Published in
Pure and Applied Optics: Journal of the European Optical Society Part A 7, issue 3, 565-574, 1998
which should be used for any reference to this work

Direct sampling for diffractive microlens encoding from a rigorous point of view

A Schilling†, P Blattner and H P Herzig

Institute of Microtechnology, University of Neuchâtel, Rue A-L Breguet 2, CH-2000 Neuchâtel, Switzerland

Abstract. Rigorous diffraction theory is applied to analyse the direct-sampling (DS) encoding method, which is based on scalar diffraction theory. For given fabrication constraints and constant sampling width of the lens function, the quantized phase profiles obtained with scalar DS are close to the optimum solutions, even for grating period to wavelength ratios (Λ_b/λ) as small as about 3. For smaller ratios, the phase profiles obtained by DS can be improved by up to 25%, using a straightforward rigorous steepest-gradient optimization. Applied to cylindrical lenses with NA = 0.5 and 0.63, coding with DS and with rigorously improved DS gives quite similar results for the total diffraction efficiency.

1. Introduction

In modern optics, diffractive microlenses are appreciated because of the additional degrees of freedom they bring to the design compared with conventional refractive optics. They can be found, for example, in collimating and focusing optics, in optical diffusers, chromatic-aberration correction optics or in athermalized hybrid elements [1–3]. In general, the desired lens function of a diffractive micro-optical element is continuous and has to be quantized into a multilevel staircase-like phase profile. This is necessary for the manufacturing of the surface-relief elements by a standard lithographic process, where only discrete phase levels are possible. The number of discrete phase levels per 2π phase steps depends on the minimum feature size (MFS), which can be fabricated. For the coding of the continuous lens function into the multilevel profile, different approaches are possible. A simple, standard way of coding is analytical quantization (AQ) [4, 5]. However, this procedure gives satisfying results only as long as four or more discrete phase levels are possible per 2π phase difference of the lens function. For high-aperture lenses or short wavelengths, the AQ gives an efficiency of the lens which is significantly lower than the optimum value [6]. A more advanced method is the radially symmetric iterative discrete on-axis (RSIDO) coding developed by Welch *et al* [6, 7]. This method also produces nearly optimum results at the edge of the lens, where only two to three phase levels are possible per grating period (Λ_b). The major drawback of this method is the fact that it utilizes nonlinear techniques like simulated annealing for the optimization of the widths and locations of the quantized phase levels. It therefore consumes considerable computing power, especially for lenses with large diameters.

† E-mail address: Andreas.Schilling@imt.unine.ch

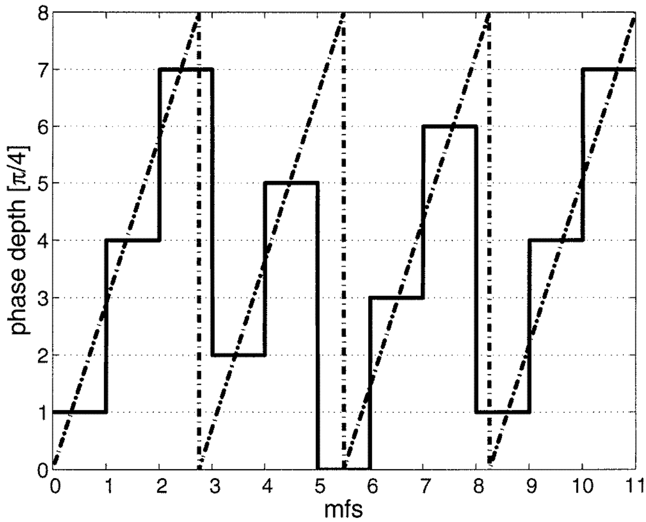

Figure 1. Illustration of the DS method for encoding: the ideal phase function of a linear blazed grating ($\Lambda_b/\mathrm{MFS} = 2.75$) is the chain curve, the full curve is the multilevel profile obtained with DS. The horizontal grid lines are the eight available phase levels.

Kuittinen and Herzig [8] described a different, more straightforward way for solving the coding problem, which they called direct sampling (DS). This method gives the optimum efficiency of a diffractive lens automatically, whereas the coding scheme convinces with its simplicity. With DS coding the complete lens function is sampled with the MFS and phase values are clipped between 0 and 2π. These phase values are then rounded to the closest available phase level. The DS method is illustrated in figure 1 for a blazed grating with $\Lambda_b/\mathrm{MFS} = 2.75$. This method is particularly suitable for fast, shaped-beam electron beam pattern generators, which only write rectangular pixels but the dose per pixel can be controlled to generate a multilevel depth profile.

In the outermost part of a diffractive high-aperture lens, only two or three phase levels are possible per 2π phase difference, because of the limitations of the MFS. We will call this part of the profile the binary region. Figure 1 shows a binary region, where 2.75 MFS fit into one grating period. While the performance of the different coding schemes is most critical in the binary region, this area can account for a large fraction of a high-aperture lens and therefore has a great impact on the total performance. For example, a spherical lens with $\mathrm{NA} = 0.32$ operating at $\lambda = 632.8$ nm has a binary area of 59% when a MFS of 1 μm is assumed.

In the binary region, assuming a MFS of 1 μm and the wavelength range of the visible spectrum, one is clearly outside the range of validity of scalar diffraction theory [9]. The DS method for encoding is based on simple scalar theory and so far has not been investigated rigorously. In this paper we present an analysis of the DS method, using exact electromagnetic grating theory, in order to investigate the performance of DS from a rigorous point of view. The local grating periods and therefore the Λ_b/λ ratios are smallest in the binary region for a given diffractive lens. This means that differences between scalar and exact electromagnetic theory are most likely to appear there.

The lens function can be considered locally as a linear blazed phase grating. Therefore, we analysed the quantization of ideally blazed surface relief phase gratings as local approximations for arbitrary two-dimensional lens functions.

2. Numerical methods

For the rigorous calculations we used a rigorous eigenmode method as described by Turunen [10]. This method provides an exact, non-iterative solution of Maxwell's equations and the accuracy of the solution depends only on the number of terms retained in the expansions of the electromagnetic fields. The implementation was modified according to Moharam *et al* [11] to avoid numerical instabilities which arise from the successive electromagnetic field matching at the boundaries between the individual grating layers. For the entire analysis, the diffraction efficiencies were calculated for a normally incident plane wave with TE polarization. The standard propagation direction was substrate to air and the refractive index of the substrate was 1.5.

In order to determine whether the quantized phase profiles obtained with DS are still the optimum solutions with the highest diffraction efficiencies, when analysed rigorously, we introduced a 'rigorous steepest-gradient optimization' (RSGO) method. The sampling of the lens function was hereby kept constant (1 MFS). The starting point for the optimization procedure was the phase-level distribution obtained with DS, such as, for example, illustrated in figure 1 by the full curve. Using this distribution, the first phase level was changed one level up ($\pi/4$ for eight level coding), or down, while all the others were kept constant, and rigorous diffraction efficiencies were calculated for the new distribution. The same was then performed with every other phase level until all levels had been changed once and corresponding efficiencies were calculated. The new phase-level distribution was then chosen as the one with the highest diffraction efficiency in the corresponding order. This new phase-level distribution was the starting point for the next iteration step. Iterations were executed until no further improvement of diffraction efficiency could be achieved.

Since the RSGO is a steepest-gradient optimization, it cannot be excluded that the algorithm ends in a local maximum. On the other hand, utilizing a nonlinear optimization technique like simulated annealing, which avoids stagnation in local maxima quite effectively, is inappropriate because of the time-consuming rigorous calculations. However, for some phase-level distributions, we used simulated annealing, and obtained the same maximum diffraction efficiencies as with RSGO.

3. Results and discussion

3.1. Rigorous analysis of DS blazed gratings in the binary region

When a blazed grating with a certain grating period, Λ_b, is DS encoded into a multilevel surface profile, the formal period of the grating increases, because, in general, the 2π phase steps of the blazed grating do not coincide with the DS steps of 1 MFS. This has to be considered for the calculation of diffraction efficiencies, since here the periodicity of the structure is of fundamental importance. The grating period of the DS blazed grating, Λ, is determined through $\Lambda = m\Lambda_\mathrm{b}$ with m obtained from the condition $m\Lambda_\mathrm{b} = n$ MFS, where m and n are the smallest non-zero integers. As illustrated in figure 1, the blazed grating with period $\Lambda_\mathrm{b} = 2.75$ MFS coincides with the DS steps of 1 MFS after $m = 4$ periods. The grating period of the DS-encoded grating is therefore 11 MFS. For a standard blazed grating the first-order diffraction efficiency is of concern. However, for the DS-encoded blazed gratings the grating period is increased by the above-mentioned factor m. As a consequence, the direction of the first order for the blazed grating corresponds to the mth order of the DS-encoded grating. Therefore, it is this order which has to be optimized in the calculations.

The rigorous analysis of the DS blazed grating was performed for eight different Λ_b/MFS ratios in the binary region, namely 2, 2.2, 2.25, 2.4, 2.5, 2.6, 2.75 and 2.8, for TE polarization. A representative part of the results is shown in figures 2(a)–(c). First, we calculated the rigorous diffraction efficiencies for the various Λ_b/MFS ratios by varying λ while keeping the grating period Λ_b constant (full curves). Then the RSGO was executed for a Λ_b/λ range of 1 to 7, again for fixed Λ_b (broken curves). For large Λ_b/λ ratios the diffraction efficiencies approached the scalar values (chain curves), as expected. However, the rigorous values are always a few per cent below the scalar values, since in scalar diffraction theory no Fresnel losses and reflected orders are taken into account. For Λ_b/λ between 2 and 3, the diffraction efficiency exhibits a strong decrease followed by a peak, as can be seen, for example, in figure 2(c). This behaviour can be understood qualitatively as an occurrence of total internal reflection at the interface [12].

The important criterion for the performance of the encoding method is the Λ_b/λ ratio until which the scalar-based coding method produces phase-level distributions which cannot significantly be improved with rigorous methods for given fabrication constraints. This ratio will be designated $(\Lambda_b/\lambda)_0$. From the analysis of our calculations in the binary region, we extracted $(\Lambda_b/\lambda)_0$ for the eight different Λ_b/MFS ratios. We determined $(\Lambda_b/\lambda)_0$ by the condition that for all Λ_b/λ ratios larger than $(\Lambda_b/\lambda)_0$ the rigorously optimized solutions show an improvement in diffraction efficiency of less than 3% over the solutions obtained with DS. The $(\Lambda_b/\lambda)_0$ ratios are given in table 1. The most important result of the analysis in the binary region is, that the values for $(\Lambda_b/\lambda)_0$ are between 2 and 3.3, which is surprisingly small. This means for the DS method, that it can be reliably applied to wavelengths of up to 0.85 MFS which corresponds to NA $= 0.43$. The rigorous diffraction efficiencies are somewhat smaller than the corresponding scalar values, the phase-level profiles are, however, under the given fabrication constraints and the constant sampling of the lens function, the best possible. These results show that simple DS encoding ensures optimum performance in the above sense over a very wide Λ_b/λ range, from the scalar regime ($\Lambda_b/\lambda > 5$) to deep into the rigorous regime ($\Lambda_b/\lambda \approx 2$–3).

Table 1. $(\Lambda_b/\lambda)_0$ values for different Λ_b/MFS ratios of the DS blazed grating in the binary region (eight equally spaced levels were available for 2π phase).

Λ_b/MFS	$(\Lambda_b/\lambda)_0$
2.0	2.0
2.2	2.2
2.25	2.3
2.4	2.8
2.5	2.3
2.6	2.9
2.75	3.3
2.8	3.3

In the regime $\Lambda_b/\lambda < (\Lambda_b/\lambda)_0$, where the DS quantized phase profiles have diffraction efficiencies as low as 11%, the rigorously optimized solutions (RSGO) show large improvements of up to 25%. In this regime the optimum phase profiles cannot be obtained

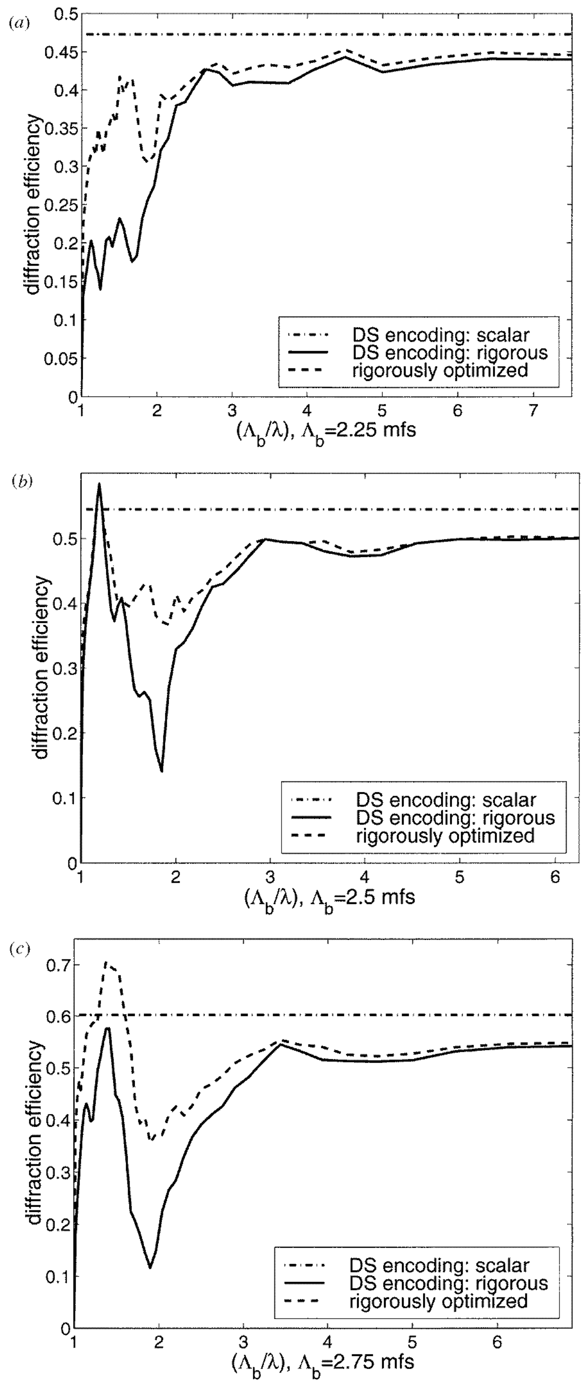

Figure 2. Calculated diffraction efficiencies versus Λ_b/λ with constant Λ_b for blazed gratings with (*a*) $\Lambda_b/\mathrm{MFS} = 2.25$, (*b*) $\Lambda_b/\mathrm{MFS} = 2.5$ and (*c*) $\Lambda_b/\mathrm{MFS} = 2.75$. For the phase depth of 2π, eight equally spaced phase levels were available. The chain curves are the scalar diffraction efficiencies. The full curves are the rigorously calculated diffraction efficiencies. The broken curves are the rigorous diffraction efficiencies for the phase-level distributions which were obtained by the rigorous optimization (RSGO).

in a simple way by a standard coding procedure. For every Λ_b/λ ratio and every profile the optimization has to be done individually.

The results shown so far were obtained for the substrate-to-air propagation direction, which is the standard case for a diffractive microlens. In the following we present results for the opposite propagation direction, i.e. air to substrate, for three selected Λ_b/MFS ratios, 2.25, 2.5 and 2.75, in the binary region. All the other parameters remained unchanged. The results of the calculations are given in table 2. The diffraction efficiencies as a function of Λ_b/λ show for the air-to-substrate propagation direction no significant difference in behaviour than for the opposite direction. The $(\Lambda_b/\lambda)_0$ ratios, which were obtained in the same way as for the substrate-to-air propagation direction, are in a similar range, between 2.3 and 3.6.

Table 2. $(\Lambda_b/\lambda)_0$ values for different Λ_b/MFS ratios of the DS blazed grating in the binary region (eight equally spaced levels were available for 2π phase). The propagation direction is reversed: air to substrate.

Λ_b/MFS	$(\Lambda_b/\lambda)_0$
2.25	2.3
2.5	3.2
2.75	3.6

3.2. Calculation of rigorous diffraction efficiencies for cylindrical lenses

For the calculation of rigorous diffraction efficiencies of cylindrical lenses which were DS encoded, we approximated the ideal lens function of a focusing diffractive lens [8] by a series of blazed gratings with piecewise constant grating periods, the grating periods

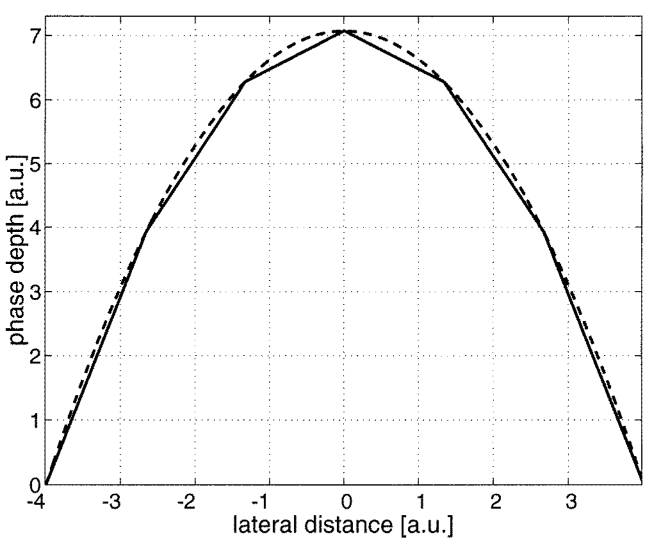

Figure 3. Phase profile of a high-aperture cylindrical lens (broken), and piecewise linear approximation (full). For the calculations we used a finer approximation than shown in the figure.

increasing towards the centre of the lens. The basic principle of this method is illustrated in figure 3. In the binary region, the same discrete grating period to MFS ratios as given in table 1 were used, namely 2, 2.2, 2.25, 2.4, 2.5, 2.6, 2.75 and 2.8. For the larger grating periods, blazed gratings with the following grating period to MFS ratios were used for the approximation: 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5 and 7. We then computed the total first-order lens diffraction efficiency, again for TE polarization, by adding up the various contributions from the different parts of the lens. The calculations were carried out for cylindrical lenses with 2 mm diameter and eight-level DS encoding. The MFS was assumed to be 1 μm and the focal length was chosen so that the lens had a local grating period of 2 MFS at the edge. Efficiencies were computed for three wavelengths, $\lambda = 0.75$, 1.0 and 1.25 μm, corresponding to NAs of 0.38, 0.5 and 0.63, respectively. Our analysis covered the outer part of the lens until a local grating period of 7 MFS was reached. Since for larger grating periods, corresponding to the innermost part of the lens, the performance is to a good approximation scalar and the influence of the coding scheme on the performance of a lens vanishes. We would like to point out that, for the high-aperture lenses studied here, the analysed fraction was the dominating part of the lens, with 76% to 81% of the total lens area. The results are summarized in table 3. The corresponding scalar values are also listed for comparison, where one has to keep in mind that the values for scalar diffraction efficiencies do not include Fresnel reflection losses. The important result is that the difference in diffraction efficiency between the rigorously optimized encoding and the simple DS encoding is quite small for all three lenses in spite of their quite high NA. For the lens with NA = 0.38 the difference amounts to only 1.7%. Furthermore, the absolute value is already quite close to the scalar one, when considering the additional Fresnel losses for the scalar calculation. With increasing NA, the difference between DS encoding and the rigorously optimized value gets larger as expected, because the Λ_b/λ ratios are smaller and one is therefore deeper in the rigorous regime. The absolute differences of 3.6% and 6.1% are, however, still very small when considering the very high NA of 0.5 and 0.63, respectively.

Table 3. Rigorous and scalar first-order diffraction efficiencies η_1 of the outer area for cylindrical lenses. DS encoding with eight phase levels.

	$\lambda = 0.75$ MFS NA = 0.38	$\lambda = 1.0$ MFS NA = 0.5	$\lambda = 1.25$ MFS NA = 0.63
Fraction of total lens:	76.7%	78.1%	80.2%
η_1: DS, rigorous	56.2%	49.0%	42.0%
η_1: rigorously optimized	57.9%	52.6%	48.1%
η_1: DS, scalar	64.5%	63.6%	62.3%

3.3. Optimization potential of extended scalar depth correction using DS

It is known from extended scalar theory [13] that for perfectly blazed gratings the highest diffraction efficiency is obtained for a grating depth, d_{est}, which is shallower than the depth corresponding to the 2π phase difference predicted by simple scalar theory [14], namely

$$d_{\mathrm{est}} = \frac{\lambda}{n - \cos(\arcsin(\lambda/\Lambda_b))}. \quad (1)$$

The incident wavelength is λ, n is the refractive index and Λ_b denotes the grating period. This reduced depth can be explained geometrically by optical path considerations.

The utility of this depth correction in order to increase the diffraction efficiency of a high-aperture lens depends on the manufacturing process. If a direct writing technique is used, the correction of the depth always leads to an improvement, because the depth can be adjusted to the local grating period. For a mask process this is different, because the phase levels for the element have to be chosen for the whole diffractive element independent of the grating period. This means that for a given wavelength the corrected depth is only correct for one specific grating period. For all other grating periods it is more or less inappropriate. The question which now arises naturally is, whether one can achieve an improvement in total diffraction efficiency for the whole lens if this depth correction is applied. Consequently, in this section we investigate rigorously the optimization potential of the grating depth correction according to extended scalar theory when applied to the DS method. Concerning the fabrication constraints, we again assumed eight available phase levels and an MFS of 1 μm.

When comparing the rigorous optimum depths of blazed gratings with the optimum depths obtained from extended scalar theory, one finds a qualitative agreement for $\Lambda_b/\lambda \geqslant 2.8$. Up to $\Lambda_b/\lambda \approx 2.2$ extended scalar theory gives at least the right tendency. On the other hand, a significant improvement of at least a few per cent in diffraction efficiency is only possible for $\Lambda_b/\lambda \leqslant 4.5$. We therefore chose for the analysis a DS blazed grating with a corrected depth according to extended scalar theory and calculated rigorous diffraction efficiencies as a function of grating period for various wavelengths. The grating depth was corrected for the middle of the binary region, $\Lambda_b = 2.5$ MFS. The results for wavelengths of 0.75, 1 and 1.25 MFS are given in figure 4. The results show, as already mentioned, that the gain in diffraction efficiency for the small grating periods is inevitably connected with a loss in diffraction efficiency for the larger ones. In order to study this effect for a practical example, we calculated the diffraction efficiencies for the three cylindrical high-aperture lenses from the preceding section. We included in the analysis the outer part of the lens until a local grating period of 8 MFS was reached. The results are summarized in table 4. The diffraction efficiencies for the cylindrical lenses with corrected depth, following extended scalar theory (equation (1)), show only a marginal improvement between 0.6 and 1.4%. The innermost part of the lenses, which amounts to about 20% in the above cases, is not included in the calculations.

Table 4. Rigorous first-order diffraction efficiencies η_1 of the outer area for cylindrical lenses. DS encoding with eight phase levels for scalar 2π depth and a corrected depth according to extended scalar theory.

	$\lambda = 0.75$ MFS NA = 0.38	$\lambda = 1.0$ MFS NA = 0.5	$\lambda = 1.25$ MFS NA = 0.63
Fraction of total lens:	78.0%	79.4%	81.4%
η_1: scalar 2π depth	56.5%	49.4%	42.1%
η_1: corrected depth	57.1%	50.7%	43.5%

When considering the results displayed in figure 4, one also has to take into account the fact that the design according to extended scalar theory introduces additional diffraction efficiency losses in the remaining central area. We therefore conclude that the design according to extended scalar theory has, in terms of total lens efficiency, no advantages over the design according to simple scalar theory for a mask-based fabrication process. The simple design for the encoding of the diffractive lens results in a better performance.

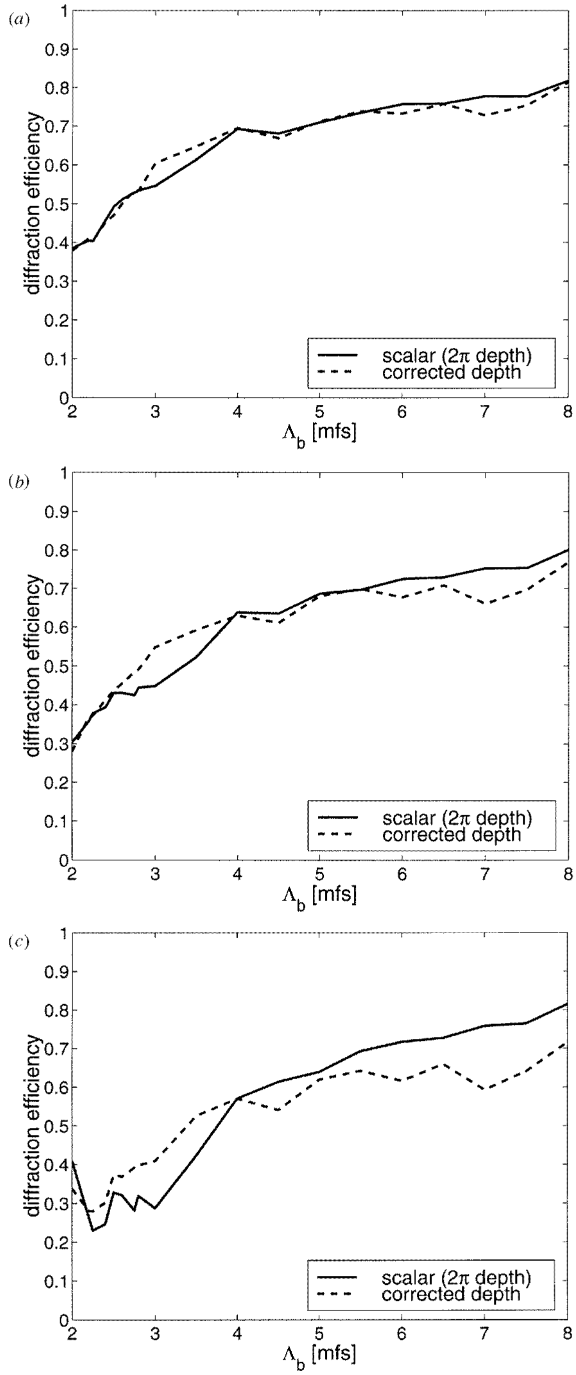

Figure 4. Rigorous diffraction efficiencies as a function of Λ_b for scalar 2π depth (full curve) and corrected depth, according to extended scalar theory (broken curve), for DS blazed grating. The grating depth was corrected for the middle of the binary region, $\Lambda_b = 2.5$ MFS. For the phase depth of 2π, eight equally spaced phase levels were available. Parts (*a*)–(*c*) are calculated for $\lambda = 0.75$, 1 and 1.25 MFS, respectively.

4. Conclusion

We have investigated the scalar-based encoding method of direct sampling (DS) from the point of view of rigorous diffraction theory. In particular, we analysed ideal blazed gratings in the binary region, that is the region where the minimum feature size (MFS) allows only two to three phase levels within one grating period. We found that quantized phase profiles obtained with scalar DS are close to the optimum solutions, even for grating period to wavelength ratios (Λ_b/λ) as small as about 2 to 3.3. This proves the usefulness of DS as an efficient encoding method that can be applied to a range of Λ_b/λ, which exceeds the validity of scalar diffraction theory. It has also been shown that, for still smaller Λ_b/λ ratios, the phase profiles obtained by DS can be improved by a straightforward 'rigorous steepest-gradient optimization'. For certain Λ_b/λ ratios the improvement is as large as 11%–36%.

We calculated rigorous lens diffraction efficiencies for high-aperture focusing cylindrical lenses, encoded with DS, by approximating them as a series of linear blazed gratings with varying grating periods. We showed that the difference in diffraction efficiency between rigorously optimized encoding and simple DS encoding is quite small even for large numerical apertures (NA $= 0.63$) of the lenses. Furthermore, the values were quite close to the expectations from scalar diffraction theory.

Additionally, we rigorously investigated the optimization potential of grating depth correction according to extended scalar theory when applied to the DS encoding. Here we found that the design according to extended scalar theory has no advantage over simple scalar theory when a mask-based fabrication process is used.

Acknowledgment

This research was supported by the Swiss Priority Programme Optique II.

References

[1] Blattner P, Herzig H P, Weible K J, Teijido J M, Heimbeck H J, Langenbach E and Rogers J 1996 *J. Mod. Opt.* **43** 1473
[2] Behrmann G P and Mait J N 1997 *Micro-Optics* ed H P Herzig (London: Taylor and Francis) p 259
[3] Singer W, Herzig H P, Kuittinen M, Piper E and Wangler J 1996 *Opt. Eng.* **35** 2779
[4] Swanson G J 1989 *Lincoln Laboratory Technical Report* 854 MIT Lincoln Laboratory, Lexington, MA
[5] Leger J R, Scott M L, Bundmann P and Griswold M P 1988 *Computer Generated Holography II (Proc. of the Society for Photo-Optical Instrument Engineering)* ed S H Lee, p 82
[6] Welch W H, Morris J E and Feldmann M R 1993 *J. Opt. Soc. Am.* A **10** 1729
[7] Hutchins R, Yang H, Morris J E and Feldman M R 1997 *Appl. Opt.* **36** 2313
[8] Kuittinen M and Herzig H P 1995 *Opt. Lett.* **20** 2156
[9] Pommet D A, Moharam M G and Grann E B 1994 *J. Opt. Soc. Am.* A **11** 1827
[10] Turunen J 1997 *Micro-Optics* ed H P Herzig (London: Taylor and Francis) p 31
[11] Moharam M G, Pommet D A, Grann E B and Gaylord T K 1995 *J. Opt. Soc. Am.* A **12** 1077
[12] Noponen E, Turunen J and Vasara A 1993 *J. Opt. Soc. Am.* A **10** 434
[13] Swanson G J 1991 *MIT Technical Report* 914
[14] Rossi M, Blough C G, Raguin D H, Popov E K and Maystre D 1996 *OSA Tech. Dig. Ser.* **5** 233