SANDIA REPORT

SAND2005-6794
Unlimited Release
Printed October 2005

Diffractive Optics in the Infrared (DiOptIR)

Shanalyn A. Kemme, David W. Peters, Eric A. Shields, Joel R. Wendt, Tony R. Carter, Sally Samora, G. Allen Vawter, Charles R. Alford

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

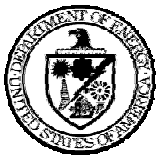
Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online



SAND2005-6794
Unlimited Release
Printed October 2005

Diffractive Optics in the Infrared (DiOptIR) LDRD 67109 Final Report

Shanalyn A. Kemme, Joel R. Wendt
Photonic Microsystems Technologies

David W. Peters
Applied Photonic Microsystems

Eric A. Shields
Remote Sensing & Exploitation

G. Allen Vawter, Charles R. Alford
RF Microsystems Technologies
Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM 87185-0603

Tony R. Carter, Sally Samora
L&M Technologies, Albuquerque, NM 87123

Abstract

This diffractive optical element (DOE) LDRD is divided into two tasks. In Task 1, we develop two new DOE technologies: 1) a broad wavelength band effective anti-reflection (AR) structure and 2) a design tool to encode dispersion and polarization information into a unique diffraction pattern. In Task 2, we model, design, and fabricate a subwavelength polarization splitter.

The first technology is an anti-reflective (AR) layer that may be etched into the DOE surface. For many wavelengths of interest, transmissive silicon DOEs are ideal. However, a significant portion of light (30% from each surface) is lost due to Fresnel reflection. To address this issue, we investigate a subwavelength, surface relief structure that acts as an effective AR coating.

The second DOE component technology in Task 1 is a design tool to determine the optimal DOE surface relief structure that can encode the light's degree of dispersion and polarization into a unique spatial pattern. Many signals of interest have unique spatial, temporal, spectral, and polarization signatures. The ability to disperse the signal into a unique diffraction pattern would result in improved signal detection sensitivity with a simultaneous reduction in false alarm.

Task 2 of this LDRD project is to investigate the modeling, design, and fabrication of subwavelength birefringent devices for polarimetric spectral sensing and imaging applications. Polarimetric spectral sensing measures the spectrum of the light and polarization state of light at each wavelength simultaneously. The capability to obtain both polarization and spectral information can help develop target/object signature and identify the target/object for several applications in NP&MC and national security.

Table of Contents

Figures and Tables

1. Introduction

Transmissive diffractive optical elements (DOEs) are increasingly utilized, particularly in long wavelength bands of interest to the remote sensing community. Often the wavelength band of interest is broad (i.e., 2-5 μm or 8-12 μm) and fortuitously centered around wavelengths not absorbed by high-index semiconductors such as silicon. However, Fresnel reflection losses for a silicon transmissive optical component are significant (30% per interface) due to the substantial refractive index contrast at a surface. The resulting large reflection loss reduces an optical system's signal-to-noise ratio and thus its ability to discriminate between the object of interest and the ambient background. This is the issue we investigate in Task 1 of this LDRD.

In Task 1, we numerically investigate the effect of a binary grating etched into silicon. Rigorous coupled wave analysis models the orders reflected and transmitted from the grating and is used to model the effect of the grating over a wide range of grating periods, from subwavelength to much larger than the wavelength. With this numerical design tool, we identify a broad wavelength band and angular band, triangular profile that performs well as an effective AR coating. This triangular profile can be fabricated with a wet-etch process, which is a simpler process than dry-etch processes, to add to an existing diffractive optic with non-planar surfaces. We fabricate components using this process and measured their transmission and reflection properties. Based upon these results, we modify the subwavelength period for the wet-etch process and complete a second design-fabricate-measurement cycle with improved results.

Also in Task 1, we encode an Iterative Fourier Transform Algorithm (IFTA) for diffractive optical elements. The application considered here is the conversion of a uniform beam of light to a desired intensity distribution in the far field. Elements that perform this operation are often called computer generated holograms (CGHs). The goal of CGH design is to calculate the height profile such that in the far-field the modified beam of light has an intensity distribution matching the desired pattern. We develop and encode several algorithms that produce an arbitrary pattern, with good uniformity, and the desired dispersion, as a function of wavelength. Thus the goal of CGH design that would result in a desired pattern that was wavelength dispersed is achieved. Our second goal of encoding a pattern that would discriminate one polarization from another is attained through a two-step process. First the CGH designs, that produce the intensity patterns for each polarization, are realized. Then the phase of each pixel is made specific to one of the two polarizations with an additive, subwavelength etch.

The large refractive index difference between semiconductors and the surrounding air puts DOEs, such as the subwavelength birefringent component described in Task 2, just within the realm of fabrication possibility. In Task 2, we first define the optical birefringence and dispersion of the subwavelength DOE with the goal that the two linear polarizations (TE and TM) can be completely separated into two angular ranges. We develop a numerical design code that calculates the effective index versus wavelength for two polarizations (dispersion and birefringence). The fabrication process for a nanoscale, subwavelength structure in GaAs is implemented for the newly-installed chemically assisted reactive ion-beam etching (CAIBE) equipment in the CSRL. Well-controlled vertical features of aspect ratios greater than 10:1 are achieved. We define a subwavelength device with angular dispersion for the wavelength range of interest and for the two polarizations. This device was fabricated and will be characterized.

2. Rectangular-Profile Anti-Reflection (AR) Gratings on Silicon

We investigate the effect of a binary grating etched into silicon as illustrated in Figure 1. Such a grating may be modeled as an effective index layer, assuming that the period of the grating is well below the wavelength of light inside the silicon. Rigorous coupled wave analysis models the orders reflected and transmitted from the grating and is used here to model the effect of the grating over a wide range of grating periods, from subwavelength to much larger than the wavelength.

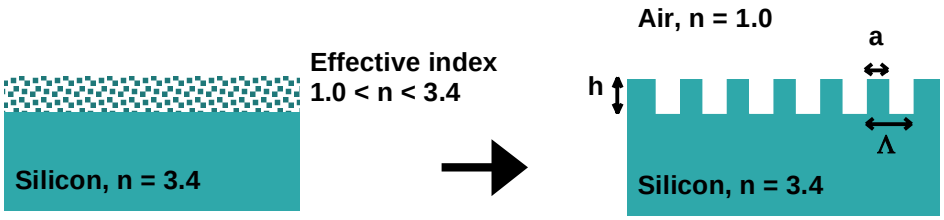


Figure 1. Thin film and subwavelength ($\Lambda<\lambda$) binary grating equivalent in silicon.

Treating the grating layer as a uniform layer of material with a refractive index between that of air and silicon is only valid in the regime where the grating period is less than the wavelength of light in the higher index material, in this case silicon. Increasing the depth of this layer (moving up a vertical path in the two-dimensional plots of Figure 2) shows a sinusoidal variation as expected. The thickness of the layer that achieves the lowest reflection depends also on the refractive index of the layer. The left plot in Figure 2 shows the reflectivity from a silicon surface that has been coated with a thin film of indicated refractive index. The right plot show the reflectivity from a silicon surface that has been coated with a subwavelength grating that acts as an equivalent thin film. Regions of almost zero reflectivity can be seen in either case for certain values of refractive index (or equivalently grating duty cycle) and depth of the layer resulting in an anti-reflection (AR) coating.

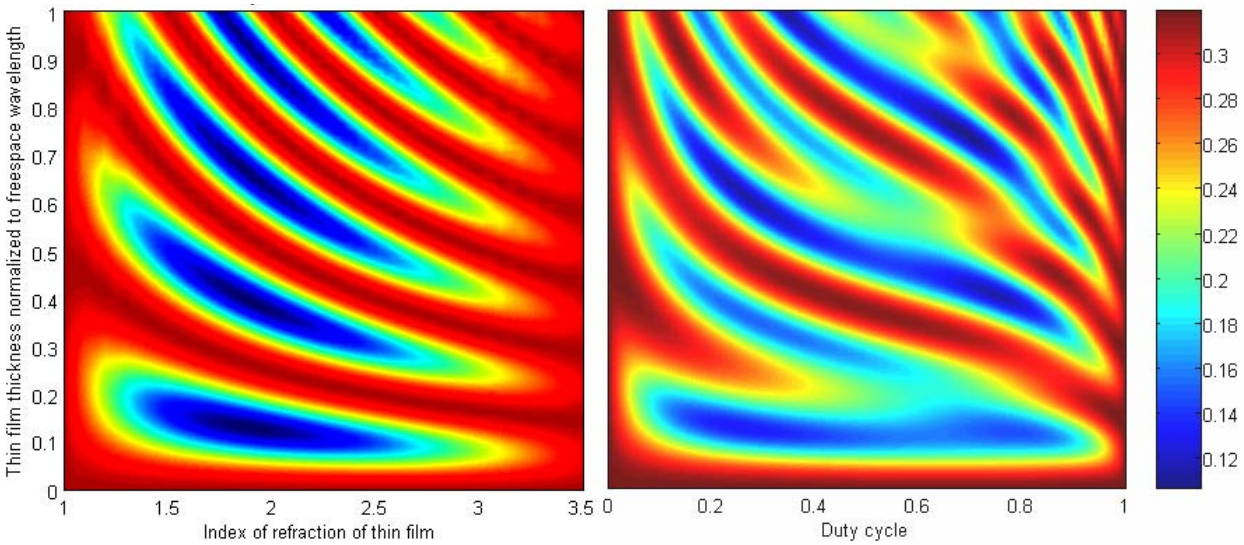


Figure 2. Reflectivity from a silicon surface with a thin film or a subwavelength ($\Lambda=\lambda/2$) grating as a function of effective index (equivalent to grating duty cycle) for a 45-degree polarization, normally-incident plane wave.

Replacing the uniform thin-film layer (left plot in Figure 2) with a subwavelength period, binary grating ($\Lambda=\lambda/2$) (right plot in Figure 2) results in a near-identical plot when the duty cycle is

scanned instead of the refractive index of a thin film. This illustrates the equivalence between a thin film layer and a subwavelength grating. The effective index of the grating may be calculated for a specified incident light polarization using the refractive indices of the two materials and the grating duty cycle. Thus it is straightforward to precisely specify the reflectivity of an effective AR coating by choosing the appropriate duty cycle and depth of the binary grating. It is clear that surface texture, as long as the feature sizes remain smaller than the incident wavelength of light will never increase the light lost due to reflection, and may make more light available for transmission.

However, there is not a one-to-one correspondence between a homogeneous thin film and a subwavelength effective-index grating as can be seen in Figure 3. As the period of the grating is increased from the subwavelength regime, additional transmitted and reflected orders appear. This figure may be broken down into three regions. The first is the region where the period of the grating is subwavelength compared to the wavelength of light in both air and silicon, roughly for periods less than 0.3. In this subwavelength regime we see no variation with period, and a sinusoidal variation of the reflectivity as a function of the grating depth. In the region where the period varies from 0.3 to 1.0, the grating is subwavelength compared to the wavelength in air, but is not subwavelength in the silicon, thus allowing higher transmitted orders, but only the zeroth order reflected wave. In this regime the reflectance is quite small and less dependent upon grating depth. Finally, in the range where the period is greater than the freespace wavelength, higher reflected and transmitted orders may propagate.

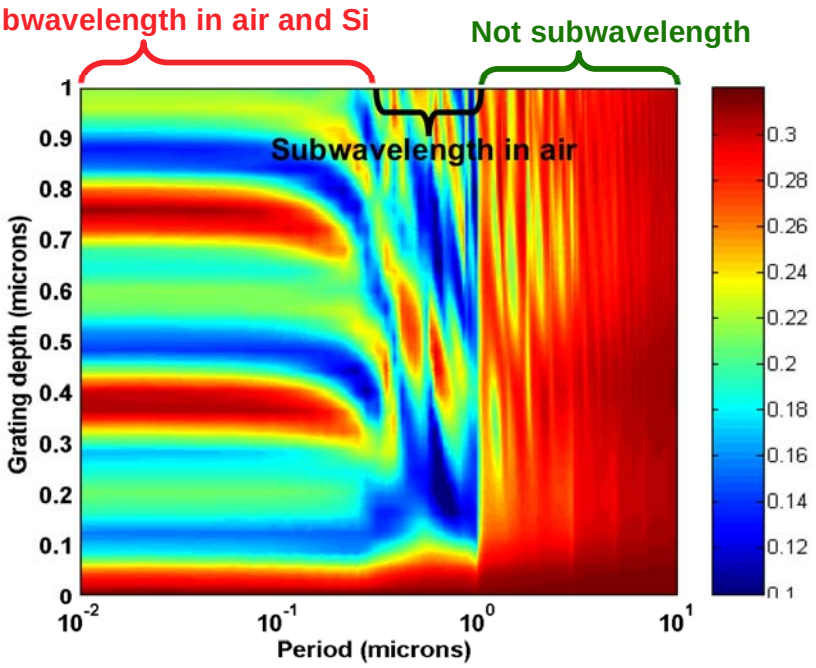


Figure 3. Reflectivity from a silicon surface as a function of grating period and grating depth for a 45-degree polarization, normally-incident plane wave.

3. Wet-Etch, Triangular-Profile Anti-Reflection (AR) Gratings on Silicon

We explored the facility of triangular-profile effective AR coatings as an alternative to rectangular-profile effective AR coatings. Our numerical analysis indicated that triangular

profile effective AR coatings have superior, wideband performance. Moreover, we have analytically demonstrated the triangular profile obtained from wet-etching single crystalline (100) Si results in a good, broadband response. Wet etching silicon is a simpler fabrication additive process than anisotropic dry etching and produces triangular-profiled ridges upon appropriate crystalline substrate orientation.

It is well known from years of silicon processing that wet etches of silicon lead to slanted sidewalls as a result of the etch following crystallographic directions. In (100) silicon this causes a triangular shape to form with an etch angle of 54.74° as in Figure 4. ((100) is a designation of the direction that a wafer is cut with respect to the crystalline structure of silicon.)

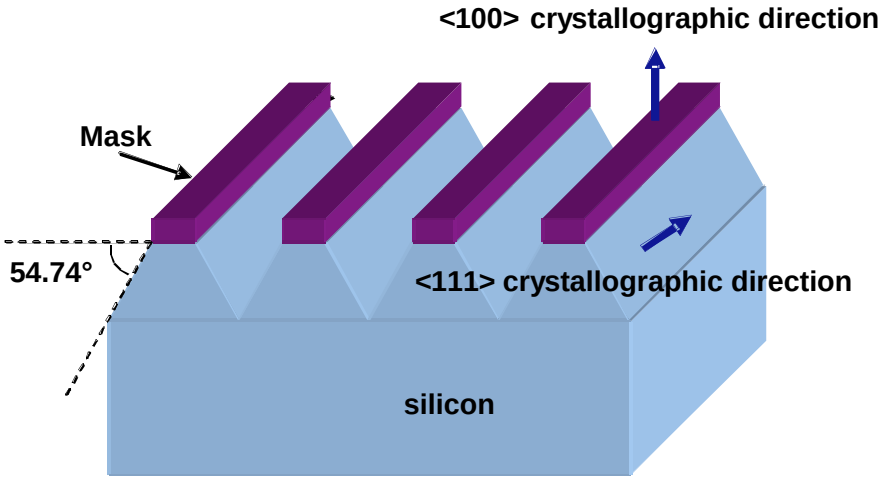


Figure 4. Etch geometry for (100) silicon wafer

These shapes look very similar to “moth eye” antireflection structures used as antireflection coatings on various surfaces. Moth eyes are pyramidal structures that act to taper the refractive index from the substrate to the air. They work best when the base is subwavelength in width and the height is over a wavelength. The natural etch angle of silicon does not allow for such a tall structure, but nonetheless we expect to see a lessening of the reflection as a region of intermediate refractive index is introduced.

Wet-etching (100) silicon leads to triangular ridges when the periodic ruled mask is parallel or perpendicular to the wafer flat. Figure 5 shows a scanning electron micrograph (SEM) of the triangular profiles obtained when wet-etching (100) silicon. The mask is aligned parallel to the major flat. The SEM on the left is from a three micron period mask, and that on the left is from a one micron period mask. Here is it easy to see the undercut of approximately 100nm obtained in a two minute etch cycle.

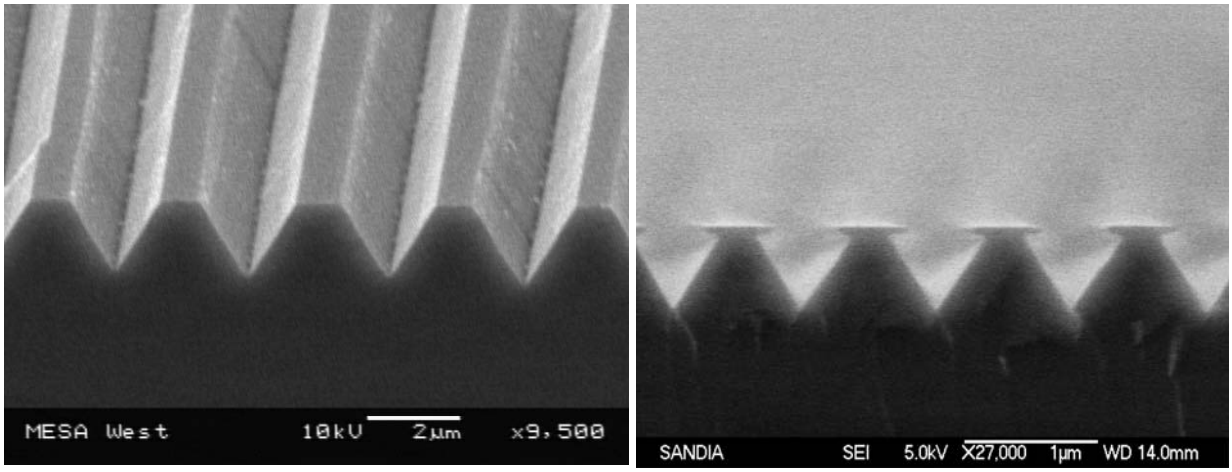


Figure 5. Scanning Electron Micrograph (SEM) of wet-etched (100) silicon. The left picture is a three micron period and the right picture is a one micron period grating. The etch mask is still visible on the one micron period grating, illustrating a 100nm undercut for the two minute etch.

The resulting triangular profiles have a shape as shown in Figure 6. The width of the top mesa depends on the width of the mask lines and the rate that the mask undercuts. The mask undercut is significantly slower than the (111) silicon plane etch, so etch time to completely remove the mask for a triangular profile is many minutes. While this is theoretically possible, it is not practical as long etches result in rough sidewalls as the mask erodes nonuniformly.

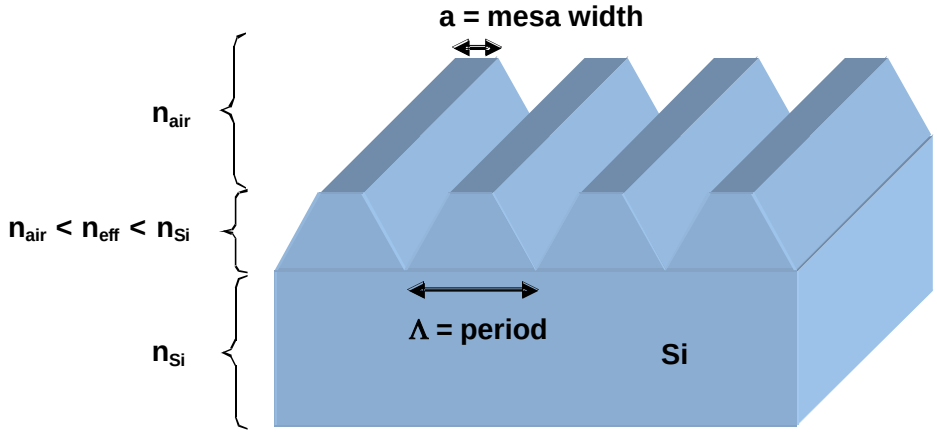


Figure 6. Resultant surface structure after wet-etch with mask consisting of lines.

3.1 Design Using a Near-Wavelength Period

To benefit from the greater depth offered by a larger period, we first considered periods of 1, 2, and 3µm. Across the 2-5µm wavelength range, these periods are not always subwavelength. Inside the silicon the wavelength is reduced by the refractive index (n_{Si} = 3.4). At times this arrangement can be beneficial, as in the air only the zero order (specular) reflection propagates, while inside the material multiple orders carry energy. Designs in this regime can use these internal higher orders to design interesting optical effects outside the material.

The structured air/silicon interface is easily and accurately modeled using rigorous coupled wave analysis (RCWA). The pyramids are sliced into cross-sections in the simulation, such that the

sidewalls are approximated as stairsteps. However, using seven layers, the individual heights are much less than a wavelength and the stairstep approximation has negligible impact on the numerical results. Figure 7 shows the total of the reflected orders from this surface for normal incidence. All of the calculations use 45° polarization to model unpolarized light. This simulation only considers the reflection at the structured surface and ignores the reflection at the back (polished) surface. Based on these calculations, we see impressively low reflectivity from the front surface.

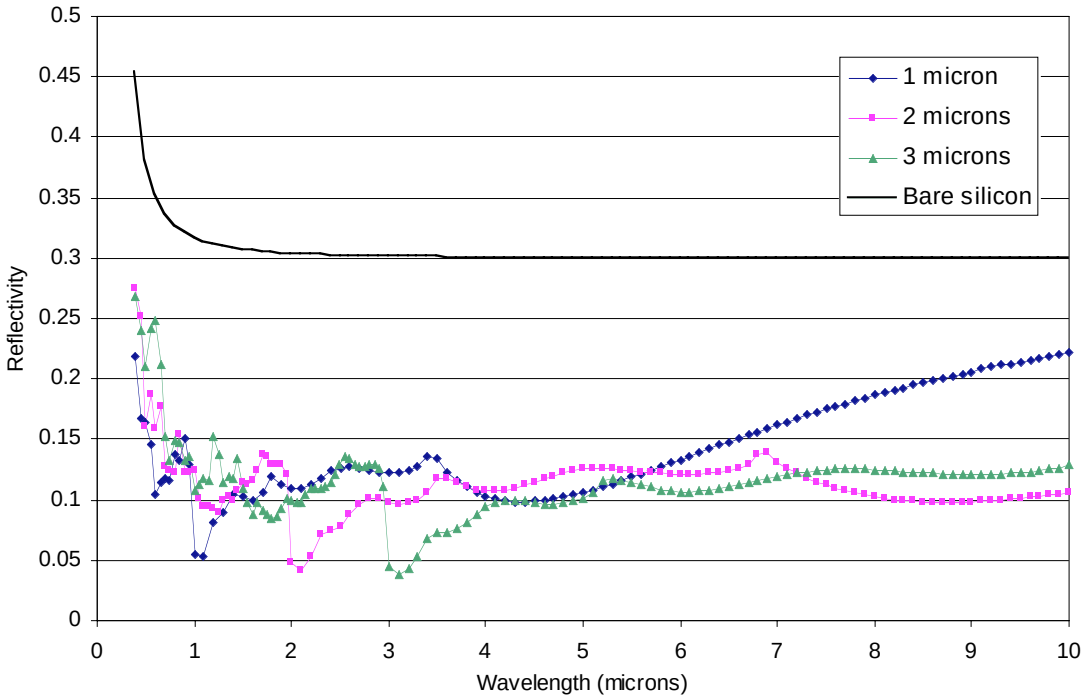


Figure 7. Total of reflected orders from etched silicon surface with periods of 1μm, 2μm, and 3μm over a broad wavelength range for normal incidence and 45° polarization.

Although optimizing the front surface to minimize the specular reflection seems at first inspection to be a logical approach, it ignores the possible interaction of higher-order transmitted energy. Since silicon has a high refractive index (~3.44 in the 2 to 5μm range) higher-order diffracted modes can exist inside the silicon, even when only the zero-order (specular) beam propagates in the air. The unintended effect of these higher orders will be demonstrated later.

For the 1μm period case, with mesa width equal to 25% of the period, the zero transmitted, the zero reflected order, and the ±1 transmitted orders propagate as shown in Figure 8. The two first order diffracted beams exit the grating at angles such that the light is bent at an angle that is wavelength dependent. The shallowest (closest to straight through the slab) angle occurs at the shortest wavelength (35° at 2μm). As the wavelength increases this angle increases and the amount of power in these non-zero orders dwindles. By 3.4μm these orders have disappeared, leaving only the zero-order beam.

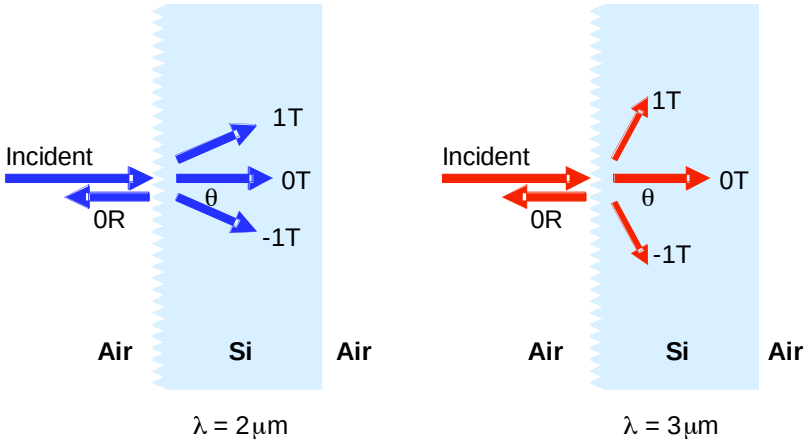


Figure 8. Propagating orders from normally-incident light on wet-etched 1μm period grating on Si wafer.

The zero order propagates through the silicon slab with some of the light transmitted through the back (polished) side of the silicon slab as governed by simple Fresnel equations. As some light is reflected by the backside, we see light reflected between the polished and etched surfaces as indicated by the red arrows in the gray box labeled "a." in Figure 9. This light will all eventually exit one side of the silicon slab; the amounts are calculated by summing an infinite series. That light exiting the front surface contributes to reflection while that exiting the back surface contributes to the transmitted signal.

At the shorter end of the wavelength range the amount of energy in the ±1T orders is greater than that in the zero order. Since the two first order beams are symmetric, we may consider only one and double the resulting reflection and transmission values obtained from following one of the orders. One might expect the ±1T orders to behave in a similar way to the zero order, with partial transmission and reflection following the Fresnel equations out the back surface. However, the angle for all wavelengths is greater than the critical angle for the flat silicon/air interface. All of this energy thus undergoes total internal reflection (Figure 9, gray box b) and strikes the grating from the silicon side.

In Figure 9 gray box c, we see the incident beam inside the silicon strikes the grating at the same angle θ as the angle of the first order transmitted beam from the first air/silicon interface. The first order transmitted beam at this silicon/air interface will exit the interface at 0°. This exiting light will thus contribute to the reflected signal. Three beams are created by the grating within the silicon (gray box c), one of which is normal to the surface, and leads to light that will ultimately contribute to the reflected or transmitted signal. The other two reflected beams continue at the same angle, θ, with opposite sign, and again undergo total internal reflection from the backside. A series of infinite geometric series give us the final amounts of light that will exit the front surface (reflected signal) and back surface (transmitted signal).

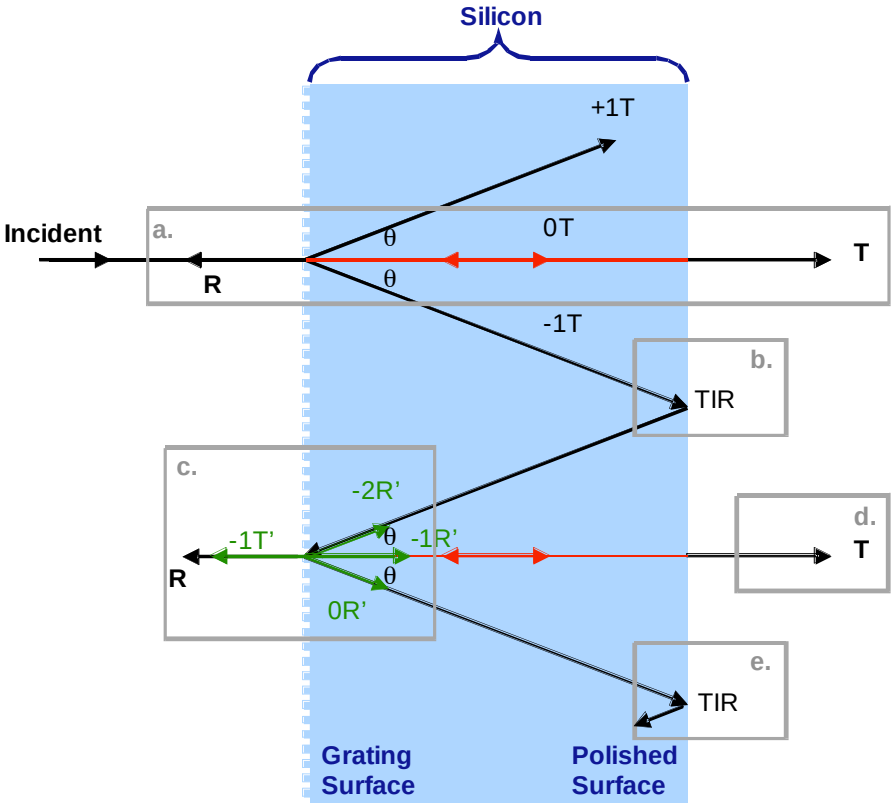


Figure 9. Diagram of diffracted orders for normally-incident light on silicon slab with wet-etched 1μm-period grating on front side only.

The front interface is surface 1, the back interface is surface 2: these numbers will be subscripts. Surface 1 may have light incident from the air, or backreflected light can strike it from the silicon side. Light from the silicon side is designated with a prime (‘). Light can be incident inside the slab at an angle of 0° or θ.

R (in gray box a) is the direct reflection from surface 1 plus the light that escapes from the infinite series represented by the red arrows.

$$R = 0R_1 + \frac{(0T_1)(R_2\angle 0°)(0T_1)}{1-(0R_1')(R_2\angle 0°)} \tag{4}$$

T (in gray box a) is the sum of the initial straight-through pass and the light that escapes from the infinite series represented by the red arrows.

$$T_a = \frac{(0T_1)(T_2\angle 0°)}{1-(0R_1')(R_2\angle 0°)} \tag{5}$$

Total internal reflection (in gray box b) requires no calculation. In gray box c, this beam leads to four diffracted beams.

The first we will consider is the beam that exits into the air. This will become part of our reflected signal. It corresponds to a diffraction into the -1T order at the first interface, and a -1T' diffraction at the first interface from the silicon side, or

$$R = (-1T_1)(-1T_1'\angle\theta). \tag{6}$$

In gray box d we see a beam that consists of the infinite series produced by the transmission from surface 2 from the light bouncing normally between the two surfaces designated by the red arrows between gray boxes c and d. This light may be represented as:

$$T_d = \frac{(-1T_1)(-1R_1'\angle\theta)(T_2)}{1-(0R_1'\angle 0°)(R_2)}. \tag{7}$$

However, this process is repeated infinitely as the 0R' and -2R' beams continue at the same angle θ and at each surface 2 interface will similarly lead to some light escaping the back surface. In this normal incidence case we may consider these two beams together:

$$CR = 0R'\angle\theta + -2R'\angle\theta \tag{8}$$

The series in Eq. (7) may be inserted into an infinite series to account for all of the exiting beams T in boxes similar to gray box d.

$$\sum T_d = \frac{T_d}{1-(CR)}. \tag{9}$$

The sum of all transmission is that seen in gray box a given by Eq. (5) and from the infinite series of gray boxes d given by Eq. (9). The second term is doubled since we considered the -1T order at the first interface of gray box a, but an identical set of beams exist for the +1T case.

$$\sum T = T_a + 2\sum T_d\ . \tag{10}$$

In the above calculations all possible paths are accounted for with the exception of the normal incidence beam from inside the slab (in red in the figure) leading to first order modes when reflecting off the grating. These orders are small, and do not significantly alter the results.

Note that all transmitted and reflected light exits normal to the surface of the silicon, despite the presence of a grating that produces higher-order modes.

3.2 Reflectivity and Transmissivity Measurements of Near-Wavelength Design

Based on the initial reflectivity calculations with the low reflectivity shown in Figure 7, a near-wavelength design with a period of 1μm was fabricated and measured. Reflected and transmitted signals from a wet-etched sample were measured as in Figure 10. The experimental setup required a 10° incidence angle for reflection measurements, while normal incidence and 10° off-normal were used for transmission.

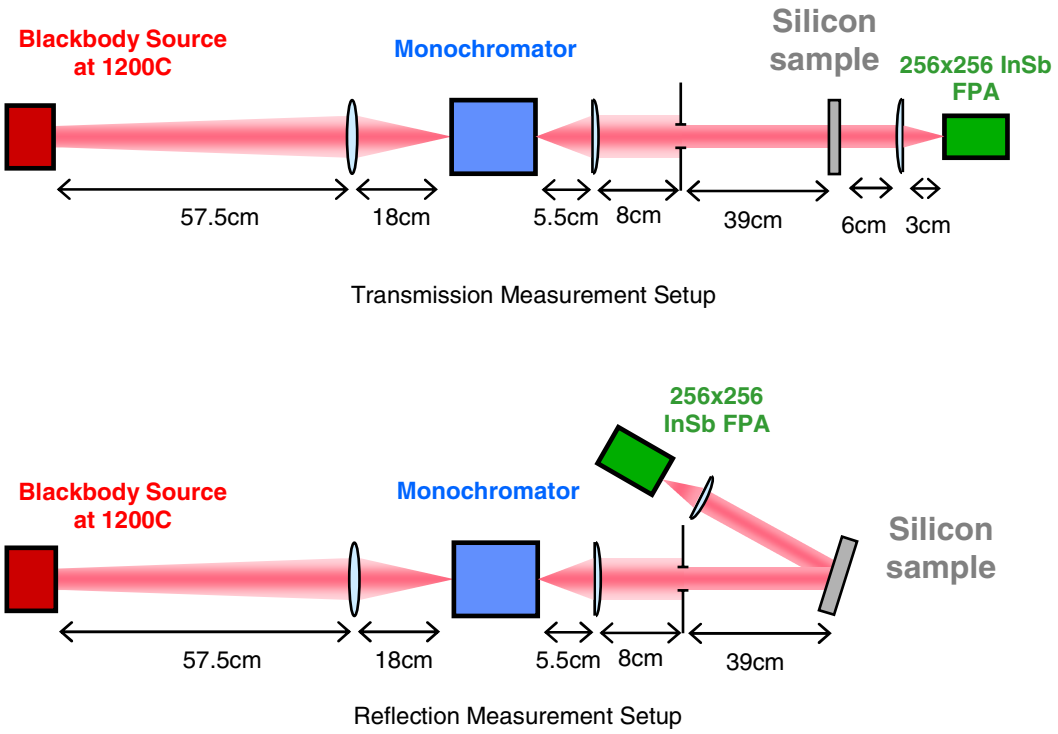


Figure 10. Experimental setup for transmission measurements (top), and reflection measurements (bottom). These schematics are not to scale, specifically, the reflected angle deviates less than 10 degrees from normal.

The light source used for the measurements was a blackbody operated at 1200°C. A 2 inch diameter 150mm focal length plano-convex lens was used to couple light into an Acton Research 275 monochromator with a 300g/mm grating and a 3.1μm blaze. The input slit was set at 100μm and the exit slit was set at 50μm. The light exiting the slit was collimated with a one-inch diameter 75mm focal length plano-convex lens. An iris limited the resulting beam to a 5mm diameter. This beam projected onto the structured surface, and the light transmitted through imaged onto a liquid nitrogen cooled 256 X 256 InSb focal plane array (FPA) with a one-inch diameter 50mm focal length plano-convex lens.

The FPA was set up and corrected for uniform gain. The target was aligned normal to the beam. Starting at 1900nm, data was collected in 100nm steps to 4000nm. For each wavelength the following data was collected: first, a background image (light blocked at input side of the monochromator), then a second image was collected with the signal light. The background image was then pixel by pixel subtracted from the signal image.

This background subtracted image was collected for

1. no sample (just air) for a 100% transmission reference.
2. the beam projected onto the structure (AR structure area is 5mm X 10mm).
3. the beam projected onto an area of unstructured Si on same wafer.
4. the beam projected onto a sample of Si polished on both sides.

The target was rotated 10° off normal, and the data collected again with the same steps. The setup was changed to measure the amount of light reflected from the target surfaces. The InSb

FPA and imaging lens were moved and aligned so that the light reflected from the target surface was imaged onto the FPA. The data was then collected in the same manner as the transmitted data.

Measured data for the transmission are shown in Figure 11 and Figure 12. The results are non-optimal in that the transmission is actually lower for the structured surface than for the polished silicon wafer. This result can be attributed to the multiple pass of the higher orders discussed previously. Similarly, the reflection at 10° incidence in Figure 13 is higher for the structured surface than for the polished silicon.

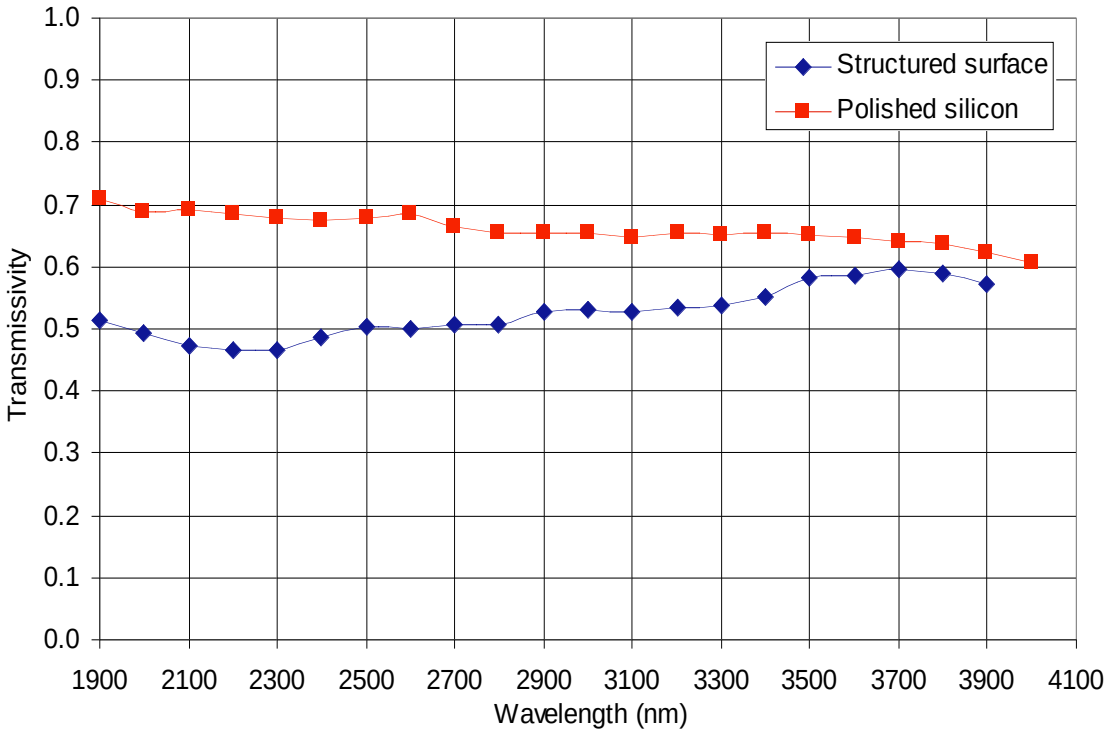


Figure 11. Measured transmission at normal incidence for polished silicon wafer and wafer with 1μm period surface structure on one side.

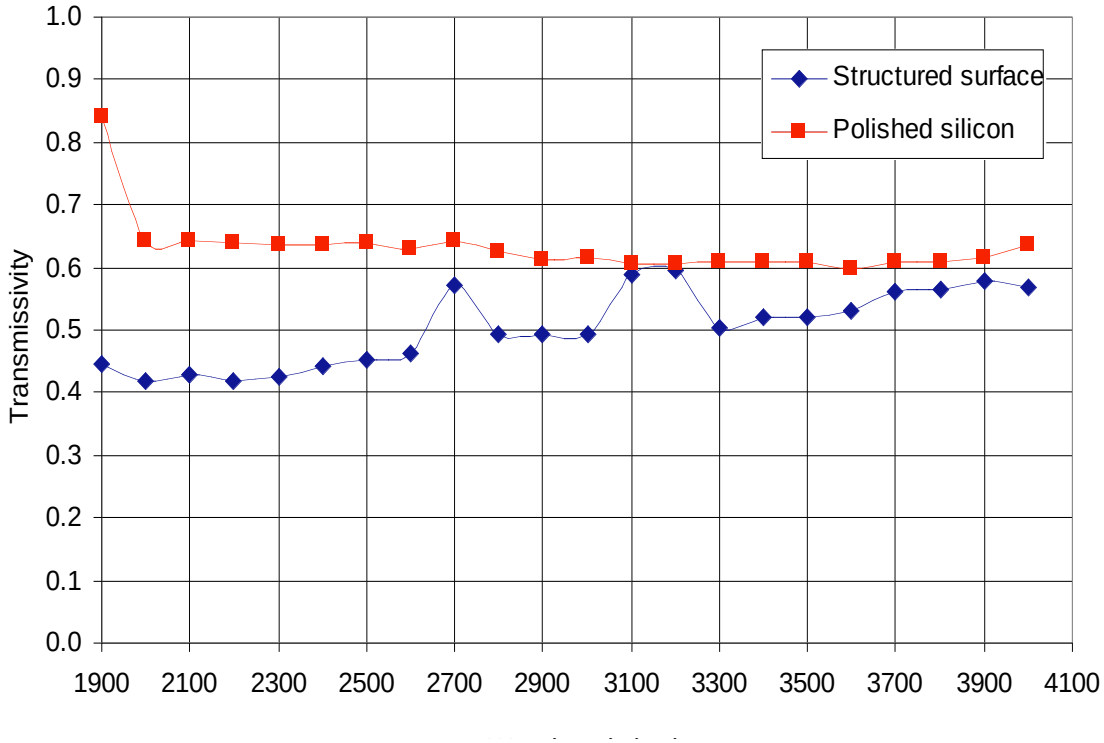


Figure 12. Measured transmission at 10° incidence for polished silicon wafer and wafer with 1μm period surface structure on one side.

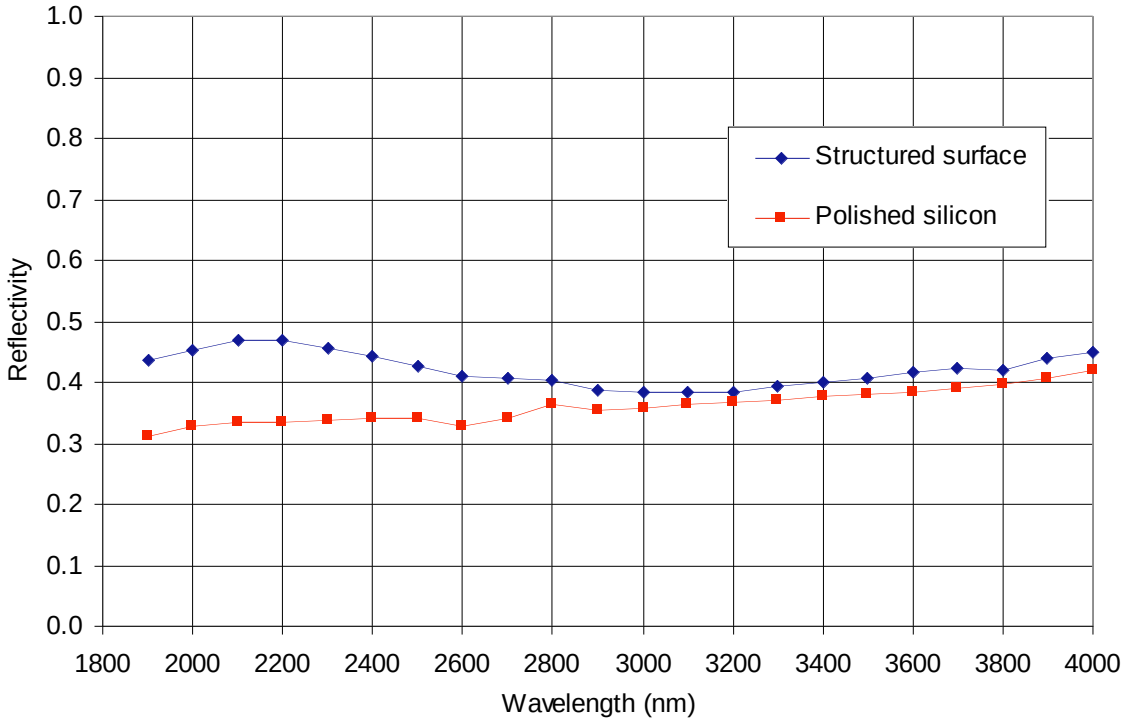


Figure 13. Measured reflection at 10° incidence for polished silicon wafer and wafer with 1μm period surface structure on one side.

Experimental results for transmission at 0° agree quite well with the calculated values at the mid to upper end of the wavelength range as shown in Figure 14. At the lower end there is a divergence, as the calculated transmission falls below the measured transmitted power. At the lower wavelength end of the plot where several orders are present, numerical simulations varied greatly with small changes in the structure height and polarization. Variation between the shape of the structure used in calculations and the one measured is the probable source of the divergence of the two curves.

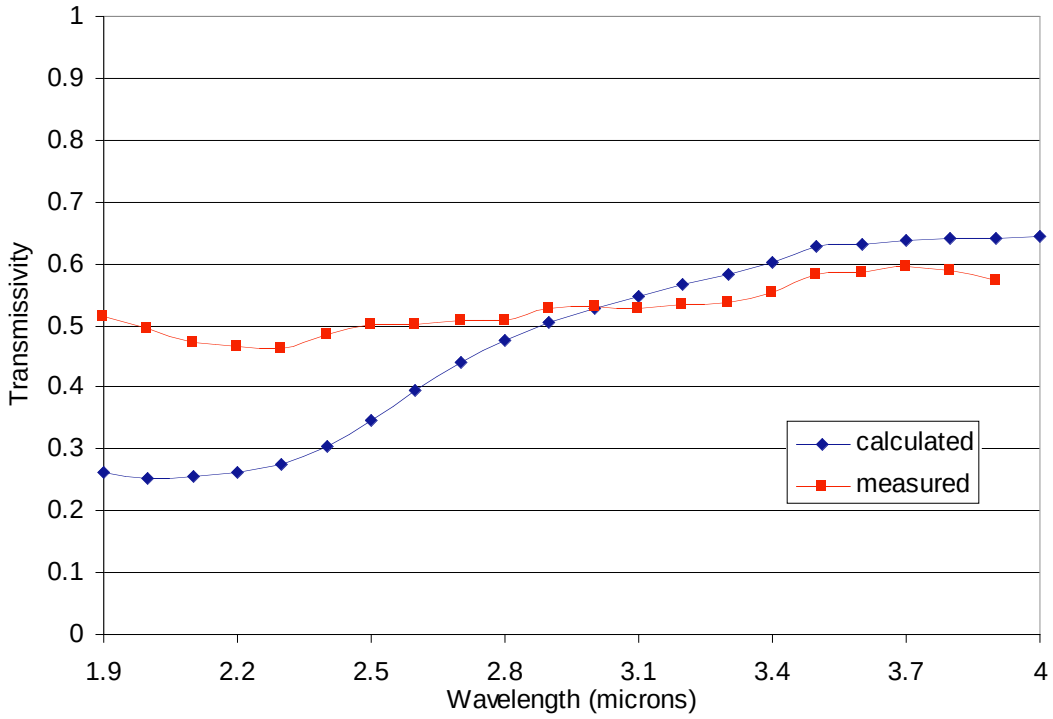


Figure 14. Comparison of calculated and measured transmissivity at normal incidence.

Overall, these measured results agree with the calculated predictions, that the higher-order modes in the silicon carry a significant portion of the transmitted signal and eventually result in an increased reflected signal. Although the longer-period device offered a lower reflectivity from the first surface, the higher-order modes increase the reflected signal.

3.3 Design Using a Subwavelength Period

Low transmission and high reflection on the sample described above is the result of the energy in the higher-order transmitted modes, the majority of which ends up as reflection. To avoid this problem, a grating with shorter period was designed. The shorter period means the grating is subwavelength compared to the wavelength of light in both the air and substrate. Higher-order modes are thus completely eliminated.

As the period is reduced, the depth is also decreased due to the fixed etch angle imposed by the crystalline structure of the silicon. Therefore the etch is not as deep as one would like for high-performance antireflection characteristics. However, we do not expect secondary reflection

resulting from higher-order modes in the silicon with this design. We thus expect a significant decrease in the reflectivity over bare silicon.

In Figure 15 we see that the zero-order reflectivity appears to decrease with increasing period, but, as we have previously discussed, the higher transmitted orders lead to greater overall reflectivity when the back surface is also considered. The optimal structure to reduce reflectivity is therefore the longest period structure that still does not allow higher-order transmitted modes in the 2-5μm range. This occurs at a period of 575nm. In Figure 15 we see that the 550nm curve (bold red) has a higher reflectivity than the longer period curves that we earlier considered in Figure 7, but we know that the reflected signal for the longer-period cases is actually considerably higher than the initial reflection that these curves indicate.

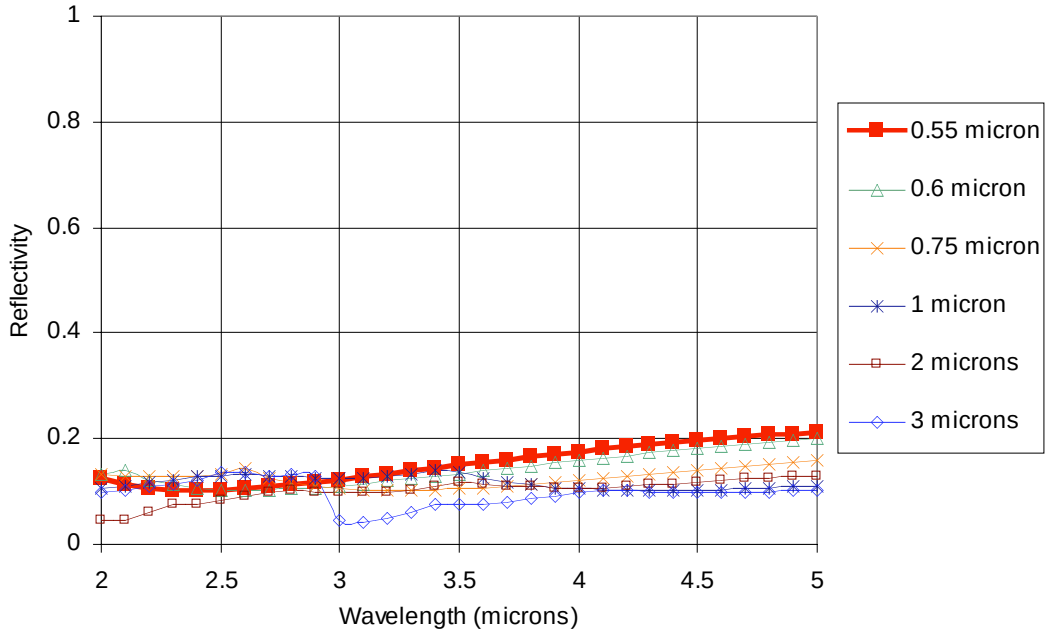


Figure 15. Sum of reflected orders from surfaced air/silicon interface for range of periods.

Fortunately, in Figure 16 we see that the zero-order transmission is quite high for the short period structures. Here we see that the transmission in the zero order is quite low for the longer-wavelength structures, as the experimental results indicated.

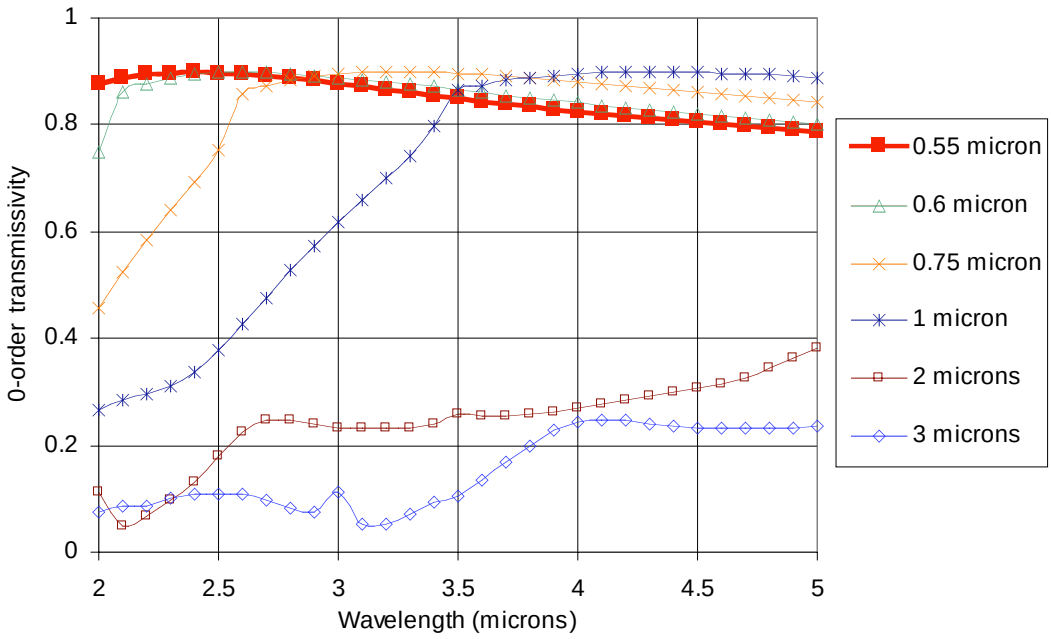


Figure 16. Zero-order transmissivity for surface air/silicon interface for range of periods.

4. Iterative Fourier Transform Algorithm (IFTA) for CGH Design

The diffractive nature of light allows one to design optical elements with unique properties. A diffractive optical element (DOE) is a component that modifies the wavefront of light via interference effects stemming from the direct control of the wavefront amplitude and phase. Most modern diffractive optical elements vary only the phase of the wavefront. Such elements are easier to fabricate since no spatially-varying amplitude modulation is required. Phase-only elements are also generally more efficient since, neglecting scatter, all of the incident beam is transmitted.

Many varieties of diffractive optical elements have been developed. One particular kind of DOE is considered in this report. A multi-level binary DOE is a surface whose height profile varies discretely both laterally and axially. Such elements are traditionally fabricated using standard photolithography and etching. When this process is used with N alignment and etching steps, an optic with 2^N different height levels can be made. Figure 17 shows the fabrication of DOE with N = 2.

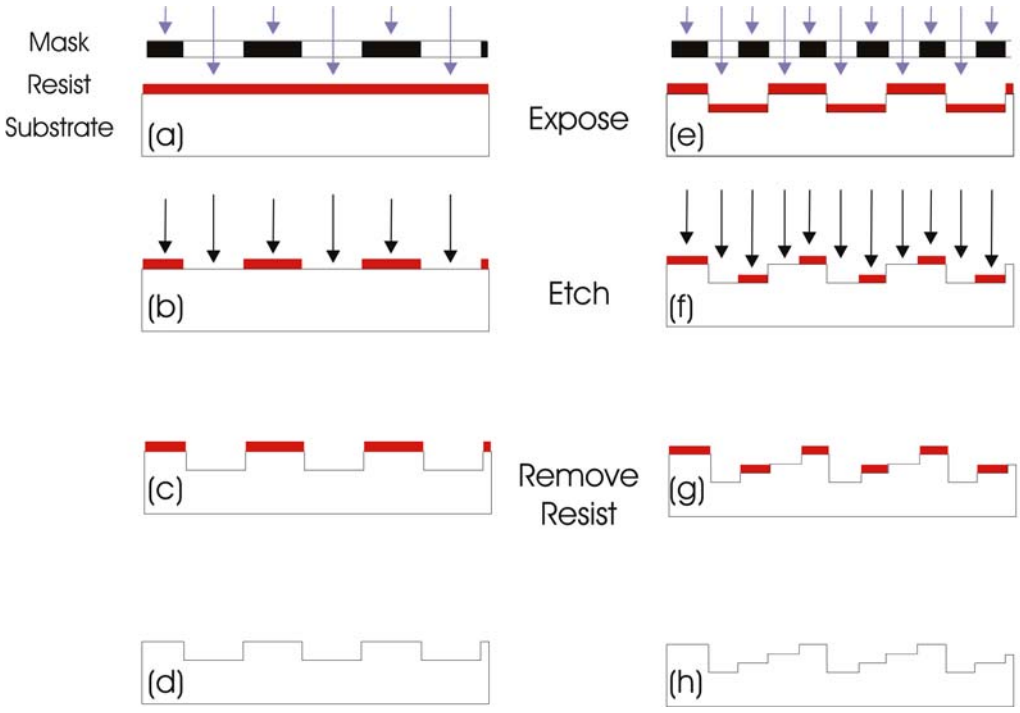


Figure 17. Fabrication of a 4-Level DOE.

Diffractive optical elements are used in a wide range of applications. The application considered here is the conversion of a uniform beam of light to a desired intensity distribution in the far field. Elements that perform this operation are often called computer generated holograms (CGHs).

The goal of CGH design is to calculate the height profile such that in the far-field the modified beam of light has an intensity distribution matching the desired pattern. This design process has been an active area of research for decades. The difficulty arises from the fact that this is an ill-posed inverse problem [1]. Ill-posedness means that one of the following conditions is true: 1) a solution does not exist, 2) the solution is not unique, or 3) the solution is unstable. For CGH design, the solution in general does not exist and multiple CGHs may yield the same diffraction pattern.

While a perfect solution cannot be found, approximations to the solution can. The nature of CGH design research is finding methods that yield improved approximations to the desired solution. A variety of different techniques have been studied, but the most widely used procedures are based on the iterative Fourier transform algorithm (IFTA, also commonly known as the Gerchberg-Saxton algorithm or the error-reduction algorithm).

4.1 Traditional IFTA

When a uniform beam illuminates the CGH, the far-field intensity distribution is proportional to the square of the modulus of the Fourier transform of the CGH amplitude transmittance function. The process also imposes constraints on the amplitudes in both the CGH plane and the far-field plane. In the CGH plane, we know that the transmittance has compact support and is a pure

phase function (the magnitude of the CGH transmittance is unity over the CGH, zero everywhere else). In the far-field plane we know that the amplitude distribution is equal to the square root of the desired intensity distribution.

These two constraints are implemented in the IFTA. In its most basic form, the IFTA consists of the following steps:

1. Determine an initial guess of the CGH transmittance function g. If necessary, g can be a random phase-only function.
2. Calculate the far-field intensity pattern G^2 by taking the Fourier transform of g.
3. Adjust G so that its amplitude distribution is equal to the square root of the desired intensity distribution. Do not change the phase of G.
4. Renormalize G so that it has the same energy as g.
5. Calculate an improved version of g by taking the inverse Fourier transform of G.
6. Adjust g so that it has unity amplitude over the CGH area. Do not change the phase of g.
7. Renormalize g to have the same energy as G.
8. Repeat steps 2-8.

It has been shown that every iteration of this algorithm reduces the mean square error (MSE) of the far-field intensity distribution. MSE is defined by:

$$MSE = \iint_S \left| |G(x,y)| - G_0(x,y) \right|^2 dxdy, \tag{1}$$

where S defines the signal area (the region where the desired intensity is non-zero) and $G_0(x,y)$ is the square-root of the desired intensity distribution.

Two other metrics are traditionally used when evaluating CGH designs. These are diffraction efficiency and non-uniformity. These metrics are most applicable to desired intensity patterns that take on only two values: zero and a constant.

Diffraction efficiency η is a measure of how much of the incident light goes to the signal area S. Non-uniformity NU is a measure of how much the final shape distribution varies from the desired distribution. Mathematically, these definitions are:

$$\eta = \frac{\iint_S |G|^2 dxdy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |G|^2 dxdy}, \tag{2}$$

$$NU = \frac{|G|_{\max} - |G|_{\min}}{|G|_{\max} + |G|_{\min}}. \tag{3}$$

The IFTA drives down the MSE. Note that the mathematical nature of the MSE differs significantly from η and NU. Thus this optimization need not arrive at the most uniform or most efficient solution. The correlation between MSE and both η and NU is rather weak and in general the IFTA will arrive at a solution with moderate uniformity and efficiency.

Figure 18 provides an example of the results of the traditional IFTA. The desired far-field pattern is the Sandia Thunderbird logo centered on a 32x32 array. The intensity distribution over the Thunderbird pattern has a uniform value and is zero everywhere else. The following figure shows the desired far-field pattern, the intensity pattern after the IFTA (also referred to as GS, for Gerchberg-Saxton algorithm) assuming that the CGH phase is a continuous function, and the intensity pattern after the IFTA assuming that the CGH phase is discretized to eight levels.

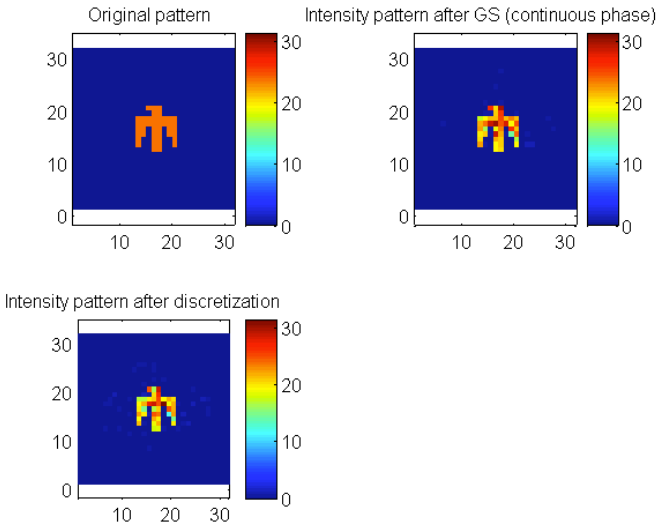


Figure 18. Desired and calculated intensity $|G|^2$ patterns for traditional IFTA.

Figure 19 shows the height distribution of the calculated eight-level CGH:

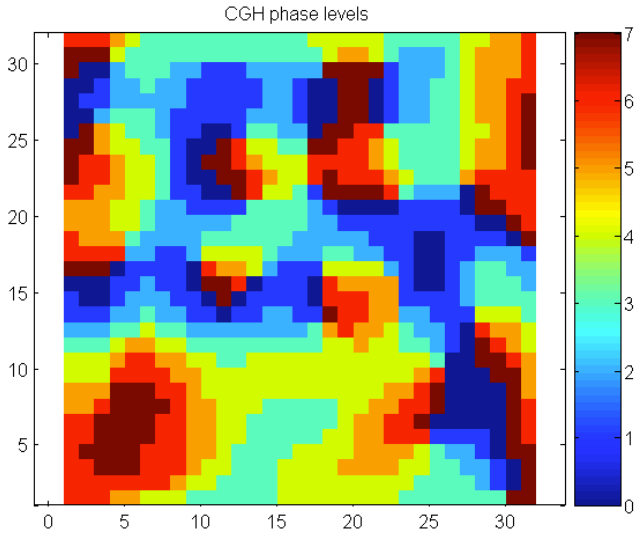


Figure 19. Traditional IFTA eight-level CGH.

Ideally, we would want a diffraction efficiency of 100% and a non-uniformity of 0. The continuous-phase CGH has a diffraction efficiency of 93.03% and a non-uniformity of 0.19. Discretization imposes a significant constraint on the transmittance function and generally

reduces both the diffraction efficiency and uniformity. The eight-level CGH has a diffraction efficiency of 88.72% and a non-uniformity of 0.25 (a higher non-uniformity corresponds to a less uniform distribution).

4.2 IFTA with Over-Compensation

There is a direct trade-off between diffraction efficiency and non-uniformity in CGH design. A design with lower diffraction efficiency allows more energy in the noise area. This introduces more freedom in the amplitude distribution which allows for improved uniformity. Conversely, a design with higher diffraction efficiency reduces the number of degrees of freedom in the amplitude distribution, thereby reducing the uniformity.

IFTA with over-compensation has been developed for adjusting the trade between diffraction efficiency and non-uniformity. One such algorithm is the so called over-compensation technique [2,3]. This technique first employs the traditional IFTA to provide a stable solution (a sufficient number of iterations are performed such that the solution converges). The over-compensation technique then performs several more iterations with modified far-field intensity constraints. A region δ is defined over the signal area S. δ is defined to be the region where the far-field amplitude is greater than a user-defined threshold value. Over δ, the magnitude of the far-field amplitude is replaced by an amplitude-symmetric distribution of the calculated amplitude about the desired amplitude. Outside of δ, the traditional IFTA constraint is applied. Furthermore, the phase of the far-field amplitude distribution is set equal to the phase of the amplitude distribution calculated with the final traditional IFTA step. Figure 20 shows the results when the traditional IFTA is followed with five over-compensation steps.

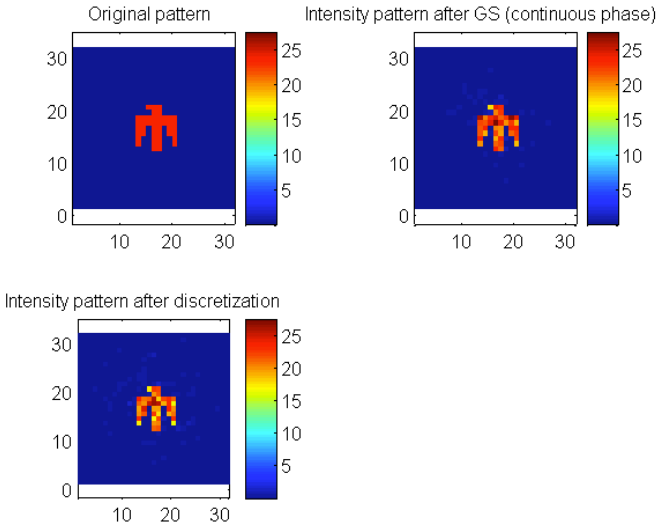


Figure 20. Intensity images for traditional IFTA followed by five over-compensation steps.

The continuous-phase CGH has a diffraction efficiency of 92.40% and a non-uniformity of 0.11. The eight-level CGH has a diffraction efficiency of 88.05% and a non-uniformity of 0.12. This technique has only slightly degraded the diffraction efficiency while significantly improving the uniformity.

4.3 IFTA with Noise at the Perimeter

Another technique used to adjust the trade between uniformity and diffraction efficiency involves allowing energy to be distributed outside a region surrounding the signal area. In this technique, the far-field amplitude is forced to zero in the noise region directly surrounding the signal area. Further away, the far-field amplitude distribution is left as-calculated. This allows energy to be deposited at the perimeter of the noise area and introduces new degrees of freedom in the amplitude distribution, thereby improving uniformity at the cost of diffraction efficiency. The results of this algorithm are shown in Figure 21.

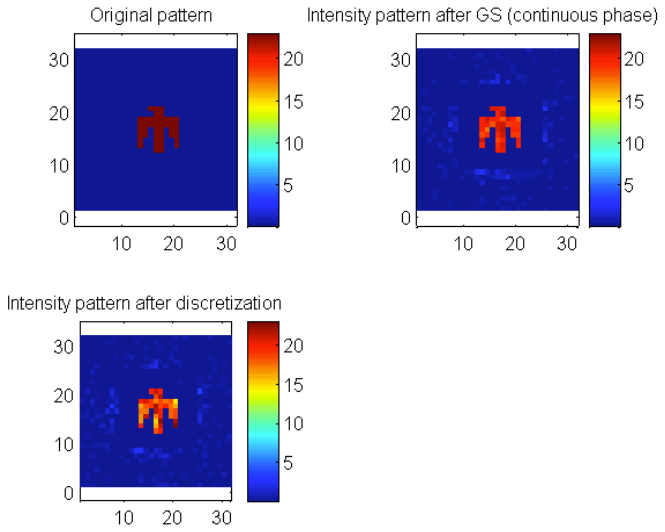


Figure 21. Intensity images for IFTA with modified constraints in the noise region.

The box where the amplitude in the noise region was forced to zero is clearly visible in both calculated intensity images. The continuous-phase CGH had a diffraction efficiency of 82.80% and a non-uniformity of 0.052. The eight-level CGH had a diffraction efficiency of 78.84% and a non-uniformity of 0.12. Compared to the traditional IFTA results, a significant amount of diffraction efficiency has been traded to improve uniformity.

4.4 IFTA with Scaling in the Noise Region

The last technique employed also changes the far-field amplitude constraint in the noise region. In this technique, the energy in the noise region is left unchanged when imposing the far-field amplitude constraint. The scaling of the entire far-field amplitude distribution is then modified so that energy is conserved. The results of this algorithm are shown Figure 22.

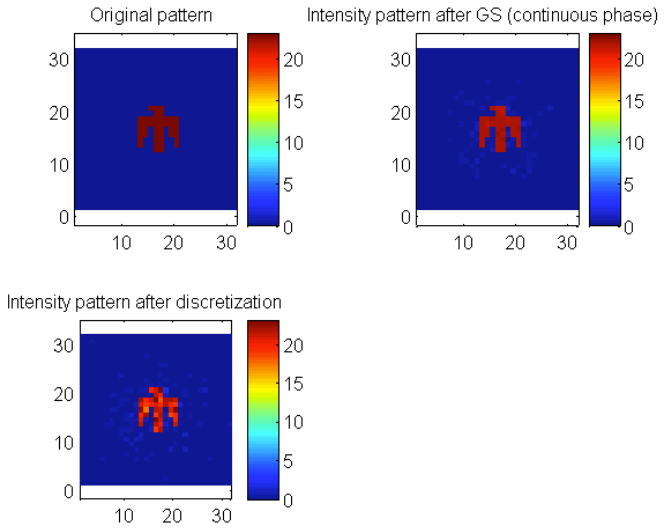


Figure 22. Traditional IFTA with scaling in the noise region.

This technique allows for more noise near the signal area when compared to the algorithm that constrained the noise region to be zero near the signal area. The diffraction efficiency of the continuous-phase CGH is 91.22% and the non-uniformity is 0.0087. The diffraction efficiency of the eight-level CGH is 86.81% and the non-uniformity is 0.07. Thus allowing noise closer to the signal area has further increased the number of degrees of freedom in the amplitude distribution and improved uniformity even further.

Another attractive algorithm was not employed because of time constraints. This technique uses an IFTA modified with Tikhonov regularization techniques and an adaptive regularization parameter distribution (ARPD) [4]. This technique provides near-optimal uniformity for a particular diffraction efficiency. If this work were to continue, implementing this algorithm would be a useful avenue of advancement.

The algorithms employed above show how diffraction efficiency and non-uniformity can traded. Table 1 summarizes the results.

Table 1. Summary of IFTA results.

	Continuous-Phase CGH		8-Level Discretized CGH	
Algorithm	Diffraction Efficiency (%)	Non-Uniformity	Diffraction Efficiency (%)	Non-Uniformity
Traditional IFTA	93.03	0.19	88.72	0.25
IFTA with over-compensation	92.40	0.11	88.05	0.12
IFTA with noise at the perimeter	82.80	0.052	78.84	0.12
IFTA with scaling in the noise area	91.22	0.0087	86.81	0.07

4.5 CGH Design Example for Uniform Bar

The previous examples were purely mathematical in nature; no consideration was given to such issues as the physical dimensions of the CGH and the desired far-field intensity pattern or the optical system under consideration. These issues were addressed in a pair of designs. These

designs were performed in only a single dimension for improved computational speed. Extrapolation to a second dimension can be done straightforwardly. The IFTA with scaling in the noise region was employed.

First, production of a uniform bar of light 500 microns wide was considered. This CGH is placed directly adjacent to a lens of focal length 62.5 mm and is illuminated by a uniform beam of light with a wavelength of 632.8 nm. A detector with 15.5 micron wide pixels is placed in the back focal plane of the lens (where the far-field intensity pattern is located).
Figure 23 shows this geometry.

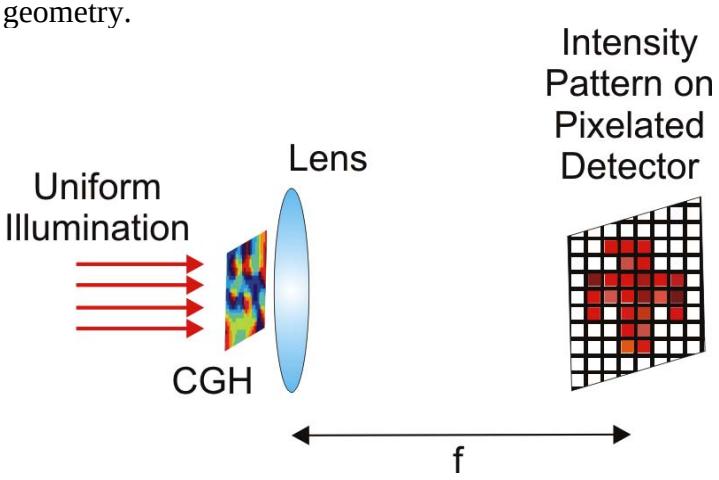


Figure 23: CGH optical system.

The CGH is discretized to eight phase levels. Each pixel is 20 microns across and the total width of the CGH is 1 cm. Such a CGH can easily be fabricated using available technologies. The final calculated CGH was over-sampled and padded to more accurately calculate the intensity distribution at the detector plane. This calculated distribution was smoothed by the point-spread function of the lens and then integrated over each detector pixel. Figure 24 shows the resulting profiles.

The continuous-phase CGH has a diffraction efficiency of 92.76% and a non-uniformity of 0.64. The eight-level discretized CGH has a diffraction efficiency of 89.40 and a non-uniformity of 0.60. The non-uniformities are high because of the left-most data point where the bar should be. This single point (with a value of approximately 0.1) skews the non-uniformity values. This demonstrates one problem with the non-uniformity metric: the non-uniformity for the entire signal region is based on the intensity values at only two discrete points.

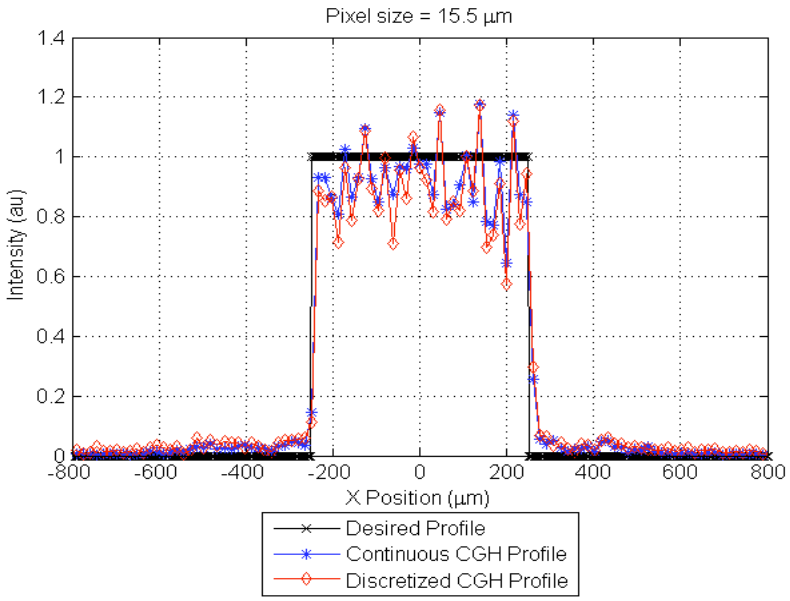


Figure 24. Intensity profiles for 1D bar.

4.6 CGH Design Example for 1-to-7 Fanout

Lastly, the fabrication of a so-called 1-to-7 fanout was simulated. This optic focuses the light to seven bars, each of which is 100 microns across. The center-to-center spacing of the bars is 250 microns. The specifications of this CGH are identical to the prior example except that 10 micron CGH pixels were used (the CGH width was fixed at 1 cm, so more pixels were used). Figure 25 shows the resulting intensity patterns.

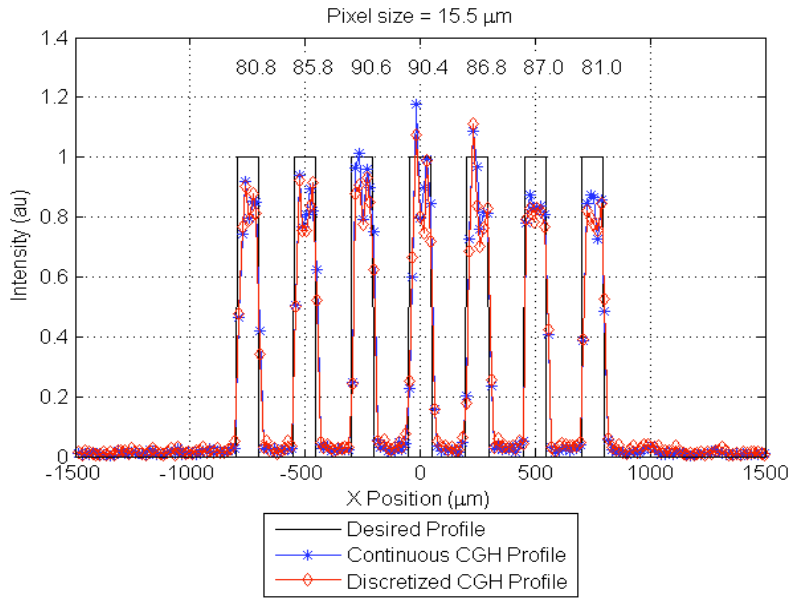


Figure 25. Intensity profiles for 1-to-7 fanout.

The number above each individual bar shows the diffraction efficiency for that bar (the amount of energy in the bar for the calculated profile divided by the amount of energy in the bar for the desired profile) as a percentage. The diffraction efficiency for each bar falls off with increased

distance from the center of the pattern. The overall diffraction efficiency for the continuous-profile CGH is 85.29% and the non-uniformity is 0.76. For the eight-level CGH these values are 82.01% and 0.75, respectively. Again, the non-uniformity values are affected by single points at the edge of each bar.

In this section we have discussed the design of a particular class of diffractive optical elements. The iterative Fourier transform algorithm was described and the trade-off between diffraction efficiency and uniformity was discussed. The IFTA and several variants were discussed and results were shown for each method. Each variant clearly demonstrated a different trade between diffraction efficiency and uniformity.

References

1 V. A. Soifer, V. V. Kotlyar, and L. Doskolovich, *Iterative Methods for Diffractive Optical Elements Computation* (Taylor and Francis, 1997).

2 J. Bengtsson, "Kinoform Design with an Optimal-Rotation-Angle Method," *Appl. Opt.* **33**(29), 6879-6884 (1994).

3 J. Liu, M. Thomson, A. J. Waddie, and M. R. Taghizadeh, "Design of Diffractive Optical Elements for High-Power Laser Applications," *Opt. Eng.* **43**(11), 2541-2548 (November 2004).

4 H. Kim, B. Yang, and B. Lee, "Iterative Fourier Transform Algorithm with Regularization for the Optimal Design of Diffractive Optical Elements," *JOSA A* **21**(12), 2353-2365 (December 2004)

Distribution List

1 MS 0972, Andy Boye, 5710
1 MS 0972, Carter Grotbeck, 5725
1 MS 0603, Jim Hudgens, 1713
4 MS 0603, Shanalyn Kemme, 1713
1 MS 0603, Rick McCormick, 1713-1
1 MS 0603, David Peters, 1713-1
1 MS 0570, Eric Shields, 5712
1 MS 0965, Jody Smith, 5712
1 MS 0603, Charles Sullivan, 1742
1 MS 0603, Allen Vawter, 1742

2 MS 9018, Central Technical Files, 8945-1
2 MS 0899, Technical Library, 9616
1 MS 0323, D.Chavez, LDRD Office, 1011