Review of iterative Fourier-transform algorithms for beam shaping applications

Olivier Ripoll
Sony International Europe GmbH
Visual Technology Department
Hedelfinger Strasse 61
D-70327 Stuttgart, Germany
E-mail: ripoll@sony.de

Ville Kettunen
Heptagon Oy, Rüschlikon/Zürich Branch
Moosstrasse 2
CH-8803 Rüschlikon, Switzerland

Hans Peter Herzig
University of Neuchâtel
Institute of Microtechnology
rue A.-L. Breguet 2
CH-2000 Neuchatel, Switzerland

Abstract. We present a comparison of some of the most used iterative Fourier transform algorithms (IFTA) for the design of continuous and multilevel diffractive optical elements (DOE). Our aim is to provide optical engineers with advice for choosing the most suited algorithm with respect to the task. We tackle mainly the beam-shaping and the beam-splitting problems, where the desired light distributions are almost binary. We compare four recent algorithms, together with the historical error-reduction and input-output methods. We conclude that three of these algorithms are interesting for continuous-phase kinoforms, and two, namely the three-step method proposed by Wyrowski and the overcompensation of Prongué, still perform well with multilevel- and binary-phase DOE. © 2004 Society of Photo-Optical Instrumentation Engineers.
[DOI: 10.1117/1.1804543]

1 Introduction

Since the invention of digital holography, many algorithms have been proposed to solve the problem of designing an element to transform a given light distribution into another desired light distribution.[1] An interesting candidate for this challenge is the iterative Fourier transform algorithm (IFTA). This term was introduced to characterize a family of algorithms that bounce back and forth between two spaces related by a Fourier transform.[2–4] The first IFTA, the error-reduction algorithm, was proposed for digital holography in the early seventies.[5] Gerchberg and Saxton adapted it for phase retrieval problems,[6] and the error-reduction algorithm is consequently also referred to as the Gerchberg-Saxton algorithm in the literature. The error-reduction algorithm was extensively studied by Fienup who introduced the input-output class of algorithms to speed up convergence.[7,8] However, both the error-reduction and the input-output algorithms were not suited for elements with discrete phase values (binary or multilevel DOE). For the design of phase-only computer-generated holograms (CGH) such as the kinoform, various improvements to IFTA were proposed in the last 20 years.[3,4,9–15] As a consequence, the optical engineer facing a design problem now has the delicate task of choosing the most suited variant of the IFTA. Algorithms with more parameters sometimes allow an improved convergence speedwise, but mostly allow avoiding being trapped into local minima. However, this wider choice of parameters is also synonymous to added complexity for the optical engineer. The designer needs more experience to master convergence and to pick the best-suited algorithm instead of the most primitive ones such as input-output and error-reduction.[16] Consequently, it is important to identify which added complexity brings a real improvement. We aim at giving some advice for this choice, based on our experience with several of the variations of the IFTA.

Among the several uses of the CGH, three are of particular interest in digital holography: beam shaping, beam splitting, and pattern or image generation. Although they are three aspects of the same design problem, they differ by the merit functions used to characterize their performance. For example, image generation will not in general require the highest possible light throughput from the optical element, while beam shaping tasks will. Similarly, beam splitting is usually used to divide a light beam into several sub-beams of equal power while image generation requires pixels of different intensity levels. This article aims at comparing various algorithms for beam shaping. Beam shaping is most commonly used in high-energy laser applications, such as laser branding or photolithographic illumination.[17,18] These applications often require minimal energy losses, implying the use of phase-only elements, such as lenses and kinoforms.

In Sec. 2, we describe the various algorithms we have tested. The criteria we use to quantify their relative performance and the beam shaping design problem we utilize are tackled in Sec. 3. Then, in Sec. 4, we present the results comparing the raw performance of the algorithms. Finally, Sec. 5 furthers the discussion for each algorithm, considering ease of use, speed of convergence, and stagnation.

2 Description of the Algorithms

The first IFTA variant, the error-reduction algorithm, is well known and will be used here as a reference. Its principle is explained in Fig. 1. At iteration k, the complex amplitude of the light field in the signal plane A_k is back-propagated to the CGH plane, leading to the complex amplitude a_k. The CGH constraints are then applied to pro-
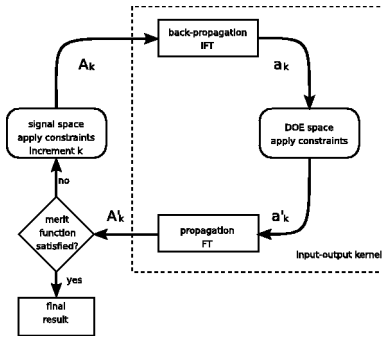
Fig. 1 The principle of the iterative Fourier transform algorithm. The dashed box denotes the input-output kernel.
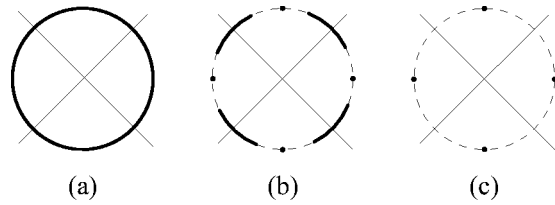

Fig. 2 Soft quantization on four levels of phase in the complex plane. At the beginning, the phases are spread over the entire angular interval of 2π (a). They are progressively projected onto the four phase levels (b) until the quantization process is ended (c).

duce a new amplitude a'_k. CGH constraints depend on the fabrication technology. In general, for kinoforms, the modulus of the amplitude $|a'_k|$ is identified to the modulus of the incident illumination. Additionally, for multilevel and binary elements, the phase of a_k has to be quantized. The new amplitude a'_k is propagated to the signal plane leading to a complex amplitude A'_k. The goal to achieve being A_{goal}, the signal plane constraints are now enforced, i.e., $|A'_k|$ is replaced by A_{goal} to produce A_{k+1}, which is the start of a new iteration. Fienup has shown that, for continuous phase elements, this algorithm succeeded in reducing the errors at every iteration in their respective plane, hence the name error-reduction.[7]

If the CGH constraints are enforced strictly, the error in the CGH space is not relevant, and only the error in the resulting signal is of interest. The two propagation steps and the application of the CGH constraints can be grouped in the input-output algorithm kernel,[7,8] symbolized by the dashed box in Fig. 1. Among the variants of the input-output algorithm, we have chosen the one given by Eqs. (9) and (10) of Ref. 7. Fienup called this variant output-output in a later article.[8] The signal orders are changed from iteration k to $k+1$ according to

$$A_{k+1} = A'_k + \beta \Delta A_{\text{driving},k} \tag{1}$$

with

$$\Delta A_{\text{driving},k} = |A_{\text{goal}}| \{2 \exp[i \cdot \arg(A'_k)] \\ - \exp[i \cdot \arg(A_k)]\} - A_k, \tag{2}$$

where β is a free parameter, usually chosen close to one. The input-output family is specially designed to give high efficiency and improved convergence, speedwise. The two other algorithms of the family are the input-output and the hybrid input-output, which will be mentioned in Sec. 5.

The two previous algorithms are known to give poor performance when the CGH is constrained to discrete phase levels. They have a tendency to converge to the nearest local optimum, failing to find the desired global optimum. To overcome this failure, many improvements have been introduced. Studies have been carried out on the role of design freedoms in the performance of the CGH.[9,19] If the desired light distribution is only constrained in a limited region of space, the value of the complex amplitude A_k

outside this area can be arbitrarily chosen by the designer. This possibility is referred to as amplitude freedom. Furthermore, in most applications, the phase of the generated light distribution can also be chosen arbitrarily by the designer. This second freedom is called phase freedom. Finally, in some applications where the fidelity of the generated signal matters more than the level of light throughput, a scale factor can be utilized by the designer to adjust the power repartition between the signal and noise areas. This scale factor freedom is the basis of the tuning of the algorithms by a technique called over-tensing, which will be covered at the end of this section.

The three freedoms presented above are not specific to discrete phase design, but happen to be more necessary in this case. A most noticeable additional improvement specific to discrete phase design is the soft-quantization concept proposed by Wyrowski.[4] Soft-quantization consists, as shown in Fig. 2, in quantifying progressively the phase values of a_k in the element plane. At a given iteration, only the points whose phase is contained inside intervals centered around the discrete phase levels are quantized. The intervals are progressively enlarged until all the phases are covered and the CGH is treated. Most of the time, this algorithm is not used alone, but as the final step of a more general scheme that we will refer to here as the three-step algorithm. The complete design process starts by generating a continuous CGH where all the power is constrained to be in the signal (use of phase freedom only). Then, the amplitude freedom is introduced, allowing the signal quality to increase at the expense of the apparition of noise outside the signal. This results in an optimized continuous CGH. Then only, the soft-quantization process is applied to obtain the discrete phase CGH. A recent variant was introduced in which the soft quantization is partially performed, and repeated several times, the last time leading always to a complete quantization.[14] This variant introduces new parameters to adjust the convergence of the IFTA.

A fourth algorithm of interest is the over-compensation algorithm proposed by Prongué for continuous beam-splitting elements.[10] Although originally introduced as a final step to optimize further results obtained by the simplex downhill method, over-compensation can be used alone with good results.[20] The amplitudes A'_k in the signal window of the output plane are modified for A_{k+1} according to

$$|A_{k+1}| = |A_k| \frac{\langle |A'_k| \rangle}{|A'_k|}, \tag{3}$$

where $\langle|A_k'|\rangle$ is the average of the $|A_k'|$. Thus, the low power orders are set to higher values and the high power orders are attenuated. This technique requires taking care of the conservation of the power, by re-normalizing the total power at every iteration. Also, one has to avoid the cases where the denominator tends toward zero and implies divergence of the corresponding value of $|A_{k+1}|$ in Eq. (3). It has to be noted that the amplitudes in Eq. (3) can be replaced by intensities, without changing much the convergence properties of the algorithm.[20] This algorithm is restricted to binary signal distributions, but can be easily extended to general distributions, by changing Eq. (3) to

$$|A_{k+1}| = |A_k| \frac{|A_{\text{goal}}|}{|A_k'|}. \tag{4}$$

Over-compensation was also proposed simultaneously by Farn[21] and has been rediscovered recently by Liu.[15]

Another variation of the IFTA worth of interest was proposed by Arrizón.[12] The author originally introduced it for para-geometrical and continuous solutions, i.e., continuous profiles that are close to solutions designed by means of geometrical optics and whose phase can be unwrapped. However, we have found it interesting to include it in our tests, even with discrete phase levels, and random start phase distributions. It is characterized by the possibility to reduce the desired energy within the signal, i.e., to use the scale freedom. Whenever the uniformity is not improving enough during an iteration, the efficiency goal is slightly decreased, allowing the uniformity error to be lowered. The algorithm stops when a given uniformity goal is reached, or in the case of stagnation. We have chosen in our test to stop when a uniformity error of 0.5% was reached.

Lastly, we studied another recent algorithm proposed by Johansson for continuous profile CGH, called up-scaling.[22,23] This algorithm is based on a concept similar to over-compensation. Two real-valued thresholds are defined in the signal space, a lower $A_{\min}$ and a higher $A_{\max}$, around the desired real signal value A_{goal}. The signal orders are changed according to

$$|A_{k+1}| = \begin{cases} A_{\max} & \text{if } |A_k'| \leq A_{\min} \\ A_{\min} & \text{if } |A_k'| \geq A_{\max}. \\ 2A_{\text{goal}} - |A_k'| & \text{otherwise} \end{cases} \tag{5}$$

The noise orders are not modified in the original algorithm.

IFTA as described here does not seem to be tunable, with the exceptions of the variations proposed by Arrizón and Fienup. However, the designer usually utilizes a technique called over-tensing, which can bring some flexibility to error-reduction and its derivatives. This technique exploits the scale factor freedom mentioned earlier. Instead of replacing the signal amplitude A_k' by A_{goal} to generate A_{k+1}, the signal orders are replaced by γA_{goal} where γ is a scalar. If γ is superior to one, the IFTA is said to be overtensed, and if inferior to one, under-tensed. We have used this technique in order to make the error-reduction, the upscaling, and the three-step algorithms adjustable. For the input-output, we have used the β parameter to tune the
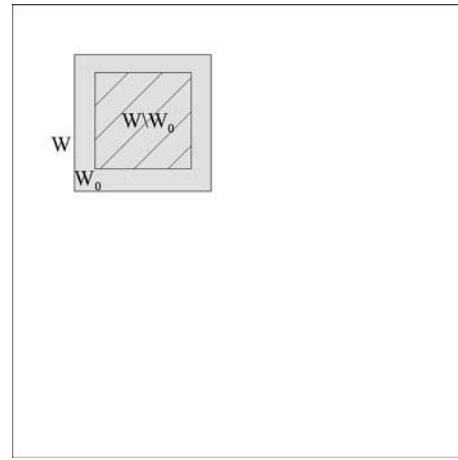

Fig. 3 The signal window W (in gray) defines the area of the output space where the constraints are applied. For beam-shaping purposes, it is often decomposed in a zero frame W_0 and a non-zero signal area $W \backslash W_0$ (dashed).

convergence. Over-compensation was not tuned, but we retained the configurations of best efficiency and best uniformity.

3 Test Pattern and Performance Criteria

The study of the various algorithms has been realized for a very simple situation, the shaping of a uniform wave into a square light distribution of uniform intensity. This corresponds to a pattern of $N_{\text{orders}} = 15 \times 15$ orders. We have reduced the study to a small CGH composed of only 128×128 points per unit cell. The start phase of the algorithms was random, but the same for every computation to allow a relevant comparison. We had the option to impose a frame of zeros around the square pattern. In this case, the signal window W contained a sub-window W_0, as show in Fig. 3. The results we present here were obtained without zero frame, but we tackle the topic in Sec. 5. Since efficiency is an important requirement of beam-shaping tasks, the element was a kinoform, i.e., a phase-only element. The task was to design such an element with continuous, eight-level, and binary profiles. Although we chose this problem with beam shaping in mind, this also corresponds to a classical beam splitting situation, where the incoming light is divided into 225 spots of equal intensity, located on a 15×15 grid of points. Consequently, this study is also valid for beam splitting applications.

Different criteria are usually encountered in the literature for measuring the performance of illumination designs. At a given iteration k, the generated complex amplitude is composed of pixels that we denote by the index j. For the sake of simplicity, we omit the iteration number and note the complex amplitude A_j. The most interesting for beam shaping is often the efficiency, the ratio

$$\eta = \frac{\sum_{j \in W \backslash W_0} |A_j|^2}{\sum_j |A_j|^2} \tag{6}$$

of light in the signal window $W \backslash W_0$ to the total light in the output space. The intensity present in the zero frame W_0

leads to a parasitic noise that cannot be accounted in the efficiency and is measured by the zero noise

$$Z = \frac{\Sigma_{j \in W_0} |A_j|^2}{\Sigma_j |A_j|^2}. \tag{7}$$

For beam-splitting elements, which are often designed to generate equally intense orders, the uniformity of the set of orders is often measured by its min-to-max uniformity error

$$U = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \tag{8}$$

where $I_j = |A_j|^2$ and $I_{min} = \min(I_j)$ and $I_{max} = \max(I_j)$ are the intensities of the weakest and the strongest orders of diffractions respectively. This error measure is also commonly used in general illumination design. Another measure of the uniformity is the normalized standard deviation of the intensity of the diffraction orders in the signal

$$\sigma = \frac{[\Sigma_{j \in W \backslash W_0} (I_j - I_{mean})^2]^{1/2}}{\Sigma_{j \in W \backslash W_0} I_j} \tag{9}$$

where I_{mean} is the mean of the signal intensities in $W \backslash W_0$.

Additionally, in the literature, one can find other criteria such as the signal-to-noise ratio (SNR) or the mean squared error E. However, they are mostly useful in design of CGH for image generation producing light distributions with several levels of intensity. We will thus not use them in our study. For beam shaping elements generating a binary intensity pattern such as the ones we are interested in, we prefer the efficiency, uniformity error, and the zero noise.

In general, a merit function can be built up from the previous formulas for any specific design task. When compared to direct optimization techniques, such as simulated annealing[24] or direct binary search,[25] it is noticeable that IFTA cannot use the merit function to accept or reject a change in the optical element, and consequently cannot use it to reliably control the convergence. The merit function is mostly a measure of the evolution of the IFTA process. Attempts have however been made to influence the algorithm based on the merit function when it is composed of antagonistic criteria, such as efficiency and uniformity error, or efficiency and SNR.[11,13,26] In this situation, one can approximately relate the value of the merit function to the value of the over-tensing factor γ. These algorithms are however difficult to master and reliable only for some specific merit functions. Additionally and more simply, it is possible to monitor the merit function during the optimization process and to retain the configuration of the best iteration, which can be different from the ultimate iteration. Of course, this strategy is only applicable with algorithms where the complete set of constraints is satisfied at every iteration, which is not the case for the three-step algorithm because of the soft-quantization nature.

If the result of the IFTA is not satisfying, one can improve the process by chaining various algorithms. One can associate several IFTA such as in the three-step method[4,9] or the recent variation of the over-compensation by Liu.[15] One can also associate IFTA to another family of algo-rithms, such as downhill simplex[10] or direct binary search (DBS).[27] However, in all these cases, the IFTA step influences strongly the final results, and the conclusions of this study are still relevant. On the other hand, combining different families of algorithms requires more experience from the engineer. Combinations limited to IFTA are thus more interesting from an engineering point of view.

4 Results of the Various Algorithms

We now tackle the quantitative results obtained for the design problem described in Sec. 3 with the algorithms presented in Sec. 2. The efficiency η is plotted versus the uniformity error U for the resulting optimized solution. For each algorithm, the different points correspond to different values of the tuning parameter, most often the over-tensing factor γ. The figures illustrate the overall performances of the algorithms and their flexibility. In such a plot, the data corresponding to the best results are located in the top-left corner.

With respect to the conditions of computation, two points have to be stated. First, as previously mentioned, every optimization process started from the same random configuration. Second, we ran the same number of iterations (1200) for each algorithm. This number, quite large for IFTA, was chosen to ensure that every algorithm would reach a stable state. An exception was made when stagnation was encountered, and the algorithm was exited to avoid useless computations. For the three-step algorithm, we divided the total number of iterations equally between the three steps.

4.1 Continuous Phase

The resulting efficiency and uniformity error for a continuous phase CGH are presented in Fig. 4. As seen, most of the algorithms perform well. It is noticeable that since soft-quantization is not used, the second and third stages of the three-step algorithm are actually equivalent to error-reduction. However, the presence of the first phase-synthesis stage improves significantly the performance of the algorithm. This emphasizes the role of choosing or building a good start distribution for IFTA, which the first stage realizes.

A further look at the curves shows that three algorithms are performing very well. Over-compensation, the three-step algorithm, and the variation by Arrizón are reaching high efficiency with low uniformity error. Concerning the last one, it should be noted that the points are located around $U = 0.5\%$. This is due to our design choice, which was, as stated in Sec. 2, to target 0.5% for this criterion. Up-scaling, three-step method, and error-reduction appear to be very tunable, with the presence of a clear trade-off between efficiency and uniformity error.

4.2 Eight-Level Phase

Figure 5 presents the plots of efficiency versus uniformity error for the design of an eight-level element. The difference with the continuous case is obvious for most algorithms. Both the algorithms aimed at continuous phase design (up-scaling and Arrizón's variant) and the first-generation algorithms (output-output and error-reduction)
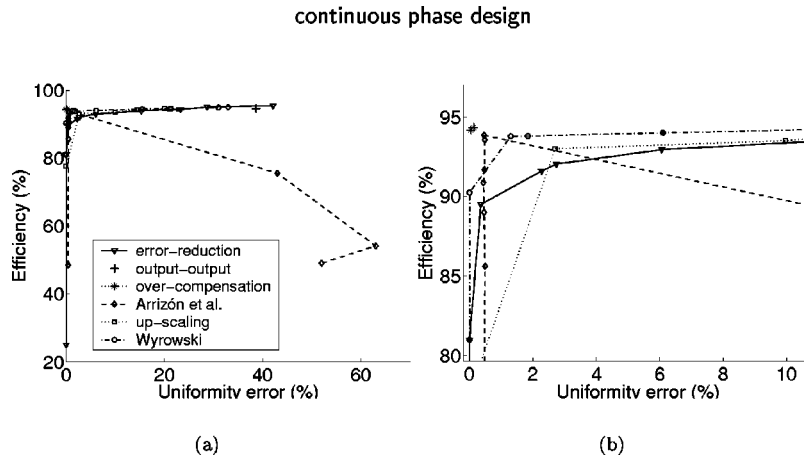
continuous phase design

Fig. 4 Comparison of efficiency and uniformity error of a continuous profile design with various algorithms; (b) is a magnification of (a).

fail in optimizing the uniformity error. We may, however, note that all the algorithms still perform well from the efficiency point of view.

Two algorithms are now clearly apart: over-compensation and three-step with soft-quantization. We will not draw conclusions from this curve about their relative performance. The difference is not significant, and we have observed opposite results for a different problem.[28] However, we observe that the three-step algorithm allows us to balance uniformity versus efficiency, which can be an interesting feature for the optical designer. The conclusion is that they both out-perform the other algorithms, and achieve similar results.

4.3 Binary Phase

With drastically reduced design freedoms, the convergence of the optimization process is no longer straightforward. The multiplicity of local optima usually traps the process, and most of the algorithms stagnate after a few iterations. This is obvious from Fig. 6, where the gap between over-compensation and the other algorithms has increased. Among the not-optimal algorithms, the three-step one still presents the option to allow either good efficiency or good

uniformity by sacrificing the other criterion. Additionally, up-scaling now seems to be the second-best algorithm for overall performance.

An interesting property can be observed in Fig. 6 and to a lesser extent also in Fig. 5. The algorithms using the scale freedom exhibit curves with chaotic behaviors, especially when compared to the regularity seen for continuous profiles. This illustrates the fact that modifying the scale factor (over-tensing and under-tensing) is not a completely reliable parameter to tune the IFTA, as previously mentioned in Sec. 3.

5 Discussion

In this section, we discuss in more details the results for each algorithm, and try to take into account the ease of use and flexibility in our appreciation. Most of the conclusions found here are not new and have already been stated in the literature. We merely want to reassert them with respect to our test.

5.1 General Observations

From the previous curves, two general conclusions can be drawn. First, the role of the start phase is important to reach
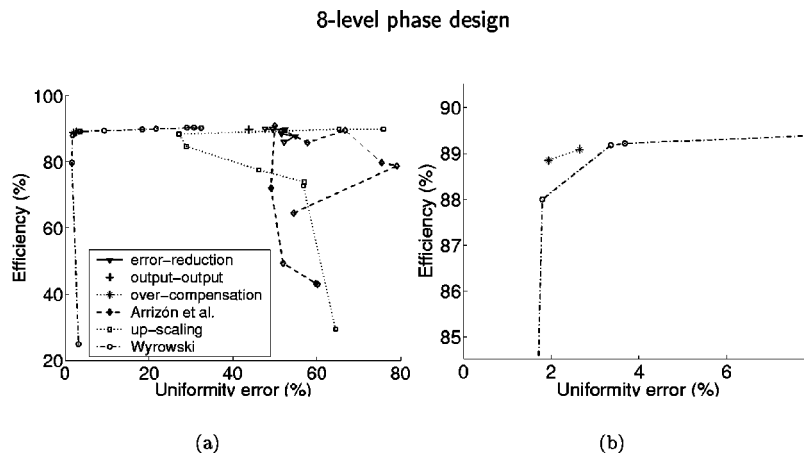
8-level phase design

Fig. 5 Comparison of efficiency and uniformity error of an 8-level profile design with various algorithms; (b) is a magnification of (a).
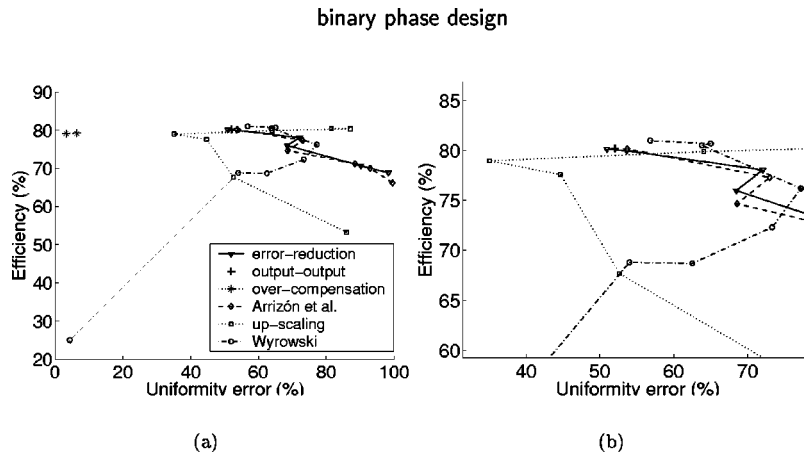
binary phase design

Fig. 6 Comparison of efficiency and uniformity error of a binary profile design with various algorithms; (b) is a magnification of (a).

5.2 Input-Output Family

As we can see from the simulations, choosing the parameter β in Eq. (1) to tune the performance of the output-output algorithm is not effective. Most of the results are similar, with high efficiency and poor uniformity, with values of β ranging from 0.5 up to 500. Also, the results are similar for the three types of DOE, which suggests that the algorithm is not able to take advantage of the additional design freedoms of the continuous profile.

As mentioned in Sec. 2, we chose the variation called output-output. The reason of this choice is illustrated in Fig. 7, where we compare the performance of the three variations. While the overall behavior is the same, the output-output exhibits a slightly better efficiency in the three tests.

Input-output algorithms are usually good for speed of convergence, and were an appreciable improvement when introduced in the early eighties. However, they should not be seriously considered anymore for design tasks.

a good solution. It is responsible for the difference between error-reduction and the three-step algorithm for continuous phase. In practice, the designer will benefit from the possibility to launch several optimization processes with different start phases. He will then compare the results and choose the best configuration. Building a good start phase is also a possibility. Para-geometrical solutions, as seen in Arrizón's original article, usually result in a higher efficiency.[12,28]

Second, the use of the scale factor to tune the efficiency-uniformity trade-off is very effective. For phase elements, since the amount of energy is conserved, this scale factor is easily integrated in the algorithm by the over-tensing technique. In general, over-tensing implies increasing efficiency, while under-tensing means increasing uniformity. The three algorithms that use this technique, namely error-reduction, three-step technique, and up-scaling, are the algorithms presenting the highest variability of efficiency and uniformity in our test. Additionally, an advantage of the over-tensing is that it allows better convergence for multi-level structures.

5.3 Over-Compensation

Over-compensation has constantly shown a remarkable performance in this test. It is noticeable, however, that this technique has not often been used in practice. We attribute this mainly to lack of knowledge, due to the low number of articles where it was described. Nevertheless, over-compensation exhibits some drawbacks that are not outlined by the test we have performed here.

First, there are design situations where over-compensation does not achieve high performance. Indeed, Eq. (3) diverges when the value of the denominator is close to zero. Consequently, over-compensation is often not efficient at imposing zero or low-intensity values inside the signal window. Thus, over-compensation is well suited for flat-top beam shaping and beam splitting, but not for image generation. Also, it is not optimal for shaping a beam to a distribution with areas of low energy such as a Gaussian profile.

Furthermore, over-compensation is not good at imposing windows of zero energy, such as the one illustrated in Fig. 3. The zero noise N is usually stronger for this algorithm, especially when compared to the three-step algorithm.
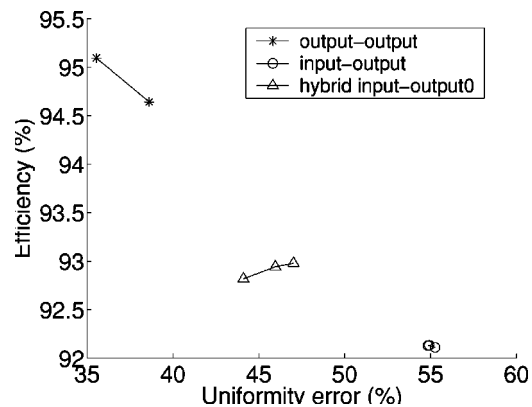

Fig. 7 Comparison of the results of the three variants of Fienup's algorithm, namely input-output, output-output, and hybrid output-output.

Lastly, over-compensation is insensitive to the use of over-tensing, and we still have to see a parameter that can allow easy tuning. This restricts the freedom of the optical engineer. Since with respect to uniformity the final iteration is not guaranteed to be the best iteration, we recommend the use of a merit function during the optimization process, in order to retain the best iteration of all.

5.4 IFTA Variation by Arrizón

The conclusion drawn from the test is that this algorithm is only to be used for continuous phase, which is the field it was proposed for. Nevertheless, our test shows that in addition to the para-geometric start phase of the original article, it works also with a random start phase. We think that it is worth implementing this algorithm, since it gives a simple control to the designer on the evolution of the optimization process. The central idea of the algorithm is that since uniformity and efficiency are competing, and since over-tensing the IFTA increases efficiency at the expense of uniformity (and vice versa), the over-tensing is decreased progressively until acceptable uniformity is reached. This idea could be applied to other algorithms, at the condition that they are affected by over-tensing.

5.5 Three-Step with Soft-Quantization

This algorithm exhibits very interesting capabilities. While it is comparable to the over-compensation in performance, it can be applied for a much wider range of designs, such as image generation, since it is not sensitive to the presence of signal orders of low value.

On the other hand, a typical drawback of this technique is the presence of isolated pixels in the CGH pattern. Application of filters such as the median filter has been proposed to overcome this issue.[29] The over-compensation, on the other hand, seems to be naturally immune to this defect, and generally generates more symmetrical patterns.

We have mentioned in Sec. 2 the existence of a recent variation of the soft-quantization step, proposed in the field of image generation. We have tested this algorithm in our beam shaping problem, and have not found it to produce significantly better CGH. Since it introduces two supplementary parameters with no straightforward physical signification and thus requires much more experience, we advise optical engineers to consider the three-step algorithm with soft-quantization as a better candidate for beam shaping and beam splitting problems. The comparison for image generation is the subject of Ref. 14.

Finally we would like to point out that our experience is that the three-step algorithm performs very well with respect to the noise N of Eq. (7) when a window of zeros W_0 is imposed, as illustrated in Fig. 3.

6 Conclusions

We have compared various variations of the iterative Fourier transform algorithm in order to give the optical engineer advice for the choice of the most suited procedure to design phase-only computer generated holograms. Three algorithms have shown good performances for continuous phase elements, namely Arrizón's variant, three-step with soft-quantization, and over-compensation. The last two are also the only ones that exhibit good performances when the

phase is quantized. We advise the optical engineer to avoid outdated algorithms and to concentrate on these three variants in order to have fast and effective design processes when aiming at beam shaping and beam splitting.

Additionally, we have stressed the need to take some care in the choice of the start point of the IFTA, and have shown the possibility to adjust the variants, which is an important parameter for the designer.

It must be noted that the problem we have used exhibited a low degree of design freedom, due to the small number of pixels used. In practice, the designer should increase the number of pixels for such a design to obtain better results. We chose this situation as it outlines the capability of the algorithms to perform well with drastic design constraints.

References

1. V. A. Soifer, V. V. Kotlyar, and L. L. Doskolovich, *Iterative Methods for Diffractive Optical Elements Computation*, Taylor & Francis, London (1997).
2. J. R. Fienup and C. C. Wackerman, "Phase-retrieval stagnation problems and solutions," *J. Opt. Soc. Am. A* **3**(11), 1897–1907 (1986).
3. F. Wyrowski and O. Bryndahl, "Iterative Fourier-transform algorithm applied to computer holography," *J. Opt. Soc. Am. A* **5**(7), 1058–1065 (1988).
4. F. Wyrowski, "Iterative quantization of amplitude holograms," *Appl. Opt.* **28**(18), 3864–3870 (1989).
5. P. M. Hirsch, J. A. Jordan, and B. J. Lesem, "Method of making an object independent diffuser," U.S. Patent No. 3,619,022 (1971).
6. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image diffraction plane pictures," *Optik (Stuttgart)* **35**, 237–246 (1972).
7. J. R. Fienup, "Iterative method applied to image reconstruction and to computer-generated holograms," *Opt. Eng.* **19**(3), 297–305 (1980).
8. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**(15), 2758–2769 (1982).
9. F. Wyrowski, "Diffraction efficiency of analog and quantized digital amplitude holograms: analysis and manipulation," *J. Opt. Soc. Am. A* **7**(3), 383–393 (1990).
10. D. Prongué, H. P. Herzig, R. Dändliker, and M. T. Gale, "Optimized kinoform structures for highly efficient fan-out elements," *Appl. Opt.* **31**(26), 5706–5711 (1992).
11. F. Fetthauer, C. Stroot, and O. Bryngdahl, "On the quantization of holograms with the iterative Fourier transform algorithm," *Opt. Commun.* **136**, 7–10 (1997).
12. V. Arrizón, M. Testorf, S. Sinzinger, and J. Jahns, "Iterative optimization of phase-only diffractive optical elements based on a lenslet array," *J. Opt. Soc. Am. A* **17**(12), 2157–2164 (2000).
13. S. Bühling and F. Wyrowski, "Improved transmission design algorithms by utilizing variable-strength projections," *J. Mod. Opt.* **49**(11), 1871–1892 (2002).
14. M. Skeren, I. Richter, and P. Fiala, "Iterative Fourier transform algorithm: comparison of various approaches," *J. Mod. Opt.* **49**(11), 1851–1870 (2002).
15. J. S. Liu and M. R. Taghizadeh, "Iterative algorithm for the design of diffractive phase elements for laser beam shaping," *Opt. Lett.* **27**(16), 1463–1465 (2002).
16. J. Yao, G. C. K. Chen, and T.-K. Lim, "Holographic diffuser for diffuse infrared wireless home networking," *Opt. Eng.* **42**(2), 317–324 (2003).
17. J. R. Leger, "Laser beam shaping," Chap. 9 in *Micro-optics, Elements; Elements, Systems and Applications*, H. P. Herzig, Ed., pp. 223–257, Taylor and Francis, London (1997).
18. F. M. Dickey and S. C. Holswade, *Laser Beam Shaping; Theory and Techniques*, Marcel Dekker, New York (2000).
19. F. Wyrowski, "Diffractive optical elements: iterative calculation of quantized, blazed phase structures," *J. Opt. Soc. Am. A* **7**(6), 961–969 (1990).
20. A. Schilling, "Diffractive and refractive optical microstructures: theory, design and applications," PhD Thesis, University of Neuchâtel (2000).
21. J. Bengtsson, "Kinoform design with an optimal-rotation-angle method," *Appl. Opt.* **33**(29), 6879–6884 (1994).

22. M. Johansson and J. Bengtsson, ''Robust design method for highly efficient beam-shaping diffractive optical elements using an iterative-Fourier-transform algorithm with soft operations,'' *J. Mod. Opt.* **47**(8), 1385–1398 (2000).
23. D. Schäfer, ''Design concept for diffractive elements shaping partially coherent laser beams,'' *J. Opt. Soc. Am. A* **18**(11), 2915–2922 (2001).
24. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, ''Optimization by simulated annealing,'' *Science* **220**(4598), 671–680 (1983).
25. M. A. Seldowitz, J. P. Allebach, and D. W. Sweeney, ''Synthesis of digital holograms by direct binary search,'' *Appl. Opt.* **26**(14), 2788–2798 (1987).
26. L. Bigué and P. Ambs, ''Optimal multicriteria approach to the iterative Fourier transform algorithm,'' *Appl. Opt.* **40**(32), 5886–5893 (2001).
27. L. Legeard, P. Réfrégier, and P. Ambs, ''Multicriteria optimality for iterative encoding of computer-generated holograms,'' *Appl. Opt.* **36**(29), 7444–7449 (1997).
28. O. Ripoll, ''Far-field beam shaping elements for deep UV lithography,'' PhD Thesis, University of Neuchâtel (2003).
29. H. Schwarzer, S. Teiwes, and F. Wyrowski, ''Non-pixelated design of computer-generated diffractive elements for increased diffraction efficiency,'' in *Diffractive Optics 97*, J. Turunen and F. Wyrowski, Eds., pp. 164–165, EOS, Savonlinna (1997).

Olivier Ripoll received his diploma in optics and photonics from the Institut d'Optique in Orsay, France, in 1997. In 1998, he became a research assistant in the Applied Optics Group at the Institute of Microtechnology of the University of Neuchâtel, Switzerland. He worked in the design and analysis of diffractive optical elements for beam shaping. He received his PhD degree in 2002. Since 2002, he has worked as an R&D engineer for Sony International Europe, Stuttgart, Germany. His research interests include diffractive optics design and illumination systems design. He is member of the French Optical Society (SFO) and the European Optical Society (EOS).

Ville Kettunen received his master's degree in physics and his PhD degree in micro-optics from the University of Joensuu, Finland, in 1995 and 1998 respectively. From 1998 to 2000 he was a senior researcher in the diffractive optics group at the University of Joensuu working in the field of micro-optics design and modeling. In 2000 he joined the Applied Optics Group at the Institute of Microtechnology of the University of Neuchâtel, Switzerland, as a senior scientist. Since 2002 he has been the senior optical engineer at the Swiss-Finnish company Heptagon Oy in Zürich, Switzerland. His current research interests include design of diffractive optical elements and novel applications of refractive and diffractive micro-optics on component and system levels.

Hans Peter Herzig received his diploma in physics from the Swiss Federal Institute of Technology in Zürich, Switzerland, in 1978. From 1978 to 1982 he was a scientist with the Optics Development Department of Kern in Aarau, Switzerland, working in lens design and optical testing. In 1983, he became a graduate research assistant with the Applied Optics Group at the Institute of Microtechnology of the University of Neuchâtel, Switzerland, working in the field of holographic optical elements, especially scanning elements. In 1987, he received his PhD degree in optics. From 1989 until 2001 he was head of the micro-optics research group at the University of Neuchâtel. Since 2002 he has been a full professor and head of the Applied Optics Laboratory. His current research interests include refractive and diffractive micro-optics, nano-scale optics, and optical MEMS. Dr. Herzig is a senior editor of the *Journal of Microlithography, Microfabrication, and Microsystems* (JM³) and a member of the editorial board of *Optical Review*. Since 2001 he has been secretary of the Swiss Society of Optics and Microscopy.