Optimization of Optical Structures Using Markov Decision Processes

Tyler Hughes and Yu (Jerry) Shi

Abstract

We present an algorithm for optimizing optical structures based on Markov Decision Processes. Our method is successfully demonstrated on a one-dimensional stack of dielectric slabs, where we wish to choose a combination of slab thicknesses and refractive indices to approximate a target reflection spectrum. We show that this algorithm is less susceptible to being caught in local minimum and has favorable computational scaling as the number of layers is increased.

1. Introduction

A major challenge in optical and electrical engineering is designing a structure with tailored optical properties. Tweaking physical parameters to optimize an optical device is often time consuming, unsystematic, and requires large amounts of expertise. While systematic optimization methods for optics and photonics exist (1) (2), methods that employ machine learning concepts are not found in the literature.

We propose a method of optimization of optical devices that is based on a Markov Decision Process (3) (MDP), which is commonly used for problems in control. This method uses reinforcement learning to construct a policy for perturbing structure parameters to minimize spectral error. The algorithm will be discussed in more detail in section 3.

We consider a physical system comprising of a one-dimensional stack of dielectric slabs as shown in figure 1. When light of a certain wavelength, λ, is incident on the stack, some fraction $r(\lambda)$ will be reflected and some fraction $t(\lambda)$ will be transmitted, depending on the properties of the layers in a highly non-linear way. We wish to design a structure that will have a spectrum, $r(\lambda)$ that is as close to some target spectrum as possible. A structure with K layers has $2K$ independent parameters, namely the thickness d_i and refractive index n_i of each layer.

Proceedings of the 32^{nd} International Conference on Machine Learning, Lille, France, 2015. JMLR: W&CP volume 37.

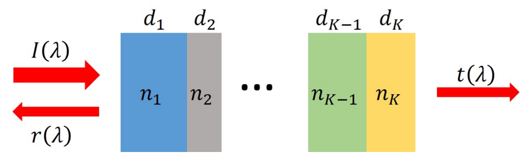

Figure 1. Schematic showing the physical system under investigation. Light of wavelength (λ) is incident on a series of slabs with thicknesses d_i and refractive indeces n_i. A fraction of optical power $r(\lambda)$ is reflected.

We wish to feed our algorithm a target spectrum and have it return the optimal structure parameters for a number of layers. A schematic of our problem is given in figure 2.

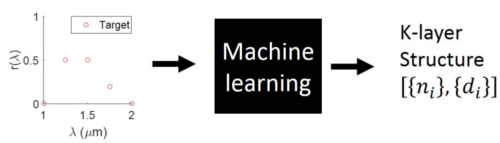

Figure 2. General schematic of the problem description. A target spectrum is proposed, which is fed through our MDP optimization algorithm, which returns a structure that has a spectrum that approximates the target.

One application of this project is in designing antireflective coatings for solar cells, which aim to minimize the amount of light that gets reflected (4). Additionally, there has been recent work on developing layered structures that reflect light in the visible spectrum and transmit in a specific window of the earth's atmospheric absorption spectrum. These devices reflect sunlight and radiate their heat into space to maintain temperatures below ambient, for use in passive refrigeration (5). Furthermore, our result can easily be generalized to other physical systems that can be parametrized by a limited number of quantities, so there are many potential applications beyond the dielectric slab stack.

We will start by explaining the method we use to calculate spectra from a set of structure parameters in section 2. Then we will discuss our MDP algorithm for optimizing these structures in section 3. Finally, we will present our results and a discussion on the usefulness and applications of our approach in section 4.

2. Transfer Matrix Method

The transfer matrix method (TMM) is a well-known optical computation technique that computes the propagation of light inside a layered medium using 2×2 matrices. When light is incident from one medium to another, the transfer matrix is written as

$$T_n(n_2, n_1) = \frac{1}{2n_2} \begin{pmatrix} n_2 + n_1 & n_2 - n_1 \\ n_2 - n_1 & n_2 + n_1 \end{pmatrix}, \tag{1}$$

where n_1 and n_2 are the incident and transmitted material's refractive indices, respectively; on the other hand, when light at wavelength λ propagates for a distance x in a material with refractive index n_r, the matrix is written as

$$T_x(n_r, x, \lambda) = \begin{pmatrix} \exp(-i\frac{2\pi n_r x}{\lambda}) & 0 \\ 0 & \exp(i\frac{2\pi n_r x}{\lambda}) \end{pmatrix}. \tag{2}$$

When we have a known geometry, the total transfer matrix, $T_{tot}(\lambda)$ for the propagation of light at each wavelength can be computed with successive multiplications of transfer matrices:

$$\begin{aligned} T_{tot}(\lambda) = T_n(n_{out}, n_k)T_x(n_k, d_k, \lambda)\times \\ \ldots T_n(n_k, n_{k-1})T_x(n_{k-1}, d_{k-1}, \lambda)\times \\ \ldots T_n(n_2, n_1)T_x(n_1, d_1, \lambda)\times \\ \ldots T_n(n_1, n_i n). \end{aligned} \tag{3}$$

In our simulations, we have assumed that the geometry is surrounded by air, so $n_{in} = n_{out} = 1$.

The overall reflection from this structure can be obtained simply by looking at two of the matrix elements of $T_{tot}(\lambda)$:

$$R(\lambda) = \left| \frac{[T_{tot}(\lambda)]_{21}}{[T_{tot}(\lambda)]_{11}} \right|^2 \tag{4}$$

To calculate a spectrum, which we will discretize and define as our 'state' in the next section, we simply sample W points in wavelengths: $\lambda = [\lambda_1, ..., \lambda_W]$.

3. Markov Decision Process

To implement the Markov Decision Process (MDP) algorithm, we consider a $K-$layer structure with $2K$ independent structure parameters, $[\{n_i\}, \{d_i\}]$. The essence of MDP is to update Bellman's equation at state s until $V(s)$ converges:

$$V(s) = R(s) + \gamma \max_{a \in A} \sum_{s' \in S} P_{sa}(s')V(s'), \tag{5}$$

where $R(s)$ is the reward at state s, $\gamma = 0.995$ is the discount rate, a is the action taken at state s, and $P_{sa}(s')$ is the transition probability from state s to s' under action a. To set up the MDP problem in the context of designing an optical structure, we need a set of discretized states S, action space A, reward $R(s)$, and value $V(s)$.

The states of MDP can be characterized by all set of possible spectra within the wavelengths of interest. In discretizing such an infinite dimensional state-space, we sample the reflection spectrum at W wavelengths and discretize the reflectivity at each wavelength into Q regions [Fig. 3]. With such disretization, we obtain Q^W states. For instance, the partition in Fig. 3 has a total of $3^5 = 243$ states, and the reflection spectrum shown in red is at state $s = 82$. We further define the reward of each state such that it penalizes spectra that deviate from the target spectrum.

$$R(s) = -||r_{target} - r_{struc}([\{n_i\}, \{d_i\}])||_2^2. \tag{6}$$

Given a structure whose spectrum is at state s, the action is taken as a perturbation of one of the structure parameters. Specifically, we vary one of the layer refractive indices by ± 0.01 or one of the layer lengths by ± 20 nm. Hence, there are $4K$ possible actions.

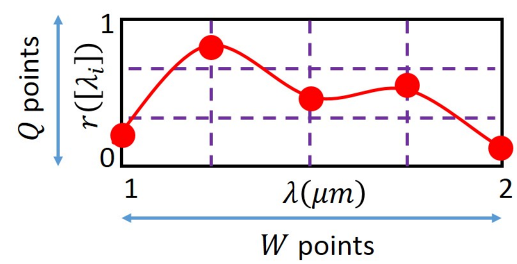

Figure 3. Discretization of a spectrum into states. The red curve is the continuous reflection spectrum, and the red dots are the sampled reflection values.

We outline the iterative MDP algorithm in Fig. 4. First, we start with a random structure whose spectrum is at state s. Then we consider taking all $4K$ actions and compute each of their rewards at the resulting states. We define the policy $\pi(s)$ as the optimum action a taken at state s such that $P_{sa}(s')V(s')$ is maximized. Upon transitioning from state s to s' under action a, we keep track of the transition rate between s and s' with action a as well as the reward at state s, such that we can continuously update $P_{sa}(s')$ and $R(s)$. For each starting structure, we exit an iteration of learning when the structure parameters deviate from realistic material parameters or when the state transition happens indefinitely between two states.

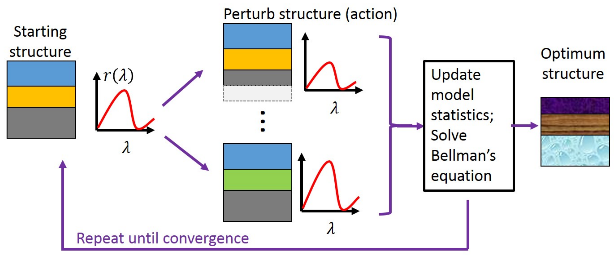

Figure 4. Schematics of the MDP process.

We denote learning convergence as the case when for every state updated with Eqn. (5), every $V(s)$ converges with exactly 1 iteration. After converging 50 times, we terminate the learning process, from which we obtain the optimal policy $\pi^*(s)$ at state s. In the next section, we demonstrate that by using $\pi^*(s)$, we can start from a random state and find the optimum structure whose reflection spectrum matches the target spectrum.

4. Results and Discussion

As proof of concept and illustration of our algorithm, we consider a system with a single slab ($K = 1$) with thickness d_1 and refractive index n_1. We specify a target spectrum over five wavelengths ($W = 5$) and with a resolution in reflection of 0.1 ($Q = 10$). Using TMM, we can calculate the mean squared error between our structure's spectrum and the target spectrum for each combination of n_1 and d_1, which is plotted in figure 5. We limit search space in n_1 and d_1 to the ranges $\{1, 4\}$ and $\{0, 1\}$, respectively. The purpose of our algorithm is to find a combination of n_1 and d_1 that minimizes the error corresponding to the error plot in figure 5.

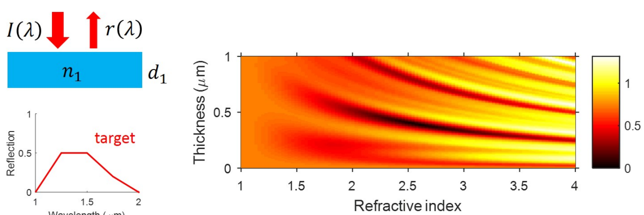

Figure 5. Top Left: Schematic of the one layer stack system being analyzed. Bottom Left: Target spectrum. Right: Brute force calculation of the error between calculated spectrum and target spectrum as a function of structure parameters.

After running a total of 426 training runs with a γ of .995, our MDP algorithm converges. We plot the progression of a randomly generated structure as it converges to the global minimum in the error plot of figure 6. The plot on the right hand side shows the spectrum of the optimized structure is in good agreement with the target spectrum. Notice that the trajectory traverses a local maximum before converging at the global minimum. This feature suggests that the MDP optimization is less likely to settle in local minima than a method such as gradient descent, which uses a local search to move through parameter space.

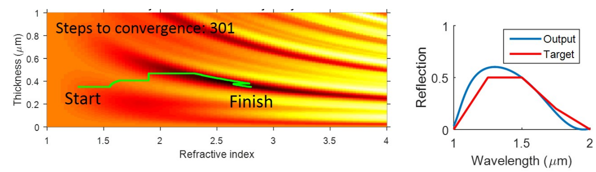

Figure 6. Left: Trajectory (green) of a structure through parameter space during a run of our MDP algorithm. Right: The final structure (blue) and target (red) spectra.

Over the course of 5000 trials, our algorithm converges to the global minimum for 23.5% of the initial random parameters, which is quite high considering the small basin of attraction of the global minimum.

When the number of layers is increased beyond $K = 1$, our algorithm has trouble converging as evidenced by Figure 7. An increase in one layer adds four more possible actions to the model (two $[+, -]$ perturbations for two $[n_K, d_K]$ parameters), which greatly increases the amount of time until convergence.

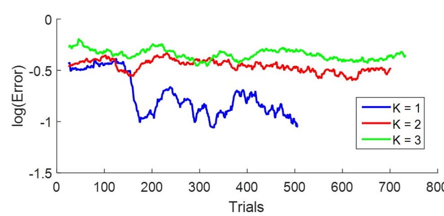

Figure 7. log(Error) between the spectrum of the optimized spectrum and the target spectrum as a function of number of trials for different layer thicknesses (K).

We believe that this physical system may be not well suited for an application of the MDP algorithm due to the large action and state space. However, by a simple counting argument, we can show that our algorithm scales with K much better than brute force calculation. The size of the action states is proportional to K, which means the number of calculations needed to populate a model $P_{sa}(s')$ is also proportional to K. Conversely, a brute force optimization method (calculation of fitness for all combinations of parameters) depends on the desired resolution "r" and K as r^{2K}.

Although our MDP optimization algorithm may not be optimal for designing the reflection spectrum of a multilayer stack, we show it does have success for limited layers. Our method was able to locate a global minimum in a small basin of attraction for a reasonable percentage of trials. Furthermore, in principle, the method could be generalized to other problems perhaps with better success if the number of states and actions is smaller.

Since the application of machine learning in optical design is an unexplored territory, this work is a first step in this direction. We anticipate there will be applications in this area that could benefit immensely from this method of analysis, but it will take some future experimentation to find systems that can take full advantage of our algorithm and other tools in machine learning at large.

References

[1] Lu, Jesse, and Jelena Vuckovic. ”Inverse design of nanophotonic structures using complementary convex optimization.” Optics express 18.4 (2010): 3793-3804.

[2] Veronis, Georgios, Robert W. Dutton, and Shanhui Fan. ”Method for sensitivity analysis of photonic crystal devices.” Optics letters 29.19 (2004): 2288-2290.

[3] Ng, A.,CS294A Lecture Notes - Stanford University.

[4] Wright, Daniel N., Erik S. Marstein, and Awe Holt. ”Double layer anti-reflective coatings for silicon solar cells.” Photovoltaic Specialists Conference, 2005. Conference Record of the Thirty-first IEEE. IEEE, 2005.

[5] Raman, Aaswath P., et al. ”Passive radiative cooling below ambient air temperature under direct sunlight.” Nature 515.7528 (2014): 540-544.