Two and three dimensional and gray scale images reconstructed from computer generated holograms designed using direct search method

Matthew Clark School of Electrical and Electronic Engineering,
University of Nottingham, University Park,
Nottingham NG7 2RD, UK.

A direct search method for the computer design of holograms is demonstrated. Two and three dimensional holograms and containing holograms with different levels of intensity (gray scales) have been designed, fabricated and optically reconstructed. The number of effective gray levels and the image intensity noise and contrast are discussed.

A modification to the *state variables* cost function used in the direct search algorithm which permits reliable control of gray scales is presented.

Optical reconstructions of 2 and 3 dimensional and gray scale binary phase computer generated holograms are presented.

1. Introduction

Computer generated holograms (CGHs) have many applications including laser ultrasonics[1], optical interconnects[2], laser machining[3] and beam shaping[4,5]. These applications require good efficiency, low noise and good contrast. CGH fabrication techniques are currently restricted to low numbers of phase levels (usually binary phase). Many design techniques produce poor results when used with low numbers of phase levels.

The direct search method presented here is a robust general method for designing CGHs which produce arbitrary images. It is not restricted to any optical geometry, hologram aperture or input wavefront. It is capable of producing CGHs with efficiencies close to the theoretical limit[6]. The method produces self-focusing phase quantised designs so the optical performance of the designs is very close to the simulated design results.

2. Direct search technique

The direct search technique[7–9] is a simple optimisation algorithm. The algorithm works by making small changes to the existing hologram phase and keeping these changes, depending on any improvement they make to the image. The phase is only permitted to take values allowed by the hologram fabrication process. Figure 1 shows a flowchart of the basic algorithm which is divided into starting, choosing, cost and stopping functions. These determine the initial solution; changes; cost, and when the algorithm stops. The CGH aperture and the image are sampled according to section 3 A. The optical amplitude at the image sample points is calculated from hologram phase distribution using a numerical model (see section 3).

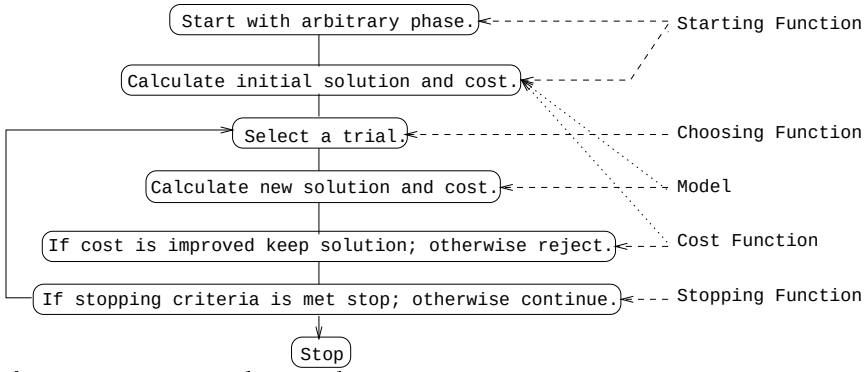

Fig. 1. A flowchart of a simple direct search algorithm

- **Starting function** The starting function is used to determine the initial state of the solution of the hologram phase. It has been demonstrated that starting the solution with no hologram samples emitting light (off), and then turning the holograms samples on when they make an improvement to the solution, can reduce optimisation time and may improve the reliability of the algorithm[10]. In this situation the hologram samples initially take an off state (not emitting light) that is additional to the usual states (on with the phase set to one of the allowed levels). Once a sample is turned on it must remain on and only its phase may be changed. This amounts to starting the optimisation process while the initial state of the image is determined from the initial state of the hologram.

- **Choosing function** The choosing function is used to select the hologram sample to be changed. This can be used to preselect samples that are more likely to result in improvement than a random choice and thus reduce the optimisation time. Exhaustive, random selection of hologram samples that lie on fringe boundaries can reduce the optimisation time without affecting the resulting image quality[7].

- **Cost function** A cost function is used to determine any improvement in the image. The cost function used in this paper is a modified version of the state-variables[7,11] which has been shown to be robust and superior to target and phase based cost functions. The normal state-variables cost function is given as

$$C = -a\bar{I} + b\sigma_I \tag{1}$$

  where C is the cost, a and b are positive "cost balancing" parameters, $\bar{I}$ is the mean intensity of the image (measured at the sampling points, see below) and is a measure of the image "signal strength" and σ_I is the standard deviation of the image intensity (at the sampling points) and is a measure of the image noise. The direct search algorithm tries to minimise the cost function. The algorithm shows good convergence[7] for $0.5 < b/a < 1.5$. Modifications have been made to the state-variables to allow the design of gray scale CGHs (see section 4).

- **Stopping function** The stopping function plays an important role in the direct search algorithm. The time taken for the algorithm to produce results of similar quality for the same design problem typically varies by as much as $\pm 20\%$ depending on arbitrary starting conditions[10].

  The algorithm is stopped when the rate of change of the hologram phase distribution or the probability of accepting a change falls below a preset value. This allows the algorithm to adjust the optimisation time and produces results with very similar performance despite arbitrary variation in the starting conditions[7,10].

3. Underlying model and sampling conditions

The algorithm can only make choices that are as good as the underlying model used to calculate the image amplitude from the hologram phase distribution. As a large number of small changes are made to the hologram phase distribution the model must allow efficient calculation of the effect of a change in hologram phase on the image amplitude. The contribution, $\triangle U_I$, from each hologram sample to the image complex amplitude, U_I, at each sample point is given using an approximation to the Kirchhoff integral by

$$\triangle U_I \cong \frac{-\mathrm{i}}{2\lambda} U_s \frac{e^{-\mathrm{i}(\phi_r+\phi_d+\phi_s)}}{R} O \triangle S \tag{2}$$

where λ is the wavelength of the incident light, U_s is magnitude of the amplitude of the wavefront at the hologram sample, ϕ_r is the phase difference due to the distance R between the hologram sample and the image sample, ϕ_d is the phase given to the wavefront by the design, ϕ_s is the phase of the incident wavefront, $\triangle S$ is the area of one hologram sample and O is the obliquity factor (taken as 1 for most designs).

A. Two dimensional sampling

Figure 2 shows a typical working layout. The hologram sample size is determined by the maximum angle of diffraction. In figure 2 this is determined by the angle required to diffract light to all parts of the image. The image only requires

sampling at the points of interests or where light is required. Outside these areas the intensity is reduced as the energy is diverted into the bright areas.

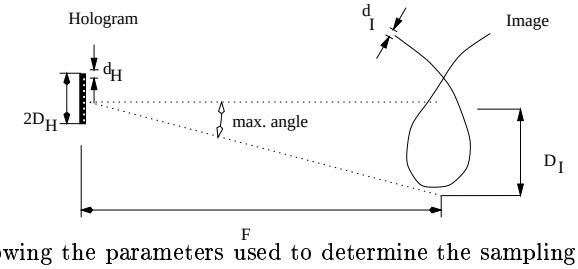

Fig. 2. Schematic diagram showing the parameters used to determine the sampling conditions. The hologram is at the left of the figure and the light travels from the left, through the hologram and then propagates a distance F to the image.

For low numerical aperture (NA) holograms the size of the hologram samples, d_H, is given by

$$d_H < \frac{\lambda F}{D_I} \tag{3}$$

where D_I is the size of the image (see figure 2) λ is the wavelength and F focal length (distance from CGH to focus). Equation 3 may require modification for large NA holograms to ensure adequate sampling. The size of the image samples, d_I, is determined by the NA of the hologram and is given by

$$d_I < \frac{1}{2}\frac{\lambda F}{D_H} \tag{4}$$

B. Three dimensional sampling

Three dimensional images require that the sample spacing along the optical axis, d_Z, is given by[12]

$$d_Z < \frac{\lambda F^2}{D_H^2} \tag{5}$$

This ensures that the samples are within one depth of focus of each other. Often the equation 4 will produce finer sample spacing in the z-direction than equation 5 and no additional samples will be required.

4. Gray scale modifications

The state-variables cost function shown in equation 1 is designed to produce images with a uniform intensity foreground. If the image is three dimensional or has different focal lengths in it the cost function still tries to equalise the intensity at all the image samples. Image samples that are further away from the hologram "see" a smaller hologram NA and this results in larger image point spread functions for these samples. This results in these image samples receiving more energy than samples nearer the hologram in order to equalise the peak intensity.

A. Simple weighting function

The relative distribution of intensity around the image samples may be affected by applying a weight to the calculated intensity of each sample. The cost function is then applied to the weighted intensity rather than the actual intensity. The cost becomes

$$C = -a\bar{I}_W + b\sigma_{I_W} \tag{6}$$

where the intensity, I, at the sample has been replaced by the weighted intensity, $I_W = I/W$ where W is the weight or required relative brightness of the image sample, $\bar{I}_W$ is the mean of the weighted intensity and σ_{I_W} is the standard deviation of the weighted intensity.

B. Ensuring convergence with gray scale images

Equation 6 only works for designs with a narrow spread of intensities. As the spread of intensity is increased the cost calculated with equation 6 can be reduced more quickly with single-hologram-sample changes if energy is diverted into the less bright image samples or the ones with the lowest weight. This can cause the developing solution to "skew" and favour the less bright images samples over the brighter ones and preventing the algorithm from converging to a satisfactory solution. This can be avoided by weighting the importance of the image samples to remove the false minima. The cost function becomes,

$$C = -a\bar{I}_{WW} + b\sigma_{I_{WW}} = -a\bar{I} + b\sigma_{I_{WW}} \tag{7}$$

where

$$\bar{I}_{WW} = \frac{\sum_{j=1}^{N} W_j I_j/W_j}{\sum_{j=1}^{N} W_j} = \bar{I}, \ \sum_{j=1}^{N} W_j = N \tag{8}$$

where W_j and I_j are the weight and intensity of an jth image sample, and N is the number of image samples. Note that the sum of the weights is normalised to the number of images samples. The doubly weighted standard deviation, $\sigma_{I_{WW}}$ of the intensity is

$$\sigma_{I_{WW}} = \sqrt{\frac{\sum_{j=1}^{N} W_j (I_j/W_j - \bar{I}_{WW})^2}{\sum_{j=1}^{N} W_j}} = \frac{1}{N}\sqrt{\sum_{j=1}^{N} W_j (I_j/W_j - \bar{I})^2} \tag{9}$$

The doubly weighted state-variable cost function (equation 7) is robust and can be used with a wide spread of intensities. It removes the false minima introduced by equation 6.

5. Optical reconstruction

Several CGHs have been fabricated at the nanostructure facility at the University of Glasgow. The holograms all had a 5.12mm aperture and were made up of 10μm pixels (hologram samples). The design was directly written onto the hologram by e-beam and dry etched into quartz substrate. The resulting etch depth was 610nm. The reconstruction wavelength was 633nm which required an etch depth of $\sim$ 650nm for maximum efficiency and consequently the reconstructions show a small residual zero order. This is visible as a circular patch of light at the centre of the images and would not be present if the CGHs had the correct etch depth.

A. Optical reconstructions

The images reconstructed from the holograms were projected onto white card and captured using a CCD camera.

Figure 3 shows three two dimensional images reconstructed from CGHs with focal lengths of 0.50m. The image sizes are $\sim$ 30mm and the reconstruction efficiency[7] was around 30 % giving a diffraction efficiency order of merit of around 40%[7].

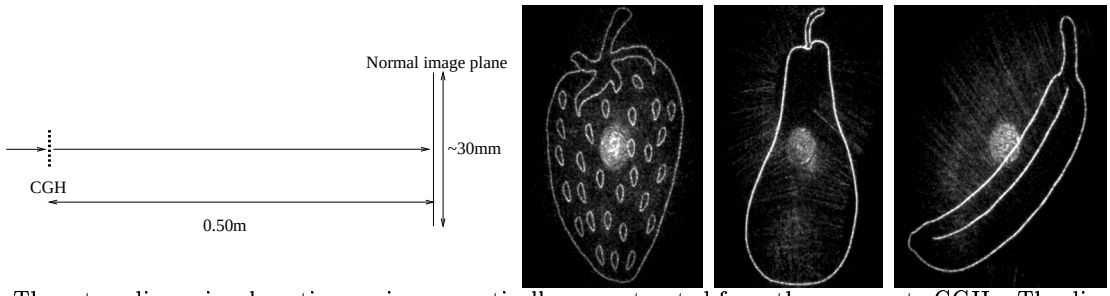

Fig. 3. Three two dimensional continuous images optically reconstructed from three separate CGHs. The diagram on the left shows the reconstruction geometry. The circular patch visible in the centre of the images is the residual zero order.

Figures 4 and 5 show reconstructions of CGHs designed to produce different images at different focal lengths. In figure 4 the images have been placed side-by-side so that the in-focus images are not obscured by the out-of-focus images. Figure 5 has the three images aligned along the optical axis so that the in-focus images are partially obscured by the out-of-focus images. The image clarity could be improved by spacing the images further apart along the optical axis so that the out-of-focus images are more diffuse.

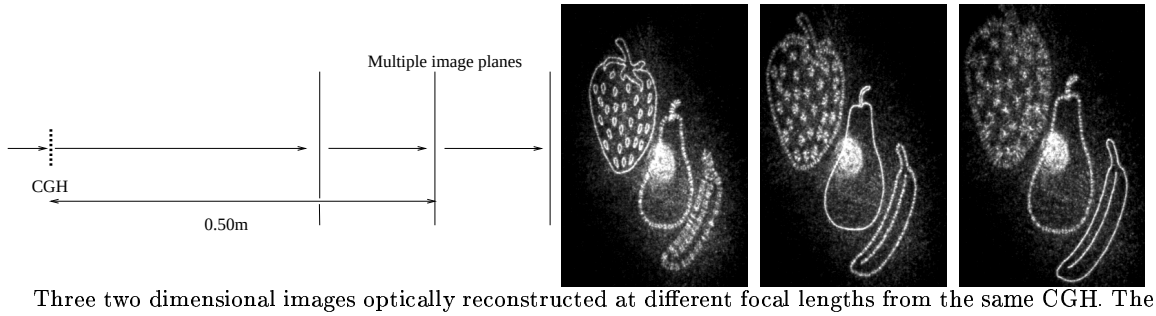

Fig. 4. Three two dimensional images optically reconstructed at different focal lengths from the same CGH. The diagram on the left shows the reconstruction geometry. The images are arranged such that the in-focus images are not covered by the out-of-focus images.

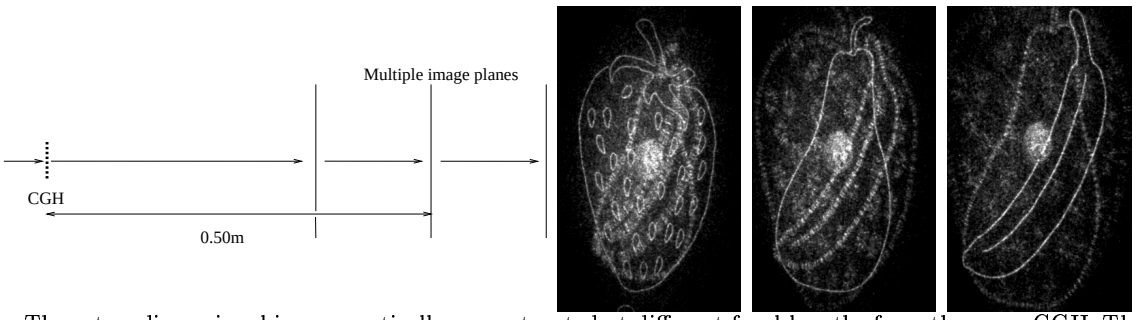

Fig. 5. Three two dimensional images optically reconstructed at different focal lengths from the same CGH. The diagram on the left shows the reconstruction geometry. The images are all centred on the optical axis so that the in-focus image is partially obscured by the out-of-focus images.

Figure 6 shows the reconstruction of a three dimensional image from a CGH. The choice of image surface for this demonstration was determined by the need to capture the image with a CCD camera. A plane surface inclined by 75° to the normal image plane was chosen. This allowed the three dimensional image design to have an image depth many times the depth of focus of the CGH and allowed the image captured in focus by a CCD camera (see the

diagram in figure 6). For comparison, a reconstruction of the three dimensional design shown on the normal image surface and reconstructions of a two dimensional design on a tilted and a normal surface are also shown.

It is clear that the three dimensional design is sharply focused over the tilted surface. The depth of the image over this surface is $\sim 5\times$ the depth of focus for the hologram. This three dimensional design does not reconstruct well over the normal image plane as no focal position can be found where it is all in focus. The two dimensional design cannot be brought entirely into focus on the tilted surface but focuses well over the normal image plane. The direct-search method is capable of producing designs that focus over arbitrary surfaces not just plane ones[11].

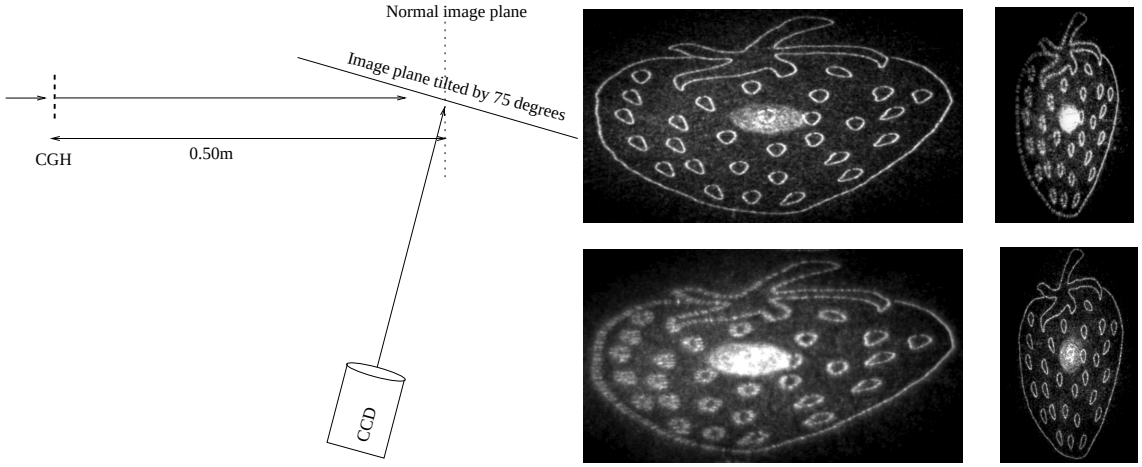

Fig. 6. Three dimensional continuous image optically reconstructed from a CGH. The diagram on the left shows the reconstruction geometry. The upper images are optical reconstructions of a CGH designed to produce a three dimensional image. The lower images are optical reconstructions of a CGH designed to produce a two dimensional image. On the left the images are reconstructed on a plane tilted by 75° to the normal image plane, on the right the images are on the normal image plane (perpendicular to the optical axis). The depth the image is $\sim 5\times$ the depth of focus of the CGH

Figure 7 shows a gray scale image and details of the image. The contrast between the lowest gray level and the highest is 10:1. This means that the areas with the lowest gray level are quite badly affected by the noise which can be seen in the details. The noise could be reduced by changing the position of the image with respect to the optical axis and separating the wanted +1 order from the unwanted -1 order (see section 6).

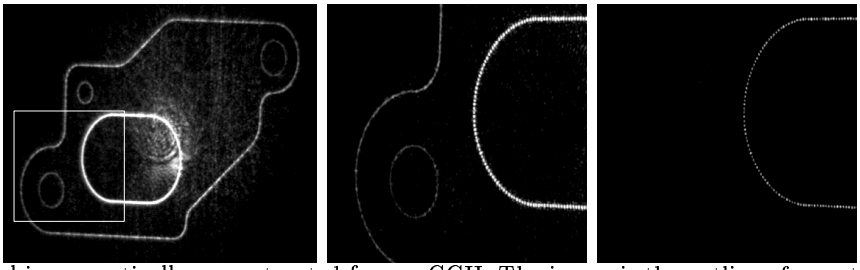

Fig. 7. A gray level image optically reconstructed from a CGH. The image is the outline of a motor vehicle gasket. This CGH contains three intensity levels with a contrast ratio between the least and most bright levels of 1:10. The images centre and right show details of the image on the left. The intensity has been reduced and the image magnified to show that the central feature is made up of discrete points. The reconstruction geometry is the same as figure 3.

6. Noise and contrast

Significant noise in the reconstructed image originates, in part, from the interference of the required image (+1 order) with other light diffracted by the holograms. In the case of a binary phase hologram the most significant contribution

to this unwanted light is the -1 order. This order appears to originate from a virtual focus positioned at -F with respect to the CGH (see figure 8). This contributes most of the energy found in the background of the image and coherently interferes with the desired foreground image. The mean intensity of the +1 order can be expected to be inversely proportional to the area of the image samples (bright regions). As the -1 order is the same as the +1 order defocused by two focal lengths ($2F$), its distribution is not random or even. However, a rough estimate of the mean intensity of the -1 order at the focal position of the +1 order can be made by considering the area the defocused image covers and assuming that the intensity is inversely proportional to the area.

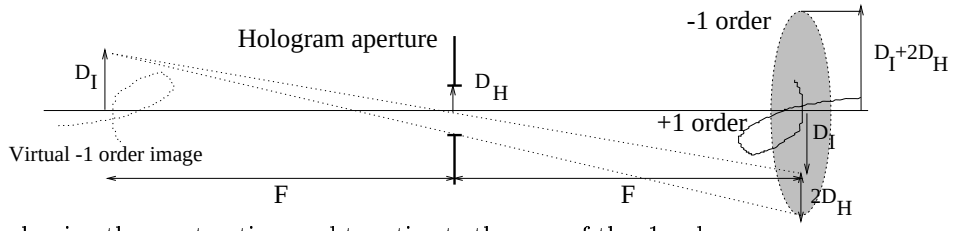

Fig. 8. Diagram showing the construction used to estimate the area of the -1 order.

The area covered by the +1 order can be estimated as $A_{+1} \sim \frac{1}{2}N\left[\lambda/(2\mathrm{NA})\right]^2$ where N is the number of image samples, λ is the wavelength and NA is the numerical aperture. The area covered by the -1 order at the focus of the +1 order can be roughly estimated as $A_{-1} \sim \pi(2D_H + D_I)^2$ where D_H and D_I are the aperture size and image size (see figures 2 and 8). The positive and negative orders contain the same energy (binary phase) and the contrast of the image (mean image intensity, $\bar{I}_{+1}$ / mean background intensity, $\bar{I}_{-1}$) can be estimated for a binary phase CGH as

$$C = \frac{\bar{I}_{+1}}{\bar{I}_{-1}} \approx \frac{A_{-1}}{A_{+1}} \sim \frac{\pi\left(2D_H + D_I\right)^2}{\frac{1}{2}N\left[\lambda/(2\mathrm{NA})\right]^2} = \frac{8\pi\left(2D_H + D_I\right)^2 D_H^2}{NF^2\lambda^2} \tag{10}$$

and the image noise for a binary phase hologram can be estimated as

$$\frac{\delta I}{I} \sim 2\frac{U_{-1}}{U_{+1}} \sim 2\sqrt{\frac{A_{+1}}{A_{-1}}} \sim 2\sqrt{\frac{1}{C}},\ \ \mathrm{SNR} \sim \frac{\sqrt{\mathrm{C}}}{2} \tag{11}$$

where U_{+1} and U_{-1} are the amplitudes of the +1 and -1 orders respectively and SNR is the signal to noise ratio. These measures are very approximate and should be taken only as guide. Particular designs may require different estimates for equations 10 and 11. as certain geometries can be used to prevent the +1 and -1 orders overlapping which will reduce the noise and improve the contrast in the required image area.

The signal to noise estimate assumes that there is a *random* phase relationship between the +1 and -1 orders. However this direct-search technique "sees" the -1 order and optimises the phase of the +1 order to match the -1 order at a particular point in space. However the phase distributions of the +1 and -1 orders change at different rates near the focus of either order so this phase matching does not occur at all positions along the depth of focus of the image[12].

Equations 10 and 11 agree within an order of magnitude with noise and contrast measurements taken from simulations of the CGHs presented in this paper. The biggest source of disagreement between the calculated and measured contrast and signal to noise arise from the assumption that the -1 order is distributed over a uniform disc. It can be seen from figure 3 that the intensity distribution of the -1 order is more complicated than this. Despite this, equations 10 and 11 provide a good rule of thumb for choosing the design parameters from the hologram performance requirements.

Increasing the number of phase levels results in more energy in the +1 order and less in the negative orders. Equations 10 and 11 would need to be modified for this and the negative orders may not be the dominant source of noise, especially with many phase levels.

A. Number of gray levels

The effective number of gray levels may be thought of as the maximum number of levels where the error introduced by quantising the gray scale is less than the fractional noise or the inverse of the contrast. Equations 10 and 11 give rough estimates for the contrast and noise for binary phase holograms. The actual noise and contrast will vary between designs and can be improved by choosing an appropriate reconstruction geometry.

7. Conclusion

The direct-search method is a robust and flexible method for the design of computer generated holograms. It is particularly effective when the number of phase levels is limited. It produces stand-alone designs which require no additional optics to reconstruct the desired image. The method is capable of producing two and three dimensional images and gray scale images. The dominant noise in the image results from interference between the desired image and the unwanted diffracted light. For binary phase holograms most of this unwanted light originates from the -1 order of diffraction. Estimates for the noise and contrast of images have been presented and are within an order of magnitude of measurements taken from simulated reconstructions. The image reconstruction noise may be reduced by separating the desired image from the -1 order of diffraction.

8. Acknowledgements

I would like to acknowledge the EPSRC (UK) and The Department of Electronics and Electrical Engineering, The University of Glasgow, UK.

---

1. M Clark, F Linnane, S D Sharples, and M G Somekh. Frequency control in laser ultrasound with computer generated holography. *Applied Physics Letters*, 72(16):1963–1965, 1998.
2. Arthur F Gmitrio, Paul E Keller, Christopher Coleman, and Paul D Maker. Design and fabrication of multi-level phase holograms for on-axis optical interconnects. In *Diffractive Optics: Design, Fabrication and Applications*, OSA Technical Digest series, pages 239–242. OSA, June 1994.
3. R G Hoptroff, P W McOwan, T J Hall, W J Hossack, and R E Burge. Two optimisation approaches to COHOE design. *Optics Communications*, 73(3):188–194, 1989.
4. Neal C Gallagher and Donald Sweeney. Computer generated microwave kinoforms. *Optical Engineering*, 28(6):599–604, 1989.
5. M T Eismann, A M Tai, and J N Cedrquist. Holographic beamformer designed by an iterative technique. *SPIE*, 1052(Holographic optics):10–18, 1989.
6. Brian K Jennison, D W Sweeney, and Jan P Allebach. Iterative approaches to computer-generated holography. *Optical Engineering*, 28(6):629–637, 1989.
7. Matthew Clark and Robin Smith. A direct-search method for the computer design of holograms. *Optics Communications*, 124(1-2):150–164, 1996.
8. Micheal A Seldowitz, Jan P Allebach, and Donald W Sweeney. Synthesis of digital holograms by direct binary search. *Applied Optics*, 26(14):2788–2798, 1987.
9. Brian K Jennison and Jan P Allebach. Direct binary search computer generated hologram: an accelerated design technique and measurement of wavfront quality. *SPIE*, 1052:2–9, 1989.
10. Matthew Clark. *A Direct Search Method for the Computer Design of Holograms*. PhD thesis, Imperial College, London, UK, 1997.
11. Matthew Clark. Enhanced direct-search method for the computer design of holograms using state variables. In *Diffractive and Holographic Optics Technology III*, volume 2689, pages 24–34. SPIE, 1996.
12. M Born and E Wolf. *Principles of Optics*, chapter 8.8. Pergamon Press, 5th edition, 1974.