Photolithographic fabrication method of computer-generated holographic interferograms

Markus Kajanto, Eero Byckling, Juha Fagerholm, Jussi Heikonen, Jari Turunen, Antti Vasara, and Arto Salin

We consider the fabrication of high-quality interferogram-type diffractive optical elements with conventional photolithographic techniques and compare the results with those achievable with electron-beam lithography. The fringes associated with the phase transfer function of the binary phase holographic interferogram are approximated with rectangles, which can be realized at submicron accuracy using a pattern generator and step-and-repeat camera. The effects of the rectangle quantization are analyzed both numerically and experimentally with the aid of diffraction patterns produced by simple focusing elements. Both resolution and diffraction efficiency of the best holograms approach their theoretical values.

I. Introduction

Computer-generated holograms (CGHs) with optimized complex-amplitude transmission functions can perform rather general wavefront transformations. Consequently, these diffractive optical elements have found use in such diverse areas as display holography, optical information processing, aberration correction in optical systems, optical interconnection, multiple-beam generation, angular spectrum shaping, laser scanning, optical testing, and pattern recognition (see Refs. 1–7 and references cited therein). Traditionally, computer-generated holograms have been fabricated using either a CRT output or computer-controlled pen and ink plotters (or, more recently, laser scanners) for patterning, followed by photoreduction with an SLR camera lens. Problems associated with the limited accuracy of those commonly used CGH fabrication methods have largely been overcome by the use of electron-beam lithography, but the equipment required is expensive and hence available at a few laboratories only. A considerably lower cost alternative is provided by the conventional photolithographic IC mask production techniques. We have applied these techniques to fabrication of various types of diffractive optical element, including holographic lenslet and grating arrays, computer-optimized aspheric lenses, and multiple-beam generators. In this paper, we concentrate on the effects of the limitations of the photolithographic apparatus in production of binary interferometric-type holograms.

A description of our photolithographic process, the apparatus used, and its accuracy is found in Sec. II. In Sec. III, an efficient method is presented for finding the fringes of a holographic interferogram defined by its phase transfer function, and for approximating the fringe structure by a network of optimally sized and located rectangles. The rectangle approximation inherently leads to quantization; a universal measure $\Delta\Phi$, termed the relative local phase error, is established for the severity of the quantization errors of this type.

A numerical analysis of the effects of quantization is performed in Sec. IV, using a simple off-axis focusing hologram as an example. Making use of the Fresnel diffraction theory, we estimate the dependence on $\Delta\Phi$ of the spot size, diffraction efficiency, and the required number of approximating rectangles. It turns out that holograms exhibiting diffraction-limited resolution and nearly theoretical diffraction efficiency can be produced using a moderate number of rectangles, and hence at a comparatively low cost. The analysis of Sec. IV may, in a way, be viewed as a 2-D generalization of the approach presented in Ref. 15 for estimation of the quantization errors occurring in electron-beam lithography.

In Sec. V, an experimental verification is provided of the ability of our fabrication method to produce high-quality diffractive optical elements. In the final sec-
tion, the merits and drawbacks of the photolithographic technique, compared with electron-beam lithography, are discussed. It will be concluded that, except in applications requiring ultimate resolution, the photolithographic method is entirely adequate.

II. Fabrication Method

A holographic interferogram is completely characterized by its phase transfer function $\Phi(x,y)$, the argument of the complex-amplitude transmittance $t(x,y)$ of the element. For a binary interferogram, the regions to be recorded if maximum diffraction efficiency is desired, are the fringes of the hologram, i.e., the bands $2\pi(n - 1/4) < \Phi(x,y) < 2\pi(n + 1/4)$, where $n$ is an integer.

Under usual circumstances, the hologram fringes may locally be viewed as arcs of circles, whose radii of curvature are large compared with the local fringe period. In this case, the fringes can be satisfactorily approximated by trains of rectangles, which are relatively long compared with their widths. Such rectangle trains can be economically recorded on a chrome mask plate using a photolithographic pattern generator. Since the minimum feature size achievable with conventional pattern generators is typically a few micrometers, and the fringe widths are usually of the order of a few wavelengths of light, the mask plate must be photoreduced by a high-quality wafer stepper. A step-and-repeat camera also makes it possible to repeat the same pattern to conveniently produce accurate periodic structures. Furthermore, it allows the combination of several pattern generator masks into a single large hologram.

The photoreduction may be done on a chrome plate (covered by photoresist) to yield an amplitude-modulated binary interferogram, with maximum diffraction efficiency of 10.1% into the first diffraction order. To improve this efficiency four times, the amplitude-type hologram can be converted into a phase-only surface-relief structure by employing suitable etching techniques, such as ion-beam milling. Alternatively, the photoreduction can be done on a plate covered only by a layer of photoresist with thickness suitably chosen to result in a phase delay of $\pi$ rad. This technique is extremely simple, but the resulting hologram must be handled with care since photoresist is not one of the most durable hologram recording materials. More resistant phase-only holograms can be made, for example, on thin silicon nitride films.

The GCA 3600 pattern generator, which we have used, can record rectangles with dimensions ranging from 4 $\mu$m to 3 mm. The accuracy of these dimensions is better than $\pm 0.5 \mu$m, and the positioning error is below 0.3 $\mu$m over an area of $5 \times 5 \text{ cm}^2$. Photoreduction (five times) with a high-quality step-and-repeat camera PAS 2000A thus gives a submicrometer feature size ($\approx 0.8 \mu$m). By careful exposure, the dimensional errors of the rectangles can be kept at about $\pm 0.1 \mu$m. Due to the very low distortion of the imaging lens, and the air-bearing stage with laser interferometer servo control, the absolute positioning errors over a 12.5- X 12.5-cm$^2$ plate are well below 0.3 $\mu$m. System performance is hence adequate for most practical applications, provided that the errors due to quantization effects caused by the rectangle approximation of the fringe pattern may be considered negligible.

III. Rectangle Approximation of the Fringes

The number of approximating rectangles in the hologram is an important factor, since it affects the time needed for recording the hologram by a pattern generator, and thus also the fabrication cost. Cost increase is not severe when the total number of rectangles $N$ in the quantized hologram structure is of the order of thousands. In large holograms, however, $N$ can be considerably larger than in typical VLSI applications and become a serious consideration. On the other hand, it is clear that the increase of the number of rectangles yields a closer approximation of the fringe pattern and thus, presumably, better performance. It is not perfectly clear how severely the quantization errors affect the performance of the hologram, although estimates have been presented for the 1-D case. To show (in the next section) that surprisingly large quantization errors can be tolerated in most applications, we now proceed by considering the calculation of the optimal locations of the rectangles and by establishing a suitable quality criterion for this kind of a rectangle approximation.

The boundaries of the fringes, the lines $\Phi(x,y) = 2\pi(n \pm 1/4)$, as well as the fringe center lines $\Phi(x,y) = 2\pi n$, can be tracked by several different algorithms. The problem may, for example, be formulated as an initial-value problem associated with an ordinary first-order differential equation. Due to the high numerical precision required, we have chosen a more direct procedure outlined briefly below.

If the phase function gradient $\nabla \Phi(x,y)$ has zeros in the hologram area (such as the center of an on-axis Fresnel zone plate), the hologram must be divided into subregions $\mathcal{D}$ that do not contain any of these zeros. Once this is done, all fringes that appear in any of the (sub)regions $\mathcal{D}$ must cross the boundary $\partial \mathcal{D}$ at least twice. By iteration along $\mathcal{B}$ it is a simple matter to find, for some $n$, the points where the lines $\Phi(x,y) = 2\pi n$ and $\Phi(x,y) = 2\pi(n \pm 1/4)$ intersect $\mathcal{B}$. Starting from the intersection points, these lines can be followed step by step inside $\mathcal{D}$ until the boundary $\partial \mathcal{D}$ is again reached. The next point on the line is searched iteratively by scanning directions close to that of the tangent of the line at the found point. In practice, it suffices to trace points on the center line $\Phi(x,y) = 2\pi n$ only. Points on the other lines will then be found in directions $\pm \nabla \Phi(x,y)|_{(x_0,y_0)}$, at distances approximately equal to $d(x_0,y_0)/4$, where the local fringe period $d(x_0,y_0)$ in the neighborhood of the point $(x_0,y_0)$ is defined as

$$d(x_0,y_0) = 2\pi |\nabla \Phi(x,y)|_{(x_0,y_0)}|^{-1}. \quad (1)$$

The above outlined tracking procedure can be repeated for all fringes (i.e., values of $n$) in the region $\mathcal{D}$, finding the starting point of the next fringe by travers-
Fig. 1. Criterion for determination of length $L$ of a rectangle: $\delta/d$ is equal to a predetermined constant.

Fig. 2. Determination of position $S$ (and width $w$) of a rectangle; the shaded and dotted areas must be equal.

The step size used during the tracking procedure must be relatively small compared with the sizes of the final rectangles. Each rectangle, can, however, be defined by giving the coordinates of three of its corner points only. It thus saves computer resources to fit the rectangles into each fringe immediately after its path has been tracked. Assuming that the fringes are arcs of circles, and that their separation remains constant, over the area occupied by one approximating rectangle, length $L$ of any rectangle can be determined by the following criterion, illustrated in Fig. 1. The deviation $\delta$, at a distance $L$, of the tangent $S$ from the fringe center line may not exceed a product $ad$, where $a$ is a predetermined constant chosen small enough to ensure that the assumptions made above are approximately valid. From the geometry of Fig. 1, we straightforwardly obtain an approximate result:

$$L = [ad(ad + 2R)]^{1/2}. \tag{2}$$

Here the fringe period $d$, defined by Eq. (1), can be calculated from the phase function, and the local radius of fringe curvature $R$ can be estimated from three tracked points on the line $\Phi(x,y) = 2\pi n$. In view of Eq. (2), a fine hologram structure (small $R$ and/or $d$) implies a small value of $L$.

Once the length of a rectangle has been determined, an equation for its optimal location (lateral deviation $D$ from the fringe center line) is obtained by straightforward calculation from the geometry of Fig. 2. The result is

$$D = R \left[ 1 - \frac{2R}{L} \arctan \frac{L}{2R} \right]^{1/2}. \tag{3}$$

The optimal width of the rectangle is naturally $w = d/2$. Using the geometry of Fig. 2, we introduce a quantity $\Delta\Phi = 2\pi D/d$, which describes the magnitude of the relative local phase error caused by replacement of a curved fringe section by a rectangle. To a good approximation, $\Delta\Phi$ is equal to $2\pi a/12$, and hence

$$a = 12\Delta\Phi/2\pi. \tag{4}$$

If we choose a particular value of $\Delta\Phi$ for all rectangles and use Eqs. (4), (2), and (3), sequentially, to calculate the network of the approximating rectangles, the quantization error caused by each rectangle is practically equal, and also independent of the particular form of the phase function. Thus, if we analyze the performance of any hologram as a function of $\Delta\Phi$, we may confidently expect that the analysis is, to a good approximation, universally valid. For example, if a complex aspheric holographic interferogram is used as an aberration compensator in an image-forming optical system, the contributions of its quantization errors to the image degradation can be estimated from the results of the next section, where the effects of quantization in simple focusing elements are analyzed.

IV. Numerical Error Analysis

The diffraction pattern in any transverse observation plane behind an arbitrary hologram, which is composed of rectangles, can be straightforwardly calculated by Fresnel diffraction theory. For each elementary rectangle, complex field amplitudes are evaluated at a grid of points in the observation plane by the well-known formula\textsuperscript{17} utilizing Fresnel integrals. The different orientations of various rectangles are taken care of by simple coordinate transformations. The superposition principle of diffraction theory is applied to sum up the contributions from all the rectangles. Finally, the total intensity distribution is found by squaring the resulting field distribution at the observation plane.

Although a NAG Library\textsuperscript{18} subroutine evaluates a Fresnel integral in just 10-µs CPU time in the IBM 3090/150VF supercomputer, calculation of the diffraction patterns of large holograms may be quite a formidable task. Holograms typically contain thousands of rectangles, each of which requires $8MN$ evaluations of a Fresnel integral, if we want $MN$ resolution points in the observation plane. However, with low focal ratios the number of rectangles and hence also the calculation time are reasonable, and as discussed above, the results obtained give good indication of the quality of more complicated holograms, which have the same value of the phase error $\Delta\Phi$.

In Fig. 3 the quantized structure of a simple off-axis focusing element with a paraxial phase transfer function $\Phi(x,y) = k(x^2 + y^2)/2F$, where $k = 2\pi/\lambda$, is dis-
Fig. 4. Diffraction pattern of the structure shown in Fig. 3. The scales are in meters.

Fig. 5. Structure of an F:50 off-axis focusing element with randomness in the length of the first rectangle of every fringe.

played for $\Delta \Phi/2\pi \approx 1/6$. The focal length of the hologram is $F = 100$ mm, the size of the (square) aperture is $2 \times 2$ mm, and the aperture is centered at a point $(x_0, y_0) = (1.5$ mm, $1.5$ mm). Figure 4 shows the calculated diffraction pattern corresponding to the hologram in Fig. 3, evaluated at a distance $F$ behind the hologram plane in $100 \times 100$ points. It is interesting to note that, although the quantization looks extremely rough, resolution (Rayleigh criterion) is, within the accuracy of Fig. 4, equal to the theoretical $62 \mu$m (for $\lambda = 633$ nm) of an equivalent lens. This somewhat surprising result shows that, if the rectangles are optimally located in the sense discussed in the previous section, rather large local phase errors $\Delta \Phi$ can be tolerated without significant loss of resolution.

Rough quantization has, however, a noticeable effect in diffraction efficiency: only 4.15% of the incident light is diffracted inside the main lobe, while the theoretical value is 8.23%. Had the calculation been performed for a phase-only hologram instead of an amplitude-type hologram, a fourfold increase of diffraction efficiency would of course have resulted. Clearly, much of the light is concentrated in the strong spurious spots located on both sides of the main lobe. Their peak intensity is $\sim 12\%$ of the peak intensity of the main lobe. These ghosts, which look familiar to users of ruled grating spectrometers, are caused by the clear vertical periodicity in the fringe structure. The diffraction efficiency into the first diffraction order (the area covered by the $100 \times 100$ grid in Fig. 4) is 9.32%, i.e., slightly lower than the theoretical efficiency of 10.1%.

To avoid the unwanted periodicities, and hence the ghost spots, we could use random reduction of the rectangle lengths. However, it has proved better to introduce randomness only in the length of the first rectangle of every fringe, i.e., that touching the boundary $B$ of $D$. In this way, the required number of rectangles remains practically unchanged, and the periodicities are in fact destroyed more completely. The result of randomness (the other parameters being unchanged) is shown in Fig. 5, and the corresponding diffraction pattern is displayed in Fig. 6. Resolution remains unaffected, as expected, and the diffraction efficiency inside the main lobe increases slightly, to 4.46%. Most notable, of course, is the almost total disappearance of the ghost peaks. The diffraction efficiency into the grid area of Fig. 6 has been reduced to 7.39%. Randomness, which causes sharper hologram structure, thus tends to smooth the ghost spots, scattering light to form a weak widespread halolike background.
To increase diffraction efficiency into the main lobe, we clearly have to reduce the quantization error \( \Delta \Phi \). In Fig. 7, we have chosen \( \Delta \Phi/2 \pi \approx 1/40 \). Now the rectangles approximate the fringes very well, and the diffraction pattern, shown in Fig. 8, looks like a pure \( \text{sinc}^2 \) distribution. Diffraction efficiency into the main lobe is as high as 8.16\%, and in the grid area it is 9.81\%. The drawback accompanied with this improvement is over a twofold increase in the number of rectangles.

The structure in Fig. 3 (corresponding to \( \Delta \Phi/2 \pi = 1/6 \)) clearly represents an upper practical bound for the tolerable quantization error. On the other hand, no significant improvement can be gained by reducing the phase error \( \Delta \Phi/2 \pi \) below 1/40 (Fig. 7). By calculating several hologram structures, diffraction patterns, and efficiencies, we obtain the following empirical dependence of the diffraction efficiency \( \eta \) on the quantization error, valid in the \( \Delta \Phi/2 \pi < 1/6 \) range:

\[
\eta = \eta_0 [1 - 18.5(\Delta \Phi/2 \pi)^2]\%.
\]

Here \( \eta_0 = 10.1\% \) for an amplitude-coded hologram, and \( \eta_0 = 40.5\% \) for a binary-phase interferogram. The corresponding figures for the diffraction efficiency inside the main lobe are \( \eta_0 = 8.23\% \) and \( \eta_0 = 32.9\% \). The law expressed in Eq. (5) is to a good approximation independent of the form of the phase function. It would be useful to have a similar law for the dependence of the number of approximating rectangles \( N \) on \( \Delta \Phi \), but \( N \) does depend on the form of the phase function.

V. Experiments

We have successfully fabricated a large number of different types of focusing elements, gratings, and holographic interconnect elements. Experimental results have been in excellent agreement with numerical predictions; for example, the ghosts were clearly visible in holograms of the type displayed in Fig. 3. To establish the ability of the fabrication technique to produce very high-quality diffractive optical elements, we present more detailed test results of the structure shown in Fig. 7.

Diffraction efficiencies of the fabricated holograms were determined by illuminating the hologram with a plane wave and comparing the first-order diffracted intensity to the intensity of the beam transmitted by an equally sized fully transparent aperture, also fabricated with a pattern generator. Diffraction efficiencies of holograms of the type shown in Fig. 7 agreed well with theory: the ratio \( \eta_{\text{measured}} / \eta_{\text{theoretical}} \) was of the order of a few percent in all cases.
Profiles of diffraction patterns were measured by scanning a detector (placed behind a pinhole) over the pattern, which was first magnified with the aid of a microscope. Perfectly diffraction-limited performance was observed; this is clearly evidenced by the scan shown in Fig. 9, in which deviations from the ideal sinc\(^2\) distribution are below 2%. A slight ripple in the sidelobes is mostly due to coherent noise originating from dust and minor imperfections in the testing system.

VI. Discussion and Conclusions

In this paper, we have aimed to demonstrate that conventional photolithography can be applied to produce high-quality computer-generated interferogram-type holograms at a relatively low cost. From both numerical and experimental results, we deduce that the inherent rectangle quantization does not appreciably reduce resolution, provided that the approximating rectangles are located optimally with respect to the fringes (Fig. 2); the quantization effects were seen only in the diffraction efficiency.

We emphasize once more that, although the numerical and experimental results presented in this paper exclusively deal with focusing elements, the fabrication and analysis techniques presented here are by no means restricted to this simple case. Rather general computer-optimized phase transfer functions can be realized; one example appears in Ref. 13. Also, the method is well suited for fabrication of holograms that are inherently composed of rectangles, e.g., Lohmann-type display holograms and Dammann grating beam splitters.\(^{14}\)

Although electron-beam fabrication of computer-generated holograms has recently gained considerable popularity due to the high precision available, it also has some drawbacks, and it is not yet clear what method might be most generally useful for fabrication of different types of holographic element (see the discussion in Ref. 19). The clear advantage of E-beam lithography over our photolithographic technique is that the minimum feature size is an order of magnitude smaller (≈0.1 μm vs 1 μm). This small feature size permits in principle fabrication of almost any hologram for visible and near-UV light, while our method is not practical if the period is below 2–5 μm, depending on the form of the hologram phase function. However, progress is still made in development of the photolithographic apparatus, and linewidths of ≈0.4 μm will be feasible in the near future.\(^{20}\) An important consideration is also the distortion in patterns written by the E-beam system. This distortion is typically a considerable portion of the wavelength of light.\(^{10}\) The effects of plotting distortion in the quality of the image formed by the hologram have, to our knowledge, not been fully investigated, but it can be anticipated that these effects narrow the difference in performance between photolithographic and E-beam techniques. Finally, at submicron feature size levels, photoresists are superior to electron resists in terms of the MTF.\(^{20}\)

This work was partially funded by the Academy of Finland and Tekniikan Edistamissaatio, Finland. Inquiries for reprints and further information should be addressed to A. Vasara.

References