Deep search methods for multilayer coating design

M. TRUBETSKOV¹,²

¹Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, Garching 85748, Germany
²OptiLayer GmbH, Watzmannring 71, Garching 85748, Germany (Michael.Trubetskov@mpq.mpg.de)

Received 17 September 2019; revised 4 November 2019; accepted 16 November 2019; posted 18 November 2019 (Doc. ID 378150); published 13 December 2019

Many existing well-known multilayer design methods are based on so-called greedy algorithms. New deep search algorithms developed for needle optimization, gradual evolution, and design cleaner methods are presented. The algorithms possess machine learning features. The advantages of the deep search methods are demonstrated on a set of examples including the OIC Design Contest 2019. © 2019 Optical Society of America

https://doi.org/10.1364/AO.59.000A75

1. INTRODUCTION

During the last two decades, the needle optimization [1] and gradual evolution [2] methods have become valuable and extremely powerful tools applied to designing the most complicated multilayer optical coatings operating in spectral ranges from far the infrared to soft x-ray. The types of multilayers include antireflection coatings, mirrors, beam splitters, polarizers, filters of various types, coatings for precise polarization and phase control, coatings for ultrafast applications (dispersive mirrors), as well as coatings for color management, for displays, for solar cells, for architectural glass, and for many other applications. Additional production-friendly tools, such as the design cleaner and thin layer removal [2,3] as well as various specialized design techniques [3] allow one to design coatings, which fulfill better feasibility demands. For a wide set of problems, the needle optimization and gradual evolution methods enable obtaining very good solutions without any starting design at all and without any additional aid from the designer. These problems include design of dielectric or slightly absorbing coatings with target requirements on reflectance, transmittance, or absorbance for different polarizations for normal and for oblique light incidence. From the mathematical point of view, these problems are problems that obey (or at least do not strongly violate) the so-called maximum principle in the optimal control theory, formulated for multilayer optical design problems in Ref. [4].

At the current state of the art, many challenging novel applications require coatings with even more sophisticated target requirements including phase on transmittance or reflectance (coatings for polarization and phase control), differential phase shifts, group delay (GD) and group delay dispersion (GDD) (coatings for ultrafast applications), coatings with metal or metal-island layers, or coatings with conductive layers. In order to solve these problems, even an experienced designer needs to apply additional expertise, since standard needle optimization and gradual evolution algorithms often do not converge to good solutions automatically and thus require quite complicated manual tuning and control of the design process. For phase-related and ultrafast design problems, a good—or at least suitable—starting design is usually required. For metal-dielectric coatings, a designer should carefully control the insertions of metal layers during the design process. All these operations require manual control of computations and application of different design tricks (see Ref. [3], for example), and they are therefore rather difficult and time-consuming. Hence, the development and harnessing of new methods that are able to solve these problems in an automatic or semi-automatic mode is a very essential and important goal, since the computational time is nowadays much cheaper than the working time of an experienced designer.

2. GREEDY ALGORITHMS IN THIN-FILM OPTICS

Many of the algorithms used in thin-film optics belong to the class of so-called greedy algorithms [5]. A greedy algorithm at each step selects the best possible choice without regard to possible future development of computations. A good example is the well-known steepest descent refinement method used in multilayer optics for the first time in Ref. [6]. At each iteration it selects the direction of the anti-gradient (the steepest descent direction) as a direction of one-dimensional search of a local minimum. This concept is very simple for understanding and for the implementation, but with this strategy the convergence to a minimum is usually very slow. As a result, the steepest descent method is very rarely used nowadays, since
many significantly faster and more efficient methods based on more sophisticated strategies are available [7,8].

The flip-flop design method [9,10] also belongs to the class of greedy algorithms—at least in its initial implementation. Indeed, as stated in Ref. [10], “The merit function is evaluated for the starting configuration and its value noted. One by one each of the thin layers is flipped from its current index to an alternative index and the merit function is again evaluated. If the merit function has improved, then the new index state is retained, and the next layer is considered.” In order to improve the performance of the flip-flop method, additional approaches, such as delayed updates, were also proposed in [10].

Evolutionary methods (for example, [11,12]) and numerous bio-inspired approaches, like genetic optimization [13,14], particle swarm optimization [15], and ant colony algorithm [16], intensively use greedy algorithms at internal steps, since it is impossible to predict the global improvement of the whole optimizing system during the iterations.

The standard needle optimization [1] and gradual evolution [2] methods also use greedy algorithms at each iteration. For example, the needle optimization at each iteration (Fig. 1) performs the computations of the perturbation function (P function [1]). This function describes the variation of the merit function δMF due to insertions of infinitely thin layers of a material with refractive index \( \hat{n} \) at any position \( z \) of the coating cross-section as

\[
\delta MF = P(\hat{n}, z)\delta z + o(\delta z).
\]

(1)

Infinitely thin layer thickness is designated as \( \delta z \); \( o(\delta z) \) are terms decreasing faster than \( \delta z \). Therefore, the \( P \) function describes variations of the merit function (MF) as the first-order linear approximation.

In order to determine the best position for a new layer insertion, the standard needle optimization searches for the place where the \( P \) function has the most negative value, since at this position the MF has the fastest decrease (Fig. 2) in the frame of the linear approximation Eq. (1).

In Fig. 2, one of the intermediate iterations in the “cissoids of Diocles” example is shown (see the next section). The \( P \) function is computed for all materials available for needle optimization, and the selection of the insertion point determines the material that should be inserted at this position. This strategy has close analogy to the steepest descent method discussed in the previous section, and it may not lead to the best solution in some cases.

The standard gradual evolution procedure increases the thickness of each layer one by one until the next local one-parametric MF minimum is located. It also tries to insert materials other than adjacent to each boundary, if the number of available materials is more than two. Usually such thickness increases or material insertions lead to an increase (degradation) of the MF. All layers are considered one by one, the step providing the smallest degradation of the MF is finally selected (the greedy algorithm), and then the obtained design is refined. Quite often, this greedy strategy is not optimal. An obvious example is the case in which one of the materials’ refractive indices is close to the substrate refractive index. In this case, the insertion of such material at the boundary between the substrate and the coating obviously provides the smallest degradation of the MF, but this is not the best step in most cases.

The standard version of the design cleaner [2] also uses the greedy algorithm, since on each iteration the design cleaner removes a layer that introduces the smallest increase of the MF without any further analysis. In many cases, another layer removal provides better results after the refinement step concluding the design cleaner operation.

3. DEEP SEARCH ALGORITHMS

Deep search algorithms use a significantly more complicated and computationally demanding approach. Instead of a simple selection of the step giving the largest decrease (or the smallest possible increase) of the MF, all possible steps are considered, and for each possible step, a refinement with respect to all layers of a new design is applied. For example, during the needle optimization step, the needle insertions are performed at all local minima of the \( P \) function (Fig. 2) one by one, with the refinement of the design afterwards. MF values obtained after these refinements are compared, and the next step is selected in order to perform the best improvement among all refined designs. Considering all possible steps and performing subsequent refinements allows one to select a much better trajectory of the design. As a result, deep search algorithms often select strategies different from greedy ones. For example, the deep search
needle method selects a different place for the layer insertion (Fig. 2), since it leads to a better MF value after the subsequent optimization step.

The only obvious disadvantage of this approach is that computational cost is rapidly growing with the number of possible variants usually proportional to the number of design layers. Most of the refinement results are rejected; only the best one is actually used, and it seems to be very wasteful in relation to the computational resources. To a reasonable extent this problem could be tackled by recording statistics of local optimization iterations and performing early termination of the refinement processes that have little chance of providing a solution with low enough MF values. For example, one can store several of the best refinement trajectories (dependencies of the MF values on the iteration counter) and on subsequent iterations compare the current trajectory with the best already available ones. With reasonable decision criteria it is possible to terminate most of iterations in very early stages and thus save a significant amount

Fig. 3. Three-line filter target (magenta crosses) and spectral performance of classical and deep search designs (two black lines, indistinguishable in this resolution).

Fig. 4. Layer thicknesses of the three-line filter obtained with the classical approach (top) and with the deep search methods (bottom).
of processor time. If the current trajectory is better than the stored ones, it should replace one of the stored trajectories; therefore, an implementation of the deep search algorithm has some elements of machine learning.

A huge advantage of the deep search methods is that it is much more reliable and has faster convergence to solutions of extremely complicated design problems as illustrated by several examples in the next section.

4. EXAMPLES

To illustrate the benefits of the developed deep search methods, problems with already published results are considered. Several examples are related to OIC Design Contests of different years, since in this case the comparison is performed with the best results known so far.

A. Three-Line Narrow-Band Transmission Filter

The first example is a three-line filter with narrow transmission zones around 450, 510, and 640 nm, already considered in Ref. [3]. Transmittance targets are shown in Fig. 3 by magenta crosses. The substrate refractive index is 1.52, the incident medium is air, and the high and low refractive indices are 2.35 and 1.45, respectively. Since the normal light incidence case is considered and layer materials are non-absorbing, the maximum principle [4] leads in this case to the conclusion that two materials are sufficient to solve this problem. Even more, the classical “greedy” versions of the gradual evolution and the needle optimization algorithms are expected to be efficient. In order to check this, both classical and deep search versions of the methods were applied to this problem. As the result, two three-line filter designs were obtained, and their spectral performances are shown in Fig. 3 by black curves. They are quite close to each other, and as a result they are indistinguishable in the scale of Fig. 3. The optical thicknesses of the layers are presented in Fig. 4, where the layers are numbered in the direction from the substrate to the incident medium. One can see that the obtained designs are rather different. In order to perform more direct comparison of these approaches, an additional cleaning of the solutions has not been applied; only layers with thicknesses less than 5 nm were removed [2].

Classical and deep search synthesis computations provided quite similar behavior of the evolution of the MF values (Fig. 5). The computational time of the deep search variant is somewhat higher than that of the classical one (7:51 min versus 3:03 min on an Intel Xeon 10 core W-2155 CPU), but this difference is quite affordable. In general, for this case, the deep search method does not provide noticeable advantages, yet the obtained solution structure is significantly different, and it can be useful for the search of designs with better manufacturability [3]. This example clearly demonstrates that classical versions of the needle optimization and gradual evolution are efficient enough for the problems with non-absorbing layers with transmittance or reflectance target requirements at normal incidence.

B. OIC 2016 Problem B, Cissoids of Diocles

Problem B1 of the OIC 2016 Design Contest [17] allowed one to use metal layers to achieve the reflectance curve having sharp turns (the cissoid of Diocles challenge; the target is shown by magenta crosses in Fig. 6). Compared to a completely dielectric case, thin metal layers are allowed to achieve significantly smaller values of the MF within the other constraints of the contest problem, limiting the maximum number of layers to 50 and minimal layer thickness to 3 nm.

In Fig. 7, the trajectories of the standard and deep search gradual evolution methods are shown. In order to proceed with iterations of the standard version of the gradual evolution, the manual increase of an outermost layer thickness was performed several times, since without this the method generated many very thin layers in the design with negligible MF decrease. The deep search method was able to proceed without any manual adjustments of the calculations. In general, the deep search method generates design solutions with noticeably smaller MF values and smaller total thicknesses. At the final stage, the deep search design cleaner procedure was applied in order to obtain solutions with the number of layers not exceeding 50 (contest
requirement). The deep search design cleaner method removes each layer separately, preserving the total optical thickness of the design, and then re-optimizes the MF. The next step (the layer to be actually removed) is selected according to the smallest refined MF after all layer removal attempts.

The deep search method solution has 

\[ MF = 0.0238 \]

this solution is already between the first (\( MF = 0.0139 \)) and second (\( MF = 0.0294 \)) place winners of the OIC 2016 Design Contest (see [17], Table 10).

C. One-Octave Dispersive Mirror

The design of coatings for phase control, including such coatings for ultrafast applications as dispersive mirrors, requires knowledge of a good starting design. Without a starting design, the standard gradual evolution often fails to provide reasonably good solutions. For example, for the extremely challenging problem of a one-octave dispersive mirror [18], a chirped mirror was used as a starting design. The new deep search gradual evolution method can find a 62-layer solution to the problem without any starting design in a completely automatic mode. In Fig. 8, the reflectance and GDD spectral dependencies of this solution are compared with the 68-layer design [18] obtained with the classical needle optimization method. The layer thicknesses of the obtained solutions are presented in Fig. 9.

One can see that practically the same performance has been achieved with a simpler design having six fewer layers. It is also interesting to note that the layer thicknesses of the first half of the design constitute a structure resembling the well-known chirp mirror, when layer thicknesses are gradually decreasing in the direction from the substrate to the incident medium (Fig. 9).

D. Design Contest 2019 Problem A

Design Contest 2019 Problem A [19] was formulated exactly as Design Contest 2007 Problem B [20]. This challenge was to design a beamsplitter that is nonpolarizing in both intensity and phase at 45° incidence with maximum possible bandwidth around the central wavelength of 550 nm. The purpose of this contest was to see the progress achieved during the last years in multilayer design. It was also not clear which intermediate material is better for this problem: one with a refractive index of \( n = 1.65 \) or one with \( n = 1.8 \), since both cases provided solutions with the highest bandwidth (see Fig. 8 in [20]).

Deep search methods were applied by the author to the solution of this problem for the first time his knowledge. Due to the extremely complicated nature of this contest problem it was necessary to add a significant amount of manual control of the design process in order to achieve the best possible results. Furthermore, another deep search method was specifically developed and applied to the solution of this problem. This method performs all possible increases and decreases of layer thicknesses by two quarter-wave values and insertions of two quarter-wave layers of the third material at all boundaries. The designs obtained with this procedure are refined, and the resulting MFs after refinements are compared. The best variant is selected for the following calculations for the deep needle or/deep gradual evolution algorithms. Of course, for this problem the quarter-wave thickness is matched to the 45° angle of incidence. Computations started without any starting design, since the gradual evolution method is able “to grow” the design in the course of the iterations.

As the result of lengthy computations involved all deep search methods, winning designs were obtained for both cases of the intermediate material selection. The highest filter bandwidth...
of 63.932 nm was obtained for the intermediate material with refractive index $n = 1.65$; this result won the contest. The reflectance and differential phase shifts on reflectance and transmittance of this design are shown in Fig. 10. The design has 373 layers, and the physical thickness was 21.631 µm.

If another intermediate material with refractive index $n = 1.8$ is selected instead of the material with $n = 1.65$, the deep search methods provided results with a slightly smaller bandwidth of 63.732 nm. The design has 386 layers, and the physical thickness was 22.892 µm. This result is still better than the rest of the results submitted to Design Contest 2019 [19]. The reflectance and differential phase shifts on reflectance and transmittance of this design are shown in Fig. 11. Compared to Design Contest 2007, the improvements of the obtained bandwidth are 8.94% and 8.60% [19].

E. Design Contest 2019 Problem B

In Problem B it was necessary to optimize a light mixing system, combining four light sources with different peak wavelengths to a single output beam [19]. Three substrates with six coated surfaces constituted the light mixing system. The goal of the design was to maximize the throughput of the system and to minimize the phase difference between polarizations. This goal was specified in Ref. [19] as a special MF taking into account these requirements.

It was possible to implement a generalization of the conventional and deep search variants of the needle optimization, gradual evolution, and design cleaner methods applicable to this light mixing system. For this purpose, the $P$ function Eq. (1) was computed for all six coatings simultaneously, and all minima of the $P$ function across all coatings were considered either by a greedy approach or by deep search method. Similarly, all layers and all interfaces of all six coatings were considered for thickness increase or layer insertion in the frame of the gradual evolution step. In the design cleaner procedure, all layers of all coatings were considered during the iterations.

In this problem it was possible to select the positions of light sources in the system in order to achieve the best possible performance. The number of different position combinations for four sources is 24. Additionally, Problem B had two subproblems: same plane of incidence and orthogonal plane of incidence [19], increasing the number of design problems up to 48. For the case of the same plane of incidence, a simple analysis of the target requirements allowed the conclusion that the position of the source with the longest peak wavelength at the direct path (position A [19]) is better for phase difference compensation, since this source has minimal relative bandwidth. The optimal
Fig. 12. Overview of the results for the same plane of incidence light mixing system subproblem. The MF of the winning result is shown by the red bar. Colors are different for the different selections of the light source at the direct path position. Source positions are shown by labels from bottom (direct path position A) to top (position D).

Fig. 13. Overview of the results for the orthogonal plane of incidence light mixing system subproblem. The MF of the winning result is shown by the red bar. Colors are different for the different selections of the light source at the direct path position. Source positions are shown by labels from bottom (direct path position A) to top (position D).

positions of other light sources, even for this variant, are less trivial and require deeper analysis. For the orthogonal plane of incidence, the choice of optimal positions is even more complicated, since in this case phase difference compensation is better due to swapping $s$ and $p$ polarizations inside the system for some optical paths.

A fully automatic version of the deep search gradual evolution method appeared to be very efficient for this problem. In all cases, the iterations started with very thin single layer “seed” designs (106 nm with refractive index 2.35). It allowed us to run massive hands-free computational experiments for all 24 light source positions and for the both subproblems.

Figure 12 shows the results for the same plane of incidence case. These results confirm the conclusion that the position of the light source with the highest peak wavelength 660 nm at the direct path is the best option. All six corresponding MF
values are significantly lower than the rest of the MFs for other cases. Nevertheless, the selection of other source positions is nontrivial. It appeared that the variant with 660, 440, 515, and 590 nm sources at positions A, B, C, and D is the best one, and it won the contest in this subproblem case.

The results for the orthogonal plane of incidence case are shown in Fig. 13. There are two cases of light source positions that are significantly better than others. The best one is the winning result of this subproblem with MF of 0.0827 and light sources 660, 515, 440, and 590 nm distributed at positions A, B, C, and D. The second result is very close to this (MF value 0.00911); it is also better than the rest of the results submitted to the contest. This result has a completely different arrangement of light sources: they are distributed at positions A, B, C, D as 440, 660, 515, and 590 nm.

5. CONCLUSIONS

New deep search versions of the needle optimization, gradual evolution, and design cleaner methods are developed and applied to a wide set of problems. The comparison with conventional versions of these methods shows similar performance in the case of problems with dielectric layers having target specifications for transmittance and reflectance. For the problems with absorbing layers like metal, metal-island films, and conductive layers and for problems with phase and phase-related requirements (GD, GDD), the new methods show superior performance and provide top-quality results. This is confirmed by the results of the OIC 2019 Design Contest, where the application of deep search methods showed top winning results.

Funding. Munich-Centre for Advanced Photonics.

Disclosures. The author declares no conflicts of interest.

REFERENCES