. Efficient thermal–light interconversions based on optical topological transition in the metal-dielectric multilayered metamaterials. *Advanced Materials*, 28(15):3017–3023, 2016.

[26] Ognjen Ilic, Peter Bermel, Gang Chen, John D Joannopoulos, Ivan Celanovic, and Marin Soljačić. Tailoring high-temperature radiation and the resurrection of the incandescent source. *Nature nanotechnology*, 11(4):320, 2016.

[27] Wei Li and Shanhui Fan. Nanophotonic control of thermal radiation for energy applications. *Optics express*, 26(12):15995–16021, 2018.