

where i' and j' are the standard projected centred film coordinates. These integer values can be calculated by examining the limiting rays as illustrated by the example in Figure 9.8 for zones AA, AB and AC. From this diagram, we see that

$$Y_{\max} = \frac{h_{2w}}{h_{2w} - h} \zeta \quad (9.35)$$

This limiting value corresponds to the integer coordinate

$$j' = 1 + \left(\frac{g-1}{N_G - 1} \right) \frac{D_\zeta h_{2w}}{D_Y (h_{2w} - h)} (N_B - 1) \quad (9.36)$$

Using this and similar arguments for the other zones and the horizontal geometry, we are led to the following results:

$$\begin{aligned} {}^{1AZ} N_J &= {}^{1BZ} N_J = 1 \\ {}^{1CZ} N_J &= 1 + (N_B - 1) \left\{ 1 - \frac{D_\zeta}{D_Y} \left(\frac{N_G - g}{N_G - 1} \right) \frac{h_{2w}}{(h_{2w} - h)} \right\} \\ {}^{2AZ} N_J &= 1 + \left(\frac{g-1}{N_G - 1} \right) \frac{D_\zeta h_{2w}}{D_Y (h_{2w} - h)} (N_B - 1) \\ {}^{2BZ} N_J &= {}^{2CZ} N_J = N_B \end{aligned} \quad (9.37)$$

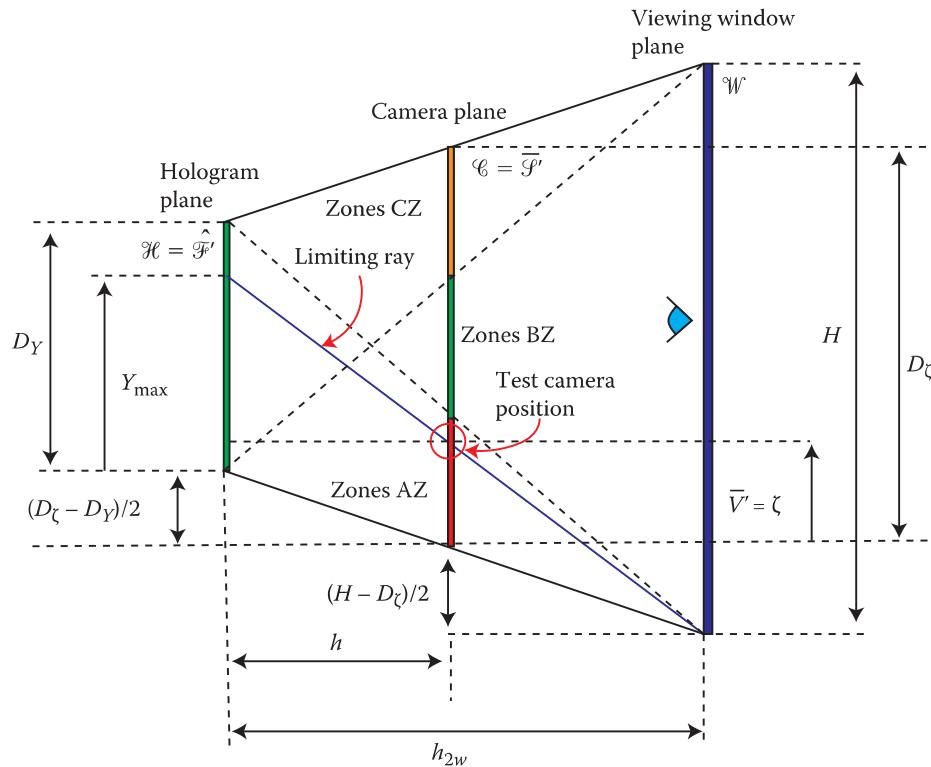


FIGURE 9.8 Side view of the ray geometry showing the limiting ray for a given value of g .

$$\begin{aligned}
{}^{1ZA}N_I &= {}^{1ZB}N_I = 1 \\
{}^{1ZC}N_I &= 1 + (N_A - 1) \left\{ 1 - \frac{D_\xi}{D_X} \left(\frac{N_K - k}{N_K - 1} \right) \frac{h_{1w}}{(h_{1w} - h)} \right\} \\
{}^{2ZA}N_I &= 1 + \left(\frac{k - 1}{N_K - 1} \right) \frac{D_\xi h_{1w}}{D_X (h_{1w} - h)} (N_A - 1) \\
{}^{2ZB}N_I &= {}^{2ZC}N_I = N_A
\end{aligned} \tag{9.38}$$

With the inequalities (Equation 9.34), these labels now define the subset of the centred film plane (i', j') required for each camera shot (k, g). This information may be programmed into a virtual camera in computer modelling software so as to produce image files of only the correct subset of the full centred camera frame.

Having rendered the data, we must then apply the general two-step processing scheme. The easiest way to do this is to define a cubic array in the main memory, which we can label $\mathcal{A}(g, i', j')$.^{*} We start with $k = 1$ and read into memory the files ${}^1g \mathbf{I}_{[i'j']}$ for all g . Now for every g outside of zone BB, the size of the file will not be $N_A \times N_B$ and so we will need to use Equation 9.37 to calculate the correct location within the $\mathcal{A}(g, i', j')$ to place each file. In general, we will load the files according to the rule

$$\mathcal{A}\left(i' + {}^{1ZZ}N_I(k) - 1, j' + {}^{1ZZ}N_J(g) - 1, g\right) = {}^k g \mathbf{I}_{i'j'} \tag{9.39}$$

with

$$\begin{aligned}
i' &\leq {}^{2ZZ}N_I - {}^{1ZZ}N_I + 1 \\
j' &\leq {}^{2ZZ}N_J - {}^{1ZZ}N_J + 1
\end{aligned} \tag{9.40}$$

Note that we have purposely labelled the camera zones using a two-letter label “ZZ” as the first label is defined by the index g and the second by k . Given g and k , one therefore knows the label “ZZ” and then Equations 9.37 and 9.38 may be used to calculate the various ${}^{ZZ}N$ values in Equations 9.39 and 9.40. This method does leave some blank spaces in the memory but because the array is overwritten by the next value of k , this is not a major concern. Of course, one may define a complex non-cubic array structure using Equations 9.37 and 9.38, but this is usually unnecessary.

Having now entered data into memory, we can simply write out the required intermediate files ${}^k [{}^g \mathbf{I}_i]_{j'}$. Like the original film files, these files are of variable size depending on the indices k and j . If we define a virtual plane \mathcal{J} as representing the plane spanned by the coordinates g and i or, equivalently, ζ and x , then we can draw a similar diagram to Figure 9.7 for these coordinates, but this time showing the subspace (i, g) instead of (i, j) . This is shown in Figure 9.9, where we introduce the new quantities ${}^{1ZZ}N_G$ and ${}^{2ZZ}N_G$, which are simply the inverse functions of ${}^{1ZZ}N_I$ and ${}^{2ZZ}N_J$:

$$\begin{aligned}
{}^{1ZZ}N_G &= 1 + \left(\frac{j' - 1}{N_B - 1} \right) \frac{D_Y (h_{2w} - h)}{D_\zeta h_{2w}} (N_G - 1) \\
{}^{2ZZ}N_G &= N_G + (N_G - 1) \frac{D_Y}{D_\zeta} \left(\frac{j' - N_B}{N_B - 1} \right) \frac{(h_{2w} - h)}{h_{2w}}
\end{aligned} \tag{9.41}$$

We then use the rule

$${}^k [{}^g \mathbf{I}_{i'}]_{j'} = \mathcal{A}\left(g + {}^{1ZZ}N_G(j') - 1, i' + {}^{1ZZ}N_I(k) - 1, j'\right) \tag{9.42}$$

^{*} This is simply a more convenient notation for $\mathcal{A}_{gij'}$.

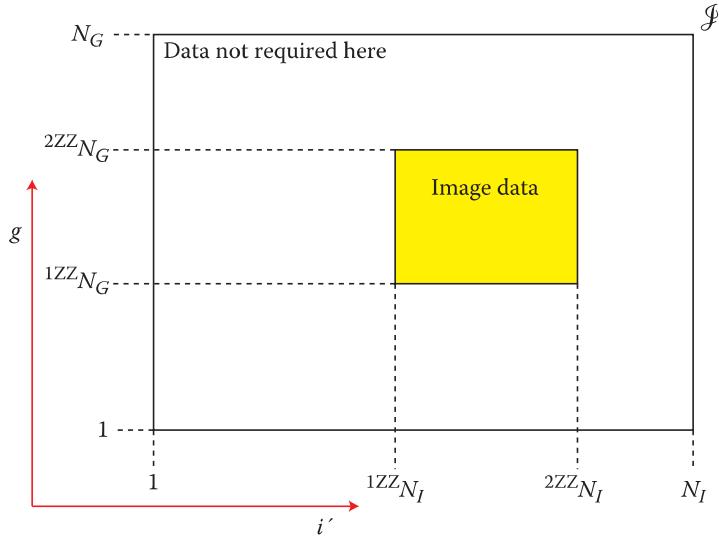


FIGURE 9.9 This diagram shows the virtual (x, ζ) plane, \mathcal{J} , corresponding to a camera located within a general zone, ZZ , of Figure 9.6. The yellow rectangle indicates the required subsection of \mathcal{J} that is required to be populated with image data. The white region of \mathcal{J} is not required and so does not need to be stored on disk.

with

$$\begin{aligned} i' &\leq {}^2ZZN_I - {}^1ZZN_I + 1 \\ g &\leq {}^2ZZN_G - {}^1ZZN_G + 1 \end{aligned} \quad (9.43)$$

to define all the intermediate files for $k = 1$. This process is then repeated for all remaining values of k , at which point we will have a full intermediate data set on disk.

The next stage is to define a new array, which we can label $\mathcal{B}(k, g, i')$. We load all $j' = 1$ files into this array using the rule

$$\mathcal{B}\left(k + {}^1ZZN_K(i') - 1, g + {}^1ZZN_G(j') - 1, i'\right) = {}^k \left[{}^g \mathbf{I}_{i'} \right]_{j'} \quad (9.44)$$

with

$$\begin{aligned} k &\leq {}^2ZZN_K - {}^1ZZN_K + 1 \\ g &\leq {}^2ZZN_G - {}^1ZZN_G + 1 \end{aligned} \quad (9.45)$$

Here, we have defined the inverse functions

$$\begin{aligned} {}^1ZZN_K &= 1 + \left(\frac{i' - 1}{N_A - 1} \right) \frac{D_x(h_{lw} - h)}{D_\xi h_{lw}} (N_K - 1) \\ {}^2ZZN_K &= N_K + (N_K - 1) \frac{D_x}{D_\xi} \left(\frac{i' - N_A}{N_A - 1} \right) \frac{(h_{lw} - h)}{h_{lw}} \end{aligned} \quad (9.46)$$

Finally, we can write out the final data set $\llbracket \bar{\mathbf{S}}_{\alpha\beta} \rrbracket = \llbracket {}^{kg} \mathbf{I}_{ij'} \rrbracket$ using the rule

$$\llbracket \bar{\mathbf{S}}_{\alpha\beta} \rrbracket = \llbracket {}^{kg} \mathbf{I}_{ij'} \rrbracket = \mathcal{B}(k + {}^{1ZZ}N_K(i') - 1, g + {}^{1ZZ}N_G(g) - 1, i') \quad (9.47)$$

with

$$\begin{aligned} k &\leq {}^{2ZZ}N_K - {}^{1ZZ}N_K + 1 \\ g &\leq {}^{2ZZ}N_G - {}^{1ZZ}N_G + 1 \end{aligned} \quad (9.48)$$

Upon repeating this last cycle for each value of the index j' , we will have completed the two-step optimised **I**-to-**S** transformation.

9.3 Resolution Requirements of DWDH Reflection Holograms

In Chapter 11, we shall review various effects that limit the practical resolution of display holograms. In particular, we shall see that the four main effects that limit the resolution of DWDH reflection holograms are source-size blurring (see Section 11.2.2 and 11.11), chromatic blurring (see Sections 11.2.3 and 11.11), speckle (see Section 11.3) and digital diffractive blurring (see Section 11.11.1).^{*} Source-size blurring is due to the intrinsic angular size of the illuminating source and, from a practical point of view, is not really affected by realistic emulsion characteristics. Chromatic blurring is due to the finite bandwidth of the illuminating source but is strongly controlled by emulsion thickness. Speckle is caused by too small a bandwidth in the illuminating source. Finally, digital diffractive blurring is caused by hogels that are generally small and thus display inherent diffraction.

We shall see in Chapter 11 that the human eye ultimately controls the best observable resolution of any display hologram. Although this varies from person to person, from a practical point of view, we can say that the human eye can just resolve two lines spaced around 1 mm from the other at a distance of 1 m. This corresponds to an angle of $\Delta\theta \sim 0.06^\circ$. When we use an illuminating light source having too large a physical diameter, too broad a spectrum or too narrow a spectrum, the resolution of the hologram will fall below the limit discernible by the human eye. However, when we choose the illumination source correctly, the observable resolution of the hologram will only be determined by the resolution of the human eye. We shall now examine the values of various critical digital parameters that are required to meet this most stringent limit.

We shall start by making the observation that if the hologram is to be viewed from a distance of more than 1 m, then a hogel size of less than 1 mm will not be resolvable by the average human eye. With HPO holography, in which the camera plane and viewing plane are collocated, we can make use of this fact because, effectively, there is no point in using a smaller hogel diameter than is resolvable by the average human eye at the viewing plane. After all, when one observes an HPO hologram from significantly closer than the viewing plane, the image effectively falls to pieces. Seeing that one seldom wants to make the viewing distance less than approximately 40 cm in HPO holography, a hogel size of approximately 0.5 mm is usually adequate in any situation.

The situation with full-parallax holograms is, however, completely different because these types of holograms can be viewed without distortion at any distance. One therefore often wants to reduce, as much as possible, the hogel size in full-parallax holography. Unfortunately, as we shall see in Chapter 11, digital diffractive blurring increases as the hogel size is diminished. For instance, a square hogel 0.1 mm in diameter will exhibit a diffractive blurring angle of approximately four times greater than the resolution angle of the human eye. The hogel size starts to limit the optimal resolution at diameters

^{*} This ignores insufficiencies in the perspective image data, printer SLM insufficiency, aberrations induced by chemical processing and printer optical aberrations.

of approximately 0.5 mm; below this value, one simply has to choose between good resolution close-up to the hologram at the expense of induced blurring at an ever-diminishing depth. This is a fundamental constraint of DWDH holography because the hologram is composed of a matrix of hogels that have all been recorded incoherently.* As such, a full-colour, full-parallax DWDH hologram will never quite have the resolution of a high-quality mirror or a physical glass window. In these cases, almost perfect images of zero to infinite depth may be seen at all distances. In the case of the DWDH hologram, which displays essentially perfect images over a range of viewing distances from infinity down to around half a metre, observation at a closer distance will inevitably reveal a substandard resolution at the hologram surface. Nevertheless, such holograms are capable of producing extremely impressive displays when illuminated with the proper illumination source. One should also point out that too large a hogel size produces a loss of resolution, predominantly at the hologram surface. When one focuses on a distant object with one's head close to the hologram, then the finite size of the hogels simply acts as an out-of-focus grill through which an effectively perfect image is seen. By paying close attention to how the hogels are written, this "grill" may be greatly minimised.

In addition to controlling the hogel size, the resolution of the human eye also puts a constraint on the number of camera views, N_K and N_G , and on the dimensions of the printer SLM, N_M and N_V . If the camera plane is taken at an h value of 1 m, then clearly we will require a spacing between camera positions of a little less than 1 mm to accommodate the average human eye. The maximum values of N_K and N_G are therefore given by dividing the window size by this distance. For a metre square hologram with a rectangular camera plane at a distance of 1 m and of a size of say 1 m × 2 m, this then equates to a camera array of $N_K \times N_G = 2000 \times 1000$ positions.

The number of pixels required of the printer SLM is similarly constrained. At a distance of 1 m, the SLM should project an image of itself with a characteristic spacing of a little less than 1 mm. In other words, two adjacent pixels should be separated by a little less than 0.06°. Given that we usually want a total horizontal field of view of around 100°, we can estimate that N_M should ideally be around 1700.

Of course, when we design a hologram to be illuminated by a light that has a greater than ideal source size or a greater than ideal bandwidth, we can relax the amount of digital data required. Values for N_K and N_G and also for the required pixel dimensions of the SLM must then be calculated using the formulae for induced blurring, which we cover in Chapter 11.

9.4 DWDH Transmission Holograms

The most successful type of ultra-realistic full-colour digital display hologram in commercial production today is the reflection hologram. Various companies including the US company Zebra Imaging Inc., the Canadian company XYZ Imaging Inc., and the Lithuanian firm Geola UAB, have all produced high-resolution large-format full-colour holograms showing many metres of image depth. On the other hand, the older large-format laser transmission holograms created by the likes of Nick Phillips and Paula Dawson produced perhaps even more stunning and deeper images using monochromatic analogue holography.

With the advent of modern laser diodes and light-emitting diodes, digital laser transmission holography now constitutes a promising and yet largely unexplored field potentially capable of producing spectacular full-colour displays. The digital image processing that we have reviewed in this chapter and in Chapter 8 is exactly the same for the monochromatic DWDH reflection hologram as it is for the full-aperture monochromatic DWDH transmission hologram. The only difference is that the reference beam must be on the same size as the object beam in the transmission scenario and the appropriate equations relevant to transmission holograms derived in Chapter 11 should be used for image predistortion.

* It is worth noting here that, in principle, one can record the hogels coherently by ensuring that the printer is interferometrically stable. This usually requires a slow step-and-repeat system incorporating fringe-locking technology. For most practical display applications, the recording of such mutually coherent hogels is just too slow.

The great advantage of the transmission hologram over the reflection hologram is that the illumination light is behind the hologram rather than in front of it. By using a compact recording geometry in which the illumination light is injected into the hologram at an angle greater than the critical angle,* it is possible to make sure that no zeroth order light is transmitted by the hologram on replay. This means that a hologram may be illuminated by laser and yet no harmful reference beam makes it through the hologram. The fact that the geometry is compact further means that a compact rear-illuminated display can be practically constructed displaying near-perfect image quality. This is likely to be a deciding factor in future applications such as integrated holographic windows, walls and floors.

Recording full-colour digital images using transmission holography of course requires a slightly different technique from that used in reflection holography. Because transmission holograms of a “normal” thickness are much less discriminating in wavelength than reflection holograms, the technique of using collinear RGB reference recording and reconstruction beams leads simply to multiple parasitic images with transmission holography. In Chapter 11, however, we shall show that a typical transmission hologram recorded on a typical silver halide emulsion of 7 µm thickness can be expected to exhibit a reference angle discrimination of approximately 15°. This is sufficiently small to allow the different coloured reference beams at recording and illumination to be arranged at differing altitudinal angles. Alternatively, the reference beams may be arranged at substantially differing azimuthal angles.

Transmission holography is likely to represent the technique of choice for future real-time holographic displays. Reflection holography has two disadvantages: first, the reference beam is in front of the hologram and so one cannot construct a fully integrated unit, and second, the reflection grating is intrinsically thick. Future submicron real-time SLM displays might realistically be expected to simply replace the fixed emulsion in transmission holographic technology, effectively converting the ultra-realistic static holographic display into a true 3D holographic television display. Colour control could then be performed using time multiplexing.

9.5 MWDH Reflection Holograms

In this chapter, we have concentrated on the DWDH technique and in particular on computational methods for preparing the required image data for this type of hologram. Arguably, DWDH is the most appropriate technique to write large-format full-parallax holograms. The reason for this is that one only needs a small laser to write each small hogel and it is easy to build such lasers to operate at relatively high repetition rates. We have seen that copying full-colour DWDH reflection holograms can be done effectively in some cases using a quasi-contact method—but at the time of writing, there are still some problems in commercialising such copying technology. As a result, DWDH holograms remain expensive.

Both full-parallax and HPO digital holograms may be generated via the MWDH technique. Before finishing this chapter, we will therefore say a few words about MWDH holograms for completeness. There are advantages and disadvantages to the MWDH technique. First, if a large hologram is to be produced, then a large RGB copy laser is required to convert the MWDH hologram to the final white light-viewable reflection hologram. Unfortunately, such lasers, although possible to build, are still very expensive at the time of writing. The speed at which copies can be produced is, however, usually far superior to DWDH combined with a contact-copy scenario. Furthermore, with MWDH, as we saw in Chapter 8, no computational change to the image plane is required—thus camera data can be used immediately in the printer without lengthy processing.

In small to medium-sized holograms, MWDH can indeed sometimes offer a superior solution over DWDH because a large costly RGB laser is not required and the intrinsic distance-copy process of

* This is usually arranged by writing the DWDH hologram using a large reference angle and then inducing an emulsion shrinkage upon processing. Replay of the hologram then requires side-illumination with the reference replay angle greater than that required for total internal reflection.

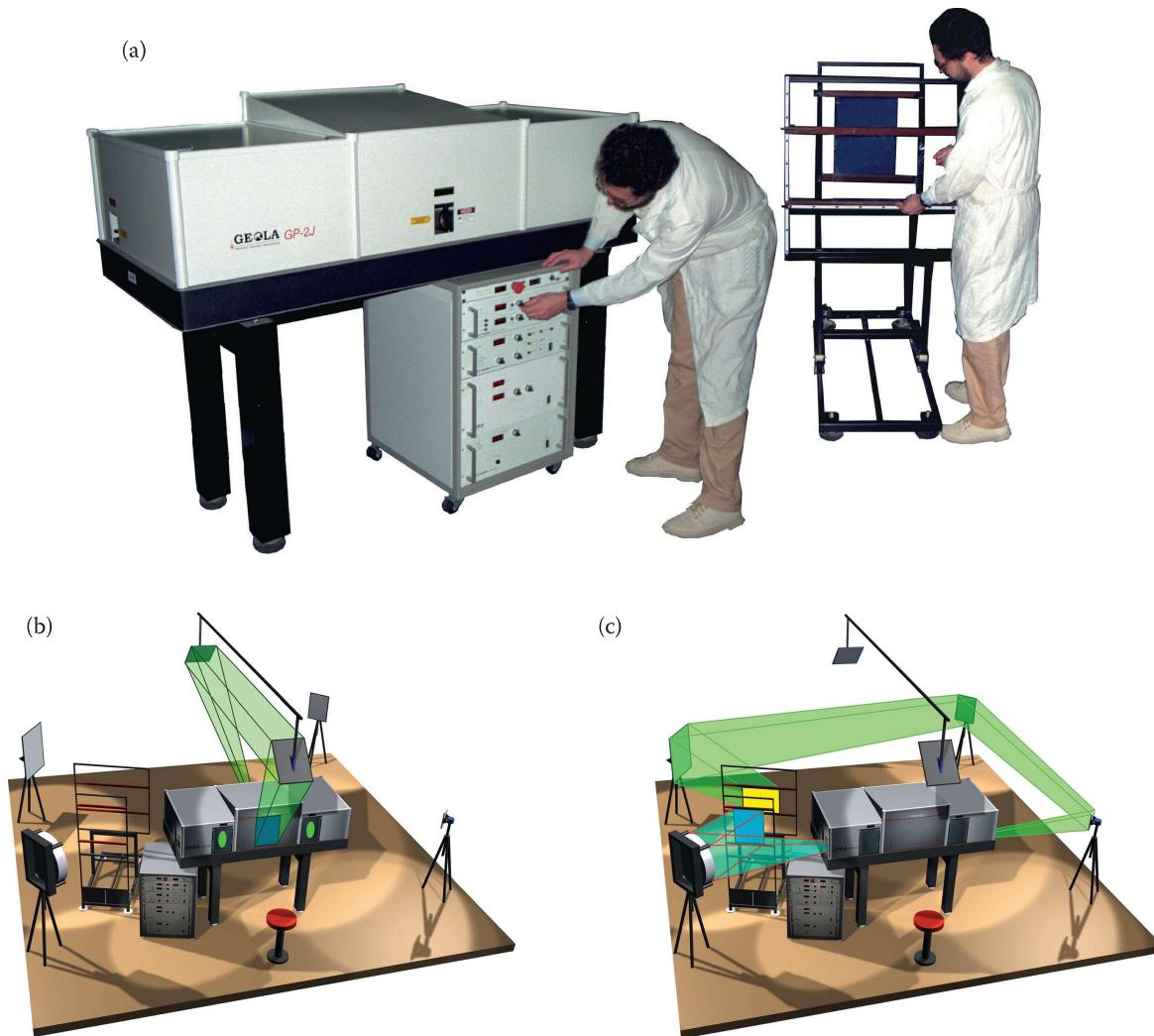


FIGURE 9.10 Photograph of a commercial large-format analogue holographic camera (made by Geola Digital UAB). The system (a), which is based on a 2J Nd:glass laser emitting at 526.5 nm, is capable of recording analogue master holograms as well as producing monochromatic reflection, achromatic or rainbow transfers of analogue or MWDH masters to an H_2 copy size of up to $0.8\text{ m} \times 1.0\text{ m}$. Diagrams (b) and (c) illustrate the operation. The creation of an analogue transmission H_1 master hologram is shown in (b). A full-aperture copy of an H_1 to a monochromatic reflection H_2 is shown in (c). Geola makes larger models based on 8J lasers capable of producing H_2 copies to $1.0\text{ m} \times 1.5\text{ m}$. (Photographs courtesy of Geola Digital UAB.)

MWDH offers more control over the quality and brightness of the final H_2 hologram.* Large-format MWDH HPO holograms of the achromatic and rainbow variety are also of significant interest as they can be copied to an H_2 hologram using commercially available large-format copying systems based on neodymium glass lasers (Figure 9.10).

9.6 Full-Parallax and HPO DWDH

We now come to the end of our discussion of the data preparation issues concerning full-parallax and HPO digital display holography. We have seen that the full-parallax and HPO cases are first and foremost distinguished simply by the amount of data implicit in the definition of the hologram. High-definition metre-square HPO holograms are usually characterised by image data of several gigabytes, whereas

* Ultimately, the choice of photosensitive material plays a vital role in defining the optimal technique for printing a digital hologram.

the corresponding full-parallax case is characterised by image data of several terabytes. However, we have also seen that there is a fundamental optical difference between HPO and full-parallax holograms. HPO holograms only present a faithful reproduction of the recorded image at a unique line in space that coincides with the camera line. Full-parallax holograms do not suffer from this restriction and, as such, at least theoretically, an aberration-free image is available at any point within the emission region of the hologram.

We have also seen in Chapter 8 how, in general, the various discrete coordinate meshes that define the image data, the digital hologram itself and the writing of the image data by the printer do not line up. This leads to the introduction of noise into the DWDH hologram due to the truncation operations in converting look-up indices required to connect image data to the data written on the hologram.

The reason for such a lack of alignment is twofold. The first is that all printer objectives usually have a finite fifth coefficient, which gives rise to Barrel distortion. The second is that even for the case of a paraxial printer objective, the various coordinate meshes only line up under certain conditions and then only at quantised values of the camera distance parameter, h . For HPO holograms, which must be viewed at the camera distance, all too often, it is therefore just impossible to use such quantised values. The result is that here one must simply accept that the coordinate meshes do not line up.

In Chapter 8, we showed how to correct the printer write data for finite objective distortion. Two methods were discussed in the context of DWDH. The first was to formulate a single “non-paraxial” image transformation capable of converting camera data to actual printer data and incorporating an inverse distortion to correct for the implicit distortion present in the printer. This approach ensures the injection of the minimum amount of interpolation or truncation noise into the hologram. The alternative is to use two separate sequential transformations, the first to convert camera data to paraxial printer data and the second to compensate for the finite distortion of the objective. The first transformation is intrinsically three-dimensional in the case of HPO holograms and four-dimensional in the case of full-parallax holograms, so one is pretty much forced to use nearest-neighbour interpolation. However, the second transformation is two-dimensional and is therefore amenable to bilinear or bicubic interpolation.

In the case of DWDH HPO holograms, by far the best practical way to deal with the non-alignment of coordinate meshes and with interpolation noise is to use oversampling. One therefore maintains complete freedom in defining the camera track and all hologram parameters at the cost of using more camera positions than normally required and, in some cases, in generating more camera data than required for each camera frame. Typical oversampling ratios can be anything from two to ten times. When the 3D image is created by computer modelling software, one usually only needs to oversample the camera positions, as a centred camera frame can be programmed to exactly line up with the hologram hogel mesh. If this cannot be arranged, as will be the case with a real physical camera (we shall study this case in more detail in Chapter 10), then typically bilinear or bicubic interpolation is used to define the proper centred camera frames from the oversampled camera frame data. In either case, a combined **I-to-S** transformation is then used to calculate the final printer data. Because the intrinsic data set of HPO holograms is much smaller than full-parallax holograms, the oversampling technique is extremely effective.

Clearly, oversampling is rarely an option with full-parallax holograms! However, here the fact that the viewing plane is not collocated with the camera plane changes everything. Unlike HPO holograms, it is frequently possible to choose a camera distance that obeys the quantisation rules and as such to achieve perfect alignment of the coordinate meshes under the paraxial objective approximation. Then simple bilinear or bicubic interpolation can be used to calculate the final non-paraxial printer data.

Of course, even with a minimal data set, full-parallax data processing requires a lot of memory. As such, we have seen that today it is simply not practical to use a one-step **I-to-S** transformation as can be applied, even with oversampling, in HPO holography. However, by splitting the problem into two steps, we have seen that in fact the processing required can usually be performed within a very reasonable time on a single PC. Of course, the problem lends itself very simply to parallelisation if increased speed is required.

MWDH inevitably presents a far simpler data-processing problem because the camera data is usually the same as or quite close to the paraxial printer data. For both full-parallax and HPO, the final non-paraxial printer data can be calculated very effectively using oversampling of frame data and bilinear/bicubic interpolation.

REFERENCES

1. S. A. Benton and V. M. Bove Jr., *Holographic Imaging*, Wiley, Hoboken, NJ (2008).
2. M. A. Teitel, *Anamorphic raytracing for synthetic alcove holographic stereograms*, Master's thesis, Massachusetts Institute of Technology, Cambridge, MA (1986).
3. M. Holzbach, *Three-dimensional image processing for synthetic holographic stereograms*, Master's thesis, Massachusetts Institute of Technology, Cambridge, MA (1986).
4. M. W. Halle and A. B. Kropp, "Fast computer graphics rendering for full parallax spatial displays," in *Practical Holography XI and Holographic Materials III*, S. A. Benton and T. J. Trout eds., Proc. of SPIE **3011**, 105–482 (1997).
5. M. E. Holzbach and D. T. Chen, *Rendering methods for full parallax autostereoscopic displays*, US Patent 6,366,370 (filed 1999, granted 2002).

10

Image Data Creation and Acquisition for Digital Display Holograms

10.1 Introduction

The image data required by digital display holograms must be derived either from a physical scene or from a virtual computer model. In this chapter, we shall consider both scenarios. In the case of image data corresponding to a physical scene, we shall see that there are different preferred techniques of data acquisition depending on whether a horizontal parallax-only (HPO) hologram or a full-parallax hologram is to be written. The most common device used is, however, some type of holocam. These devices are based on a mechanically animated high-resolution digital camera. We shall discuss in detail both one-dimensional and two-dimensional holocam systems. Although these systems are, in some respects, intrinsically simple, they alone are capable of recording images of sufficient accuracy for such applications as the holographic archival of cultural heritage or biological specimens.

In the context of computer modelling, we shall discuss typical solutions to setting up and animating a virtual camera using the well-known program 3D StudioMax. Two examples will be discussed in detail. The first is a simple animation procedure that can be graphically programmed to generate the images necessary for HPO holograms. The second discusses the use of the MAXScript scripting language to program a virtual two-dimensional holocam system within Max. We chose Max because we like it; however, there are many three-dimensional (3D) design programs available today. One example of an open-source program of particular interest at the time of writing is Blender. Hopefully, you will be able to apply the general ideas presented in Max to other platforms without much difficulty.

Although nearly all of this chapter is about either real or virtual holocam systems, we will spend a little time talking about structured-light scanning. We would classify this type of data acquisition system as a *secondary* system because, with such systems, one must first generate a 3D computer model of a real-life scene and then one uses a virtual holocam to generate the perspective view data required by the hologram. This type of system is particularly interesting when one wants to record intrinsically dynamic data with full-parallax information or when real and virtual objects must be integrated into the same hologram.

10.2 Image Acquisition from a Physical Scene: HPO Holograms

To write an HPO hologram, we require camera image data taken along a centred horizontal line in front of the object or scene we wish to record. Figure 10.1 shows an illustration of the basic process. A digital camera is mounted on a precision linear translation rail and a stepper or servo motor is used to move it along the rail at a constant velocity. A computer generally controls both the camera motion and also acquires the frame data from the camera.

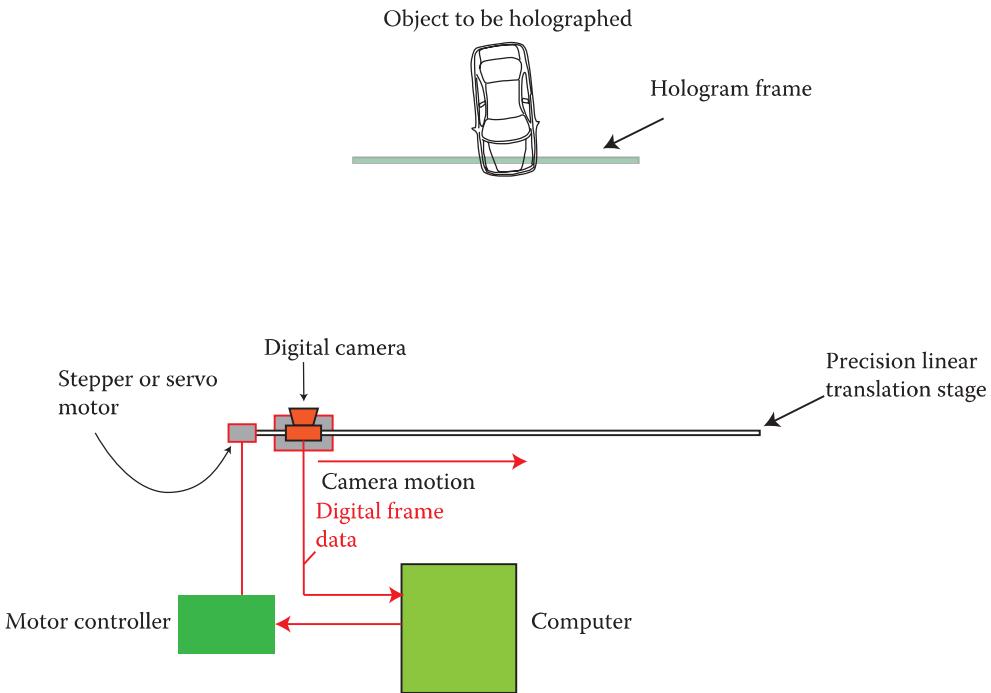


FIGURE 10.1 Basic scenario of image acquisition of a real scene: a digital camera is mounted on a precision linear translation rail and a stepper or servo motor is used to move it along the rail at a constant velocity. A computer generally controls both the camera motion and also acquires the frame data from the camera.

10.2.1 Simple Translating Camera

The simplest way to generate the required camera data is to arrange for a front-facing digital camera to be mounted on an automated linear translation stage. Such a configuration, which we referred to in Chapter 8 as the *simple translating camera*, is illustrated in Figure 10.2. This is the simplest form of a holocam device. A motor drives the camera at a constant velocity from one end of the rail to the other. The simple translating camera, however, suffers from an important defect. The field of view (FOV) of

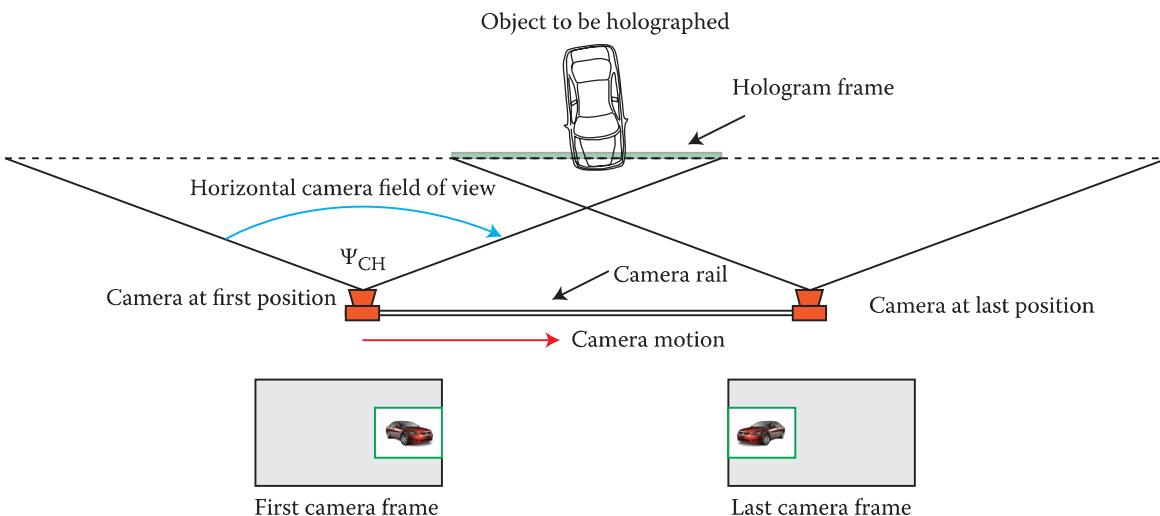


FIGURE 10.2 Simple translating camera configuration. The camera simply moves from one side of the rail to the other at a constant speed taking photographs at a fixed rate. The camera needs to have a large FOV to collect all the data required. The horizontal FOV, Ψ_{CH} , is shown for the first and last pictures. The vertical FOV (not shown) is Ψ_{CV} .

the camera objective must usually be very large to include all the data required by the hologram. As a consequence, only a small part of each camera frame actually falls within the optical field of the hologram and so most of the camera data is not used. A simple example will make this very clear. Suppose we want to record a $1\text{ m} \times 1\text{ m}$ HPO digital hologram with a rectangular window of a horizontal size of 2 m situated at a distance of $h = 1\text{ m}$. If the hologram image to object size is 1:1, we will then need to place our camera rail at a distance of 1 m from the centre of the object (assuming here that the centre of the object coincides with the centre of the hologram). The camera rail will of course need to be 2 m in length. Simple geometry then tells us that the minimum horizontal FOV of the camera must be 112.6° . If we take the aspect ratio of the camera charge-coupled device (CCD) to be 1:0.75, then each frame contains only 14.8% of useful data; the other 85.2% falls outside the hologram optical field! Such a large FOV is of course possible with modern camera lenses, but invariably, there will be significant distortion. Such distortion may usually be compensated for computationally, but the price to pay is a greater requirement on the CCD resolution. Given that one is only using approximately 15% of the image field, one ends up requiring an extremely large CCD. Even worse, if one extends the viewing zone to 3 m wide, the useful CCD area now falls to 1/12th!

This is not to say that the simple translating camera is never useful. It can be used for small holograms of limited FOV very successfully, but as soon as the hologram size or the required holographic FOV becomes large, a simple translating camera provides a rather inefficient means for the recording image data.

10.2.2 Hybrid Translating/Rotating Camera with Fixed Target

A better solution to efficient image acquisition is to use a camera that translates and rotates at the same time [1]. Figure 10.3 illustrates the general idea. By arranging the camera in such a way that it always points to a position in space corresponding to the centre of the physical hologram surface, the FOV required of the camera becomes much smaller. In addition, much more of each camera frame carries useful image data. Taking the example used above of a $1\text{ m} \times 1\text{ m}$ hologram with a 2 m camera track at $h = 1\text{ m}$, the required horizontal camera FOV of 112.6° drops to only 56.14° *.

We shall now look at the geometry of the hybrid camera and see how the data produced by such a configuration can be converted to the data corresponding to the ideal centred camera configuration of Section 8.8.3. Figures 10.4a and b show different sectional views through the relevant geometry, and Figure 10.4c shows a projection onto the camera plane of the hologram frame. As in Chapters 8 and 9, we use the label, ξ , to represent distance along the camera track starting from an origin at the extreme left-hand side. We shall now define two right-handed orthogonal Cartesian coordinate systems. The first is the system (x,y) , which describes the raw camera frame. The second is the system (\hat{x},\hat{y}) , which describes the “centred” camera frame as defined in Section 8.8.3. Figure 10.5 shows the plane (x,y) and the projection of (\hat{x},\hat{y}) onto this plane. We shall also need the 3D orthogonal Cartesian system (X,Y,Z) . These are the world coordinates whose origin is the extreme left-hand side of the camera rail. The coordinate X points along the rail to the right, the coordinate Y points upwards, and the coordinate Z points from the object to the rail.

With reference to Figures 10.4 and 10.5, we now define, in world coordinates, the direction vector of a ray that intersects both the camera origin, O , and a general point on (\hat{x},\hat{y}) :

$$\mathbf{k} = \begin{pmatrix} W/2 - D_X/2 - \xi + \hat{x} \\ \hat{y} - D_Y/2 \\ -h \end{pmatrix} \quad (10.1)$$

* In fact, in this case, the FOV of the camera is more strongly controlled by the vertical geometry—here, we require a vertical FOV of 58.6° , which given an aspect ratio of 1:0.75, corresponds to an effective horizontal FOV requirement of 73.6° .

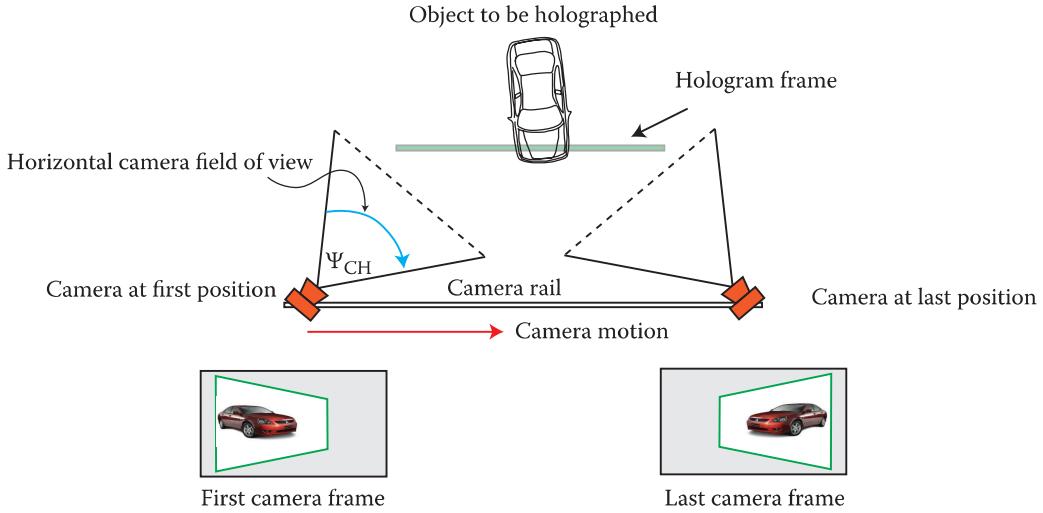


FIGURE 10.3 Hybrid translating/rotating camera configuration with fixed central target. By arranging that the camera always points to a position in space corresponding to the centre of the physical hologram surface, the FOV required of the camera becomes much smaller. This type of system can usually produce higher-quality data sets.

We also define a general point on the raw camera frame:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} h \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} + W/2 \\ y - \sqrt{h^2 + (W/2 - \xi)^2} \tan\left(\frac{\Psi_{CV}}{2}\right) \\ -h + (W/2 - \xi) \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} \end{pmatrix} \quad (10.2)$$

It is then easy to see that we can define the intersection of our test ray with the raw camera plane by extending its direction vector by a factor λ

$$\begin{pmatrix} \xi \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} W/2 - D_x/2 - \xi + \hat{x} \\ \hat{y} - D_y/2 \\ -h \end{pmatrix} = \begin{pmatrix} h \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} + W/2 \\ y - \sqrt{h^2 + (W/2 - \xi)^2} \tan\left(\frac{\Psi_{CV}}{2}\right) \\ -h + (W/2 - \xi) \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} \end{pmatrix} \quad (10.3)$$

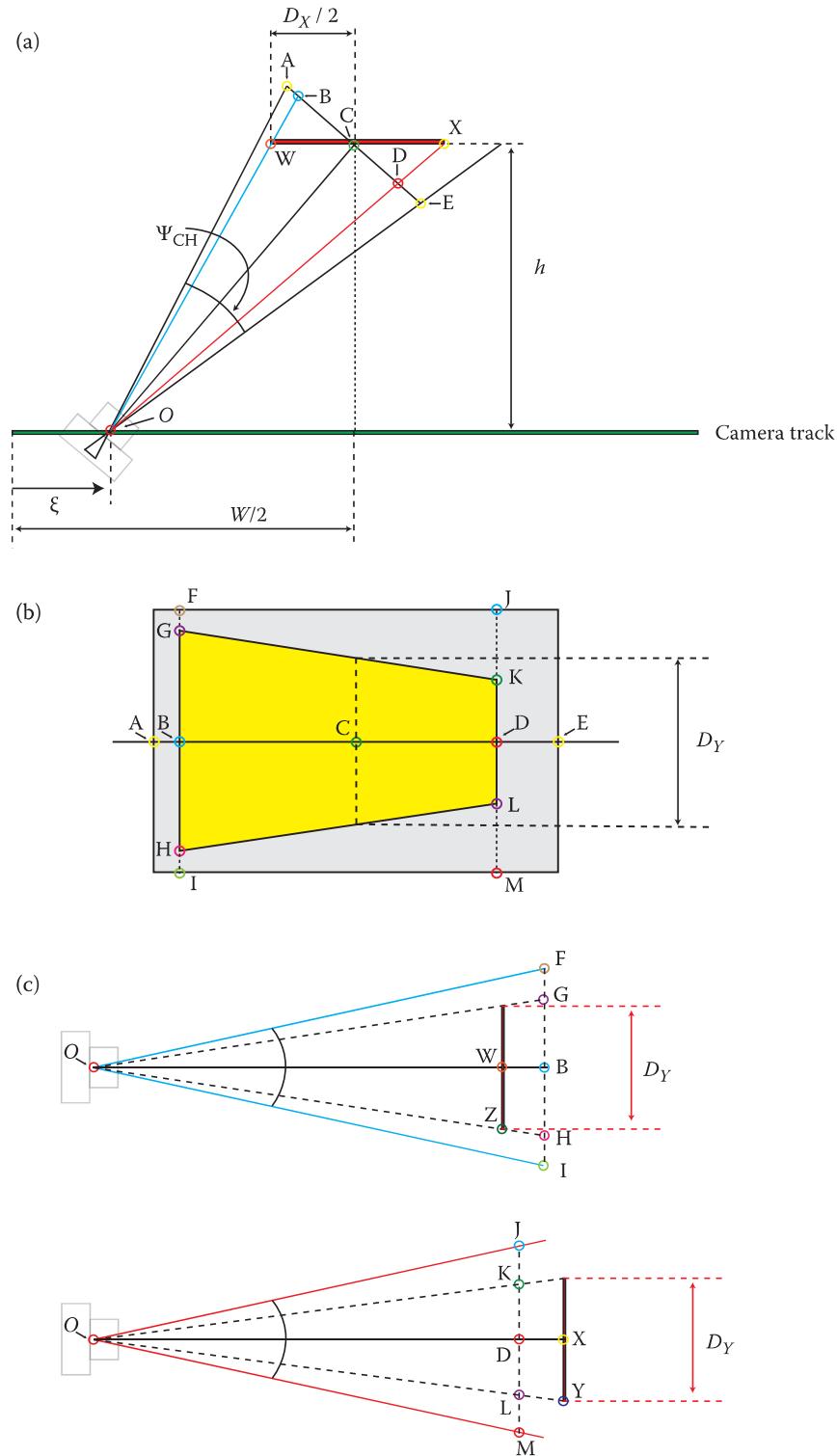


FIGURE 10.4 Geometry of hybrid translating/rotating camera configuration with fixed central target. (a) Plan view showing a midsection through the geometry. Camera track of length W is shown in green, hologram frame in red. The centre of the object is labelled C . Plane labelled ABCDE is the plane of the projected camera frame. This plane is also shown two-dimensionally in (b). The hologram frame and hologram surface are projected onto the projected camera plane in yellow. Two vertical slices are shown in (c) of the geometry from camera point O along the blue and red lines of (a). Letters indicate the correspondence of the points marked on the various diagrams.

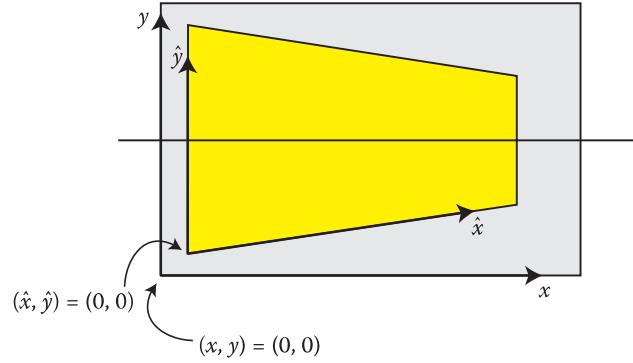


FIGURE 10.5 Projected raw camera frame plane described by the coordinates (x,y) and the projection onto this plane of the “hologram” plane (shown in yellow) described by the coordinates (\hat{x},\hat{y}) . The hologram plane is defined by the area in the real object scene that corresponds exactly, in position and dimension, to the hologram surface in the final holographic reproduction.

This yields the correspondence between the (x,y) and (\hat{x},\hat{y}) systems:

$$\hat{x} = D_x/2 + \xi - W/2 + \frac{h \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} + (W/2 - \xi)}{1 - \frac{(W/2 - \xi)}{h} \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\}} \quad (10.4)$$

$$\hat{y} = \frac{D_y}{2} + \frac{\left(y - \sqrt{h^2 + (W/2 - \xi)^2} \tan\left(\frac{\Psi_{CV}}{2}\right) \right)}{1 - \frac{(W/2 - \xi)}{h} \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\}} \quad (10.5)$$

We can also express such correspondence equally well using the inverse relations

$$x = \sqrt{h^2 + (W/2 - \xi)^2} \left\{ \tan\frac{\Psi_{CH}}{2} + \frac{\hat{x} - \frac{D_x}{2}}{h + \frac{(W/2 - \xi)}{h} \left[\hat{x} - \frac{D_x}{2} + W/2 - \xi \right]} \right\} \quad (10.6)$$

$$y = \left(\hat{y} - \frac{D_y}{2} \right) \left[1 - \frac{(W/2 - \xi)}{h} \left\{ \frac{\hat{x} - \frac{D_x}{2}}{h + \frac{(W/2 - \xi)}{h} \left[\hat{x} - \frac{D_x}{2} + W/2 - \xi \right]} \right\} \right] + \sqrt{h^2 + (W/2 - \xi)^2} \tan\left(\frac{\Psi_{CV}}{2}\right) \quad (10.7)$$

To calculate the centred camera data set, ${}^k\hat{\mathbf{I}}_{ij}$, from the raw camera data set, ${}^k\mathbf{I}_{ij}$, we must now introduce integer coordinates (i,j) and (\hat{i},\hat{j}) corresponding to the real systems (x,y) and (\hat{x},\hat{y}) according to Section 8.8.3:

$$\begin{aligned} x &= 2\sqrt{h^2 + (W/2 - \xi)^2} \tan\left(\frac{\Psi_{\text{CH}}}{2}\right) \left(\frac{i-1}{N_I-1} \right) \quad i = 1, 2, 3, \dots, N_I \\ y &= 2\sqrt{h^2 + (W/2 - \xi)^2} \tan\left(\frac{\Psi_{\text{CV}}}{2}\right) \left(\frac{j-1}{N_J-1} \right) \quad j = 1, 2, 3, \dots, N_J \end{aligned} \quad (10.8)$$

$$\begin{aligned} \hat{x} &= D_X \left(\frac{\hat{i}-1}{\hat{N}_I-1} \right) \quad \hat{i} = 1, 2, 3, \dots, \hat{N}_I \\ \hat{y} &= D_Y \left(\frac{\hat{j}-1}{\hat{N}_J-1} \right) \quad \hat{j} = 1, 2, 3, \dots, \hat{N}_J \end{aligned} \quad (10.9)$$

We must also digitise the camera position

$$\xi = \frac{(k-1)}{(N_K-1)} D_\xi \quad k = 1, 2, 3, \dots, N_K \quad (10.10)$$

We can then use Equations 10.6 through 10.9 to define the two real functions, $\Omega(\hat{i},\xi)$ and $\Upsilon(\hat{j},\xi)$, which constitute real estimates of the integers i and j :

$$\begin{aligned} \Omega(\hat{i},\xi) &= \frac{1}{2}(N_I+1) \\ &+ \frac{D_X h (2\hat{i} - \hat{N}_I - 1)(N_I - 1) \cot\left(\frac{\Psi_{\text{CH}}}{2}\right)}{4(\hat{N}_I - 1)h^2 + 2(W/2 - \xi) \left[D_X (2\hat{i} - \hat{N}_I - 1) + (\hat{N}_I - 1)(W - 2\xi) \right]} \end{aligned} \quad (10.11)$$

$$\begin{aligned} \Upsilon(\hat{i},\hat{j},\xi) &= \frac{1}{2}(N_J+1) \\ &+ \frac{D_Y (\hat{N}_I - 1)(N_J - 1)(2\hat{j} - \hat{N}_J - 1) \sqrt{h^2 + (W/2 - \xi)^2} \cot\left(\frac{\Psi_{\text{CV}}}{2}\right)}{(\hat{N}_J - 1) \left\{ 4h^2 (\hat{N}_I - 1) + (W - 2\xi) \left[D_X (2\hat{i} - \hat{N}_I - 1) + (\hat{N}_I - 1)(W - 2\xi) \right] \right\}} \end{aligned} \quad (10.12)$$

These functions may then be used to calculate the centred camera data set through the transformation

$${}^k\hat{\mathbf{I}}_{ij} = {}^k\mathbf{I}_{ij} \quad \forall \hat{i} \left\{ \hat{i} \in \mathbb{N} \mid \hat{i} \leq \hat{N}_I \right\} \quad \forall \hat{j} \left\{ \hat{j} \in \mathbb{N} \mid \hat{j} \leq \hat{N}_J \right\} \quad \forall k \left\{ k \in \mathbb{N} \mid k \leq N_K \right\} \quad (10.13)$$

where, if we use nearest integer interpolation,

$$\begin{aligned} i &\equiv \|\Omega(\hat{i}, \xi)\| \\ j &\equiv \|\Upsilon(\hat{j}, \xi)\| \end{aligned} \quad (10.14)$$

Alternatively, we may use a bilinear interpolation,* in which case, Equations 10.13 and 10.14 must be replaced by

$$\begin{aligned} {}^k\mathbf{I}_{\hat{i}\hat{j}} &= \frac{{}^k\mathbf{I}_{[\Omega]\lfloor\gamma\rfloor}}{(\lceil\Omega\rceil-\lfloor\Omega\rfloor)(\lceil\gamma\rceil-\lfloor\gamma\rfloor)} (\lceil\Omega\rceil-\Omega)(\lceil\gamma\rceil-\gamma) \\ &+ \frac{{}^k\mathbf{I}_{[\Omega]\lfloor\gamma\rfloor}}{(\lceil\Omega\rceil-\lfloor\Omega\rfloor)(\lceil\gamma\rceil-\lfloor\gamma\rfloor)} (\Omega-\lfloor\Omega\rfloor)(\lceil\gamma\rceil-\gamma) \\ &+ \frac{{}^k\mathbf{I}_{\lfloor\Omega\rfloor\lceil\gamma\rceil}}{(\lceil\Omega\rceil-\lfloor\Omega\rfloor)(\lceil\gamma\rceil-\lfloor\gamma\rfloor)} (\lceil\Omega\rceil-\Omega)(\gamma-\lfloor\gamma\rfloor) \\ &+ \frac{{}^k\mathbf{I}_{\lfloor\Omega\rfloor\lceil\gamma\rceil}}{(\lceil\Omega\rceil-\lfloor\Omega\rfloor)(\lceil\gamma\rceil-\lfloor\gamma\rfloor)} (\Omega-\lfloor\Omega\rfloor)(\gamma-\lfloor\gamma\rfloor) \end{aligned} \quad (10.15)$$

The above transformations let us calculate a centred camera data set from raw camera data when we are explicitly given the camera parameters—that is, when we know the horizontal and vertical FOVs of the camera (Ψ_{CH}, Ψ_{CV}) and the numerical dimensions of the camera CCD ($N_i \times N_j$). But, how does one calculate the most desirable horizontal and vertical FOVs?

To answer this question, we consider the behaviour of the lower left-hand corner of the projection of the centred data set onto the raw camera frame for the left-hand half of the camera trajectory. The minimum horizontal FOV that our camera may possess without clipping any required data is then given by the constraint

$$x = 0 \quad \text{if} \quad (\hat{x}, \hat{y}) = (0, 0) \quad \forall \xi \in \left\{ \Re \mid 0 \leq \xi \leq W/2 \right\} \quad (10.16)$$

This leads to a constraint for each camera position

$$\Psi_{CH} \geq 2 \tan^{-1} \left\{ \frac{D_x}{2h + \frac{2}{h}(W/2 - \xi)(W/2 - \xi - D_x/2)} \right\} \quad (10.17)$$

The value of $(W/2 - \xi)$ corresponding to the most stringent limit is given by differentiating this expression with respect to $(W/2 - \xi)$. This gives $(W/2 - \xi) = D_x/4$, which then defines the general limit

$$\Psi_{CH} \geq 2 \tan^{-1} \left\{ \frac{8hD_x}{16h^2 - D_x^2} \right\} \quad (10.18)$$

* See Appendix 7 for a more detailed description of bilinear and bicubic interpolation methods.

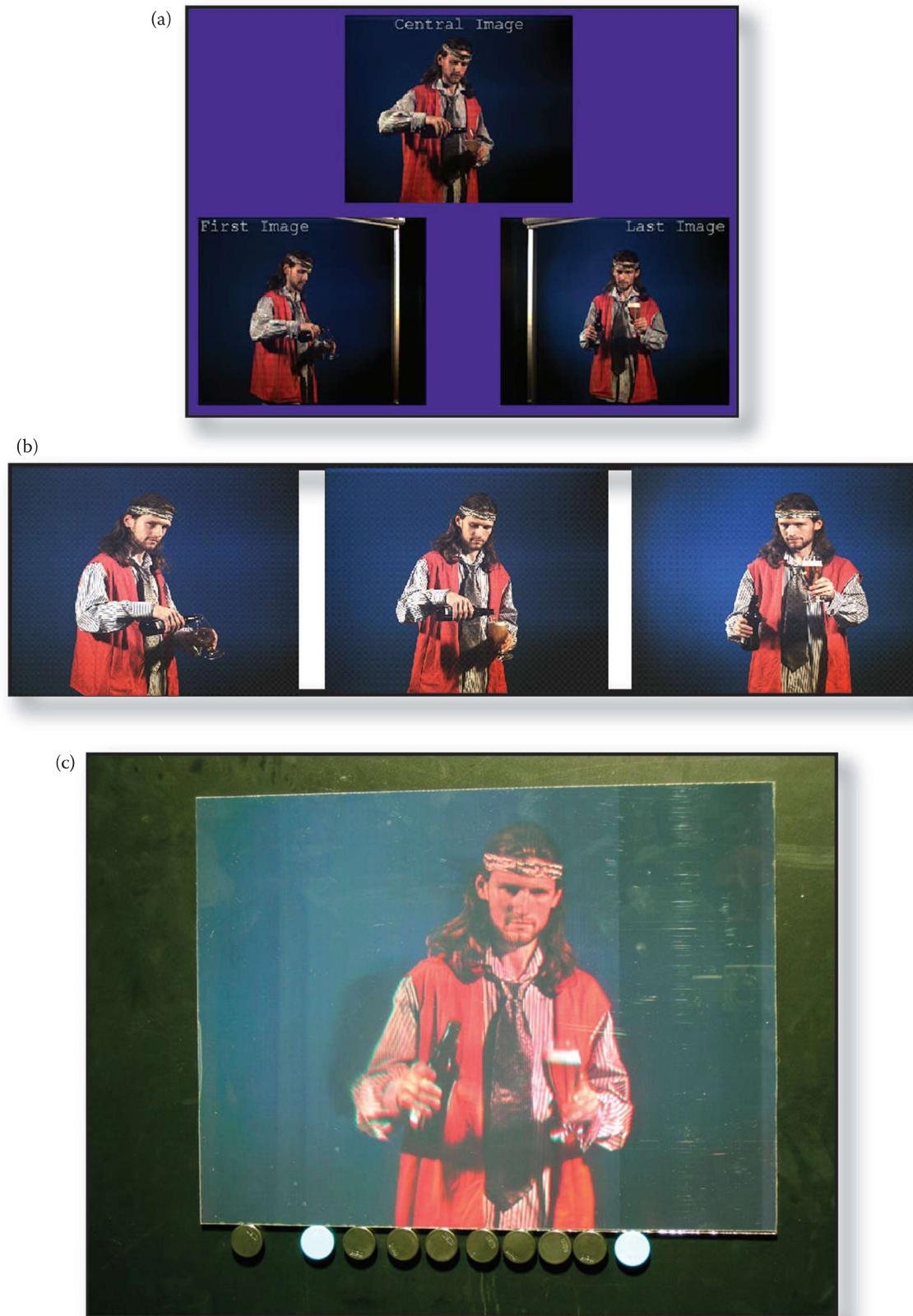


FIGURE 10.6 Example of the data processing required by a hybrid translator/rotator system. (a) Some hundreds to some thousands of raw camera frames are processed using Equation 10.15 to produce a centred data set (b), which is then used to print the final DWDH hologram (c).

The vertical FOV is similarly defined by the condition

$$y = 0 \quad \text{if} \quad (\hat{x}, \hat{y}) = (0, 0) \quad \forall \xi \in \left\{ \Re \mid 0 \leq \xi \leq W/2 \right\} \quad (10.19)$$

which leads to the general limit

$$\Psi_{cv} \geq 2 \tan^{-1} \left\{ \frac{D_Y \sqrt{h^2 + \beta^2}}{2(h^2 + \beta^2 - \beta D_X / 2)} \right\} \quad (10.20)$$

where

$$\beta = \frac{2(6^{1/3} h^2) - \left\{ \sqrt{81 D_X^2 h^4 + 48 h^6} - 9 D_X h^2 \right\}^{2/3}}{6^{2/3} \left\{ \sqrt{81 D_X^2 h^4 + 48 h^6} - 9 D_X h^2 \right\}^{1/3}} \quad (10.21)$$

Equations 10.18 and 10.20 then define the most desirable FOV parameters for a centre-targeted fixed-zoom camera. If one uses these values, then the raw camera data will contain the maximum amount of useful image information. Of course, any camera is characterised by a certain aspect ratio determined by its CCD sensor and as such only one of these constraints will, in practice, be attainable with a spherical optical zoom lens.

10.2.3 Hybrid Translating/Rotating Camera with Optimised Target

The hybrid translating/rotating camera can be somewhat improved by optimising the target point. In the previous section, we considered the case in which the camera targeted the centre of the hologram. We will now consider moving the target position as a function of camera position. This is shown in Figure 10.7, where we have introduced the parameter $\rho(\xi)$ to describe a changing target. As before, the direction

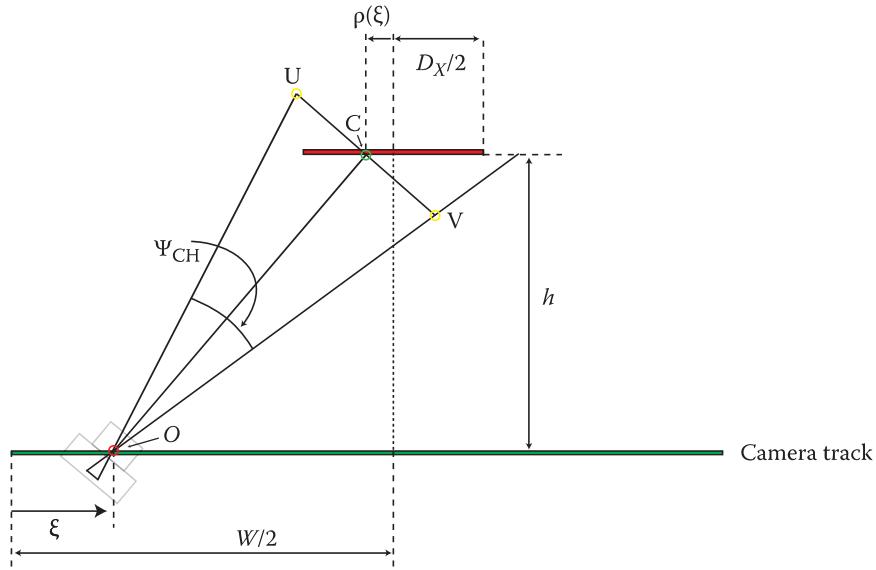


FIGURE 10.7 Geometry of hybrid translating/rotating camera when target position is now moved as a function of camera position. The geometry shown corresponds to a plan view of a meridional section. The target position is characterised by the function $\rho(\xi)$.

of a test ray, expressed in world coordinates that intersects both the camera origin, O , and a general point on (\hat{x}, \hat{y}) is given by

$$\mathbf{k} = \begin{pmatrix} W/2 - D_x/2 - \xi + \hat{x} \\ \hat{y} - D_y/2 \\ -h \end{pmatrix} \quad (10.22)$$

However, now a general point on the (x, y) plane is

$$\begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} = \begin{pmatrix} h \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi - \rho(\xi))^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} + W/2 - \rho(\xi) \\ y - \sqrt{h^2 + (W/2 - \xi - \rho(\xi))^2} \tan\left(\frac{\Psi_{CH}}{2}\right) \\ -h + (W/2 - \xi - \rho(\xi)) \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi - \rho(\xi))^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} \end{pmatrix} \quad (10.23)$$

Following our previous methodology we extend the test ray by a factor λ to arrive at the following three equations relating the (x, y) and (\hat{x}, \hat{y}) systems

$$\begin{pmatrix} \xi \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} W/2 - D_x/2 - \xi + \hat{x} \\ \hat{y} - D_y/2 \\ -h \end{pmatrix} = \begin{pmatrix} h \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi - \rho(\xi))^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} + W/2 - \rho(\xi) \\ y - \sqrt{h^2 + (W/2 - \xi - \rho(\xi))^2} \tan\left(\frac{\Psi_{CV}}{2}\right) \\ -h + (W/2 - \xi - \rho(\xi)) \left\{ \frac{x}{\sqrt{h^2 + (W/2 - \xi - \rho(\xi))^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} \end{pmatrix} \quad (10.24)$$

These equations then give the following rules:

$$\hat{x} = D_x/2 - \beta - \frac{h^2 \left\{ \frac{x}{\sqrt{h^2 + (\beta - \rho)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} + h(\beta - \rho)}{(\beta - \rho) \left\{ \frac{x}{\sqrt{h^2 + (\beta - \rho)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} - h} \quad (10.25)$$

$$\hat{y} = \frac{D_y}{2} + \frac{y - \sqrt{h^2 + (\beta - \rho)^2} \tan\left(\frac{\Psi_{CV}}{2}\right)}{1 - \frac{(\beta - \rho)}{h} \left\{ \frac{x}{\sqrt{h^2 + (\beta - \rho)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\}} \quad (10.26)$$

where

$$\beta = W/2 - \xi \quad (10.27)$$

Alternatively, we can write these relations in an inverse fashion:

$$x = \sqrt{h^2 + (\beta - \rho)^2} \left\{ \frac{h(\hat{x} - D_x/2 + \rho)}{(h^2 + (\hat{x} - D_x/2 + \beta)(\beta - \rho))} + \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} \quad (10.28)$$

$$y = \left(\hat{y} - \frac{D_y}{2} \right) \left[1 - \frac{(\beta - \rho)}{h} \left\{ \frac{x}{\sqrt{h^2 + (\beta - \rho)^2}} - \tan\left(\frac{\Psi_{CH}}{2}\right) \right\} \right] + \sqrt{h^2 + (\beta - \rho)^2} \tan\left(\frac{\Psi_{CV}}{2}\right) \quad (10.29)$$

Equations 10.28 and 10.29 may then be used to define the centred camera data set in terms of the raw camera data using exactly the same method as described in the previous section. However, this assumes that we know the camera pointing function $\rho(\xi)$.

To calculate $\rho(\xi)$, we follow a similar optimisation procedure as introduced for the case of the centrally targeted camera. In the previous section, we required (in the horizontal direction)

$$x = 0 \quad \text{if} \quad (\hat{x}, \hat{y}) = (0, 0) \quad \forall \xi \in \left\{ \Re \mid 0 \leq \xi \leq W/2 \right\} \quad (10.30)$$

This gives an expression similar to before

$$\tan\left(\frac{\Psi_{CH}}{2}\right) \geq \frac{h(\rho - D_x/2)}{(D_x/2 - \beta)(\beta - \rho) - h^2} \quad (10.31)$$

But now we have the supplementary condition

$$\hat{x} = D_x \quad \text{if} \quad (x, y) = \left(2\sqrt{h^2 + (\beta - \rho)^2} \tan\left(\frac{\Psi_{CH}}{2}\right), 0 \right) \quad \forall \beta \in \left\{ \Re \mid -W/2 \leq \beta \leq 0 \right\} \quad (10.32)$$

This condition, which makes sure that the right-hand side of the raw and centred camera frames line up, yields a further equation

$$\tan\left(\frac{\Psi_{CH}}{2}\right) = \frac{h(D_x + 2\rho)}{(D_x + 2\beta)(\beta - \rho) + 2h^2} \quad (10.33)$$

Together, Equations 10.31 and 10.33 define ρ and Ψ_{CH} as functions of β

$$\begin{aligned} \rho &= \frac{4h^2 + D_x^2 + 4\beta^2 - \sqrt{16h^2D_x^2 + (-4h^2 + D_x^2 - 4\beta^2)^2}}{8\beta}, \\ \tan\left(\frac{\Psi_{CH}}{2}\right) &= \frac{-4h^2 + D_x^2 - 4\beta^2 + \sqrt{16h^2D_x^2 + (-4h^2 + D_x^2 - 4\beta^2)^2}}{4hD_x} \end{aligned} \quad (10.34)$$

By differentiating the second expression with respect to β , we can find which value of β corresponds to the greatest required value of $\tan\left(\frac{\Psi_{CH}}{2}\right)$. This is at $\beta = 0$ where $\rho = 0$ and $\tan(\Psi_{CH}/2) = D_x/2h$. Therefore, by moving the target point of the camera according to the rule

$$\rho = \frac{4h^2 + D_x^2 + 4\beta^2 - \sqrt{16h^2D_x^2 + (-4h^2 + D_x^2 - 4\beta^2)^2}}{8\beta} \quad (10.35)$$

we are able to use a horizontal FOV given by

$$\Psi_{CH} = 2 \tan^{-1}\{D_x/2h\} \quad (10.36)$$

This is better than when the camera is targeted to the centre of the hologram as in this case ($\rho = 0$) we must use an FOV of

$$\Psi_{CH} \geq 2 \tan^{-1} \left\{ \frac{8hD_x}{16h^2 - D_x^2} \right\} \quad (10.37)$$

The horizontal linear resolution of the acquired data can therefore be improved by a factor, f , by following the modified rule for the target point where

$$f = \frac{\frac{8hD_x}{16h^2 - D_x^2}}{\frac{D_x}{2h}} = \frac{16h^2}{16h^2 - D_x^2} \quad (10.38)$$

For a 1 m-wide hologram with a viewing distance of 1 m, this amounts to a potential increase in resolution of approximately 7%. The effect is most useful for wide holograms. For example, a 2 m-wide hologram with a viewing distance of 1 m can benefit from a 33% increase in resolution.

To determine the most desirable vertical FOV, we use the constraint we mentioned previously, namely,

$$y = 0 \quad \text{if} \quad (\hat{x}, \hat{y}) = (0, 0) \quad \forall \xi \in \{\Re | 0 \leq \xi \leq W/2\} \quad (10.39)$$

This leads to the relation

$$\tan \left(\frac{\Psi_{CV}}{2} \right) = \frac{D_Y \sqrt{h^2 + (\beta - \rho)^2}}{2 \{ h^2 + (\beta - \rho)(\beta - D_x/2) \}} \quad (10.40)$$

This must be maximised numerically as no analytic solution exists; that is, one must solve

$$\frac{\partial}{\partial \beta} \left\{ \frac{D_Y \sqrt{h^2 + (\beta - \rho(\beta))^2}}{2 \{ h^2 + (\beta - \rho(\beta))(\beta - D_x/2) \}} \right\} = 0 \quad (10.41)$$

over the range $-W/2 \leq \beta \leq 0$.

Figure 10.8 shows a comparison of what the raw camera frames look like using the three methods we have discussed—the simply translating camera (a), the hybrid rotating/translating camera with fixed central target (b) and finally the rotating/translating case with an optimised target (c). Clearly, the most inefficient of the three variants is the simple translating camera, followed by the fixed-target hybrid, with the most efficient variant being the optimised target hybrid camera. Figure 10.9 shows the camera target position of Figure 10.8c versus the camera position. Also shown is the theoretical minimum vertical FOV required versus camera position (derived from Equation 10.40). The value used in Figure 10.8c corresponds to the greatest value of this FOV, which occurs just below $\xi = 1$.*

* You will note that although we have plotted in Figure 10.9 $\Psi(\xi)$ for $0 \leq \xi \leq 3$ m, Equation 10.40 only gives meaningful values for $0 \leq \xi \leq 1.5$ m here.

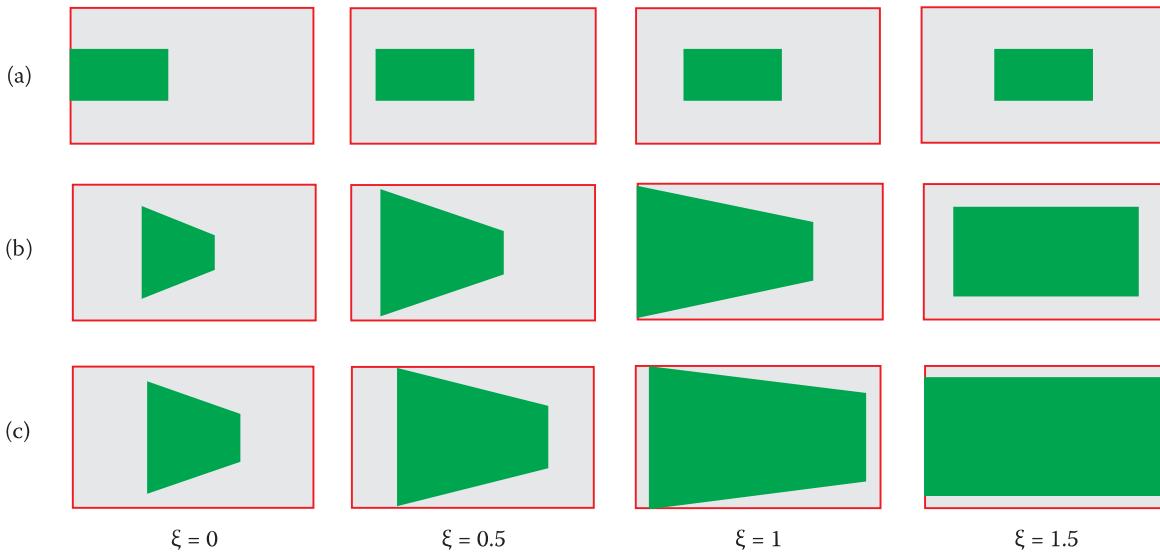


FIGURE 10.8 Comparison of raw camera frames using the three methods discussed for the case of a hologram 1 m high by 2 m wide using a camera distance of 1 m and a rail length of 3 m—(a) the simply translating camera, (b) the hybrid translating/rotating camera with fixed central target, and finally, (c) the translating/rotating case with an optimised target. All cases use optimally calculated horizontal camera FOVs. Case (a) uses a vertical FOV calculated to maintain the CCD aspect ratio equal to that of cases (b) and (c).

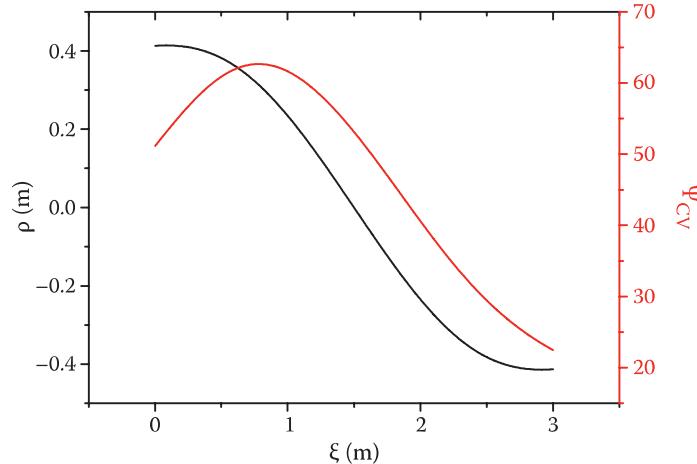


FIGURE 10.9 Optimal camera target position, ρ , of Equation 10.35 used in Figure 10.8c versus camera position, ξ . Also shown is the theoretical minimum vertical FOV required versus camera position (derived from Equation 10.40). The value used in Figure 10.8c corresponds to the greatest value of this FOV, which occurs just below $\xi = 1$.

10.2.4 Resolution Requirements

In Chapter 9, we saw that to meet the resolution requirements of the human eye, the distance between camera positions along the camera line had to be less than 1 mm per unit hologram-camera distance. In other words, if one wants to view a hologram at a distance of 0.5 m, then the camera line will need to be located at $h = 0.5$ m and one will need 2000 camera positions per metre length of camera track. We also saw in Chapter 9 that oversampling of camera positions was desirable in HPO holography due to the intrinsic misalignment of the various coordinate meshes associated with the printer image data and

the raw camera data. When large maximum-resolution HPO holograms are to be written with horizontal windows of several metres in length, it is not uncommon to require more than 10,000 camera frames.

Attention must also be given to the resolution of the CCD used in the camera. We have seen in the previous sections how a hybrid rotating/translating camera with an optimised tracking target can allow one to use the CCD much more efficiently than in the case of a simply translating camera—and this generally translates to a much better hologram for a given camera CCD. Whatever the configuration used, however, one must try to ensure that the camera resolution is several times better than the minimum angular spacing between hogels. In practice, this can be more difficult to achieve at large angles of view as will be clear from the $\xi = 0$ diagrams of Figure 10.8a, b or c.

10.2.5 Image Reduction and Magnification

The holographic image may be reduced or magnified by proportionately reducing or expanding all physical sizes at recording with respect to the corresponding values at hologram replay. Let us denote the distance between the camera track centre and the hologram centre as h_1 , and the corresponding viewing distance on replay of the hologram by h . Similarly, let us denote the physical camera rail length as W_1 and the width of the viewing box on replay by W . Finally, we denote the hologram frame size on recording by $D_{X1} \times D_{Y1}$ and by $D_X \times D_Y$ at replay. Then we can define the magnification of the holographic image as

$$M = h/h_1 = W/W_1 = D_X/D_{X1} = D_Y/D_{Y1} \quad (10.42)$$

If $M = 2$, then the holographic image will appear twice as big as the actual physical object. Of course, to make large reductions—for example, reducing a building to a 30×40 cm hologram—takes a very large rail system. Large magnifications can be achieved using microstages and this is a definite area of future interest in the field of biological specimen archival.

10.2.6 Commercial Holocam Systems

At the time of writing, several companies offered commercial holocam systems for sale. Most notable are the UK company Spatial Imaging Ltd. and the Lithuanian company Geola Digital UAB. Both companies use the efficient hybrid rotating/translating geometry. Figure 10.10 shows a standard studio system manufactured by Geola, which was built around a 4 m electromechanical rail system. The rail was built by the German company Isel. Early models of this device used high-resolution, high-speed CMOS camera systems requiring bright studio lights. However, in the past few years, camera technology has improved radically, with the result that all of Geola's current models now only require normal room lighting.

Figure 10.11 shows a portable holocam system that has proven to be the most popular type of system. The system is based on a short electromechanical rail just over 1 m in length, mounted on two photographic tripods. The holocam is supplied with its own laptop running highly automated proprietary software that takes care of alignment, shooting and data processing. A typical camera now being shipped with the device (Figure 10.12) is the Xacti HD electronic camera with a pixel resolution of 1920×1080 , capable of operation at 60 frames per second. The system can be transported easily in the back of a car and assembled in a few minutes. Operation is likewise extremely easy and the resulting files can be sent via the Internet straight to Geola for printing. The electromechanical rail is made by Isel in Germany and camera rotation is assured by a precision electromechanical rotation stage produced by the Lithuanian company Standa UAB.

One of the intrinsic problems with holocam systems is their cost. For top-quality results, the animation of the camera must be very precise, and as a result, the electromechanics required is often rather costly. A good camera can also be a costly element. To tackle this problem, Geola has just released a cheap holocam system (Figure 10.13), which currently markets for less than £1000. This is a small device that uses cheap components to ensure the required hybrid camera animation. Instead of the ultraprecise rail from Isel, the system uses a rather simpler rail in which the camera is mounted on a platform which has rubber wheels and which rides a type of “railway” track. The camera rotation is assured by a mechanical



FIGURE 10.10 Professional holocam system manufactured and available commercially from the Lithuanian company Geola Digital UAB. Designed for installation in photographic studios, this fully computer-controlled holocam, which is a hybrid translator/rotator system, comprises a 4 m-long precision electromechanical rail with high-resolution fast digital camera. It is suitable for the production of high-resolution image data sets required by high-quality large-format digital colour reflection holograms.



FIGURE 10.11 Portable commercial holocam system using hybrid translator/rotator configuration. This self-contained system is capable of easy transport and quick setup—ideal for capturing data sets when out of the studio. A Wi-Fi Internet connection allows the user to beam the data back for printing. Although not as large as the studio models offered by Geola, the company has sold many more of these smaller devices.



FIGURE 10.12 The Xacti HD electronic camera with a pixel resolution of 1920×1080 capable of operation at 60 frames per second was recommended by Geola Digital UAB for use in its portable holocam systems at the time of writing.



FIGURE 10.13 A new concept from Geola is the affordable holocam. Made from cheaper components, the unit allows smaller good-quality holographic portraits to be taken effectively. The system is a hybrid translator/rotator in which the rotation is assured mechanically by a metal arm instead of an expensive precision rotation stage. The unit, which comes with its own netbook, is completely self-contained and yet markets for less than £1000. The system can be supplied with a variety of digital cameras. (a) Side view showing rotator arm, (b) supplied netbook with installed proprietary software, (c) camera mount and (d) bottom view showing motor and wheel-based system for camera translation.

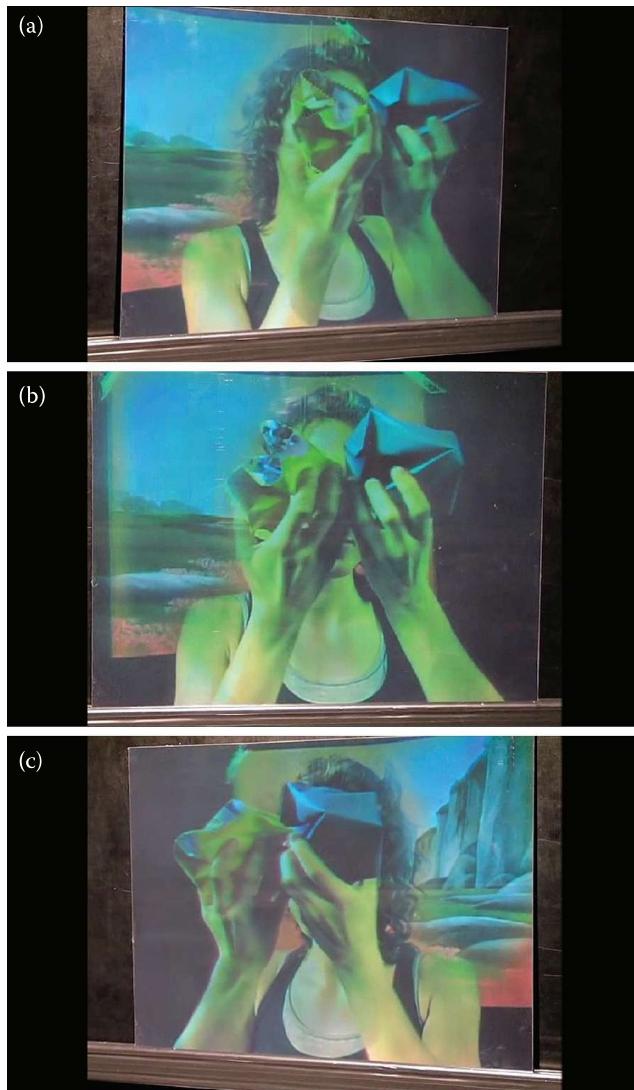


FIGURE 10.14 Example of an HPO DWDH full-colour reflection hologram ($40\text{ cm} \times 60\text{ cm}$) created using a portable Geola holocam system. (a) View from the left, (b) view from the centre and (c) view from the right. (Artist: Martin Richardson 2010.)

arm that physically turns the camera as it rotates. This cheap holocam is principally useful for portraiture at a fixed format. As such, it is a little restrictive, but the results are not so bad.

10.2.7 Alternative Strategies—Rotating the Object

An alternative strategy to the use of rail-mounted camera systems is the use of a static camera and a rotating platform. This geometry can most successfully be used with small objects or human subjects destined for the creation of small- to medium-sized display holograms. The system has a number of advantages, the most obvious of which is the potential simplicity offered by a simple rotating platform—if the mass of the platform is sufficiently high, then manual rotation will often produce acceptable results. Motorisation is, however, preferable, but again, it is much simpler to achieve than any of the linear-rail options. The main disadvantage is that people look at a hologram from a constant perpendicular distance; when the hologram is mounted on a wall, for example, people tend to walk past it, describing a linear trajectory parallel to the hologram. For small angles of view, there is not a great difference between a linear viewing line and the curved viewing line characteristic of the rotating platform geometry. However,

at larger angles (Figure 10.16), the difference becomes important and the normal observer finds that the HPO hologram distorts when viewed obliquely.

The image processing for the static camera and rotating platform configuration is a little different from the case of a linear rail as the viewing line becomes the segment of a circle. There are basically two parts to the image processing problem. The first is the transformation of the raw camera data into a centred camera data set. This is very similar to the problem we encountered in the context of a hybrid camera on a linear rail. In Section 10.3.1, we will solve this problem for the case of a horizontal rotating platform combined with a simple vertical tracking of the camera; this more general geometry is useful for full-parallax image acquisition. The present problem represents a specific case of this configuration, and as such, we shall simply refer the reader to Equations 10.71 through 10.75 where, for the problem at hand, we must put $D_\zeta = 0$.

The second part of the problem entails, for a given hogel, a derivation of the relationship between the horizontal spatial light modulator (SLM) coordinate and the camera shot coordinate. We will need this relationship to write down a modified I-to-S transformation appropriate for the rotating platform configuration. This is somewhat different from what we discussed in Chapter 8 in so much as the camera line and the viewing line are now segments of a circle rather than simple lines.

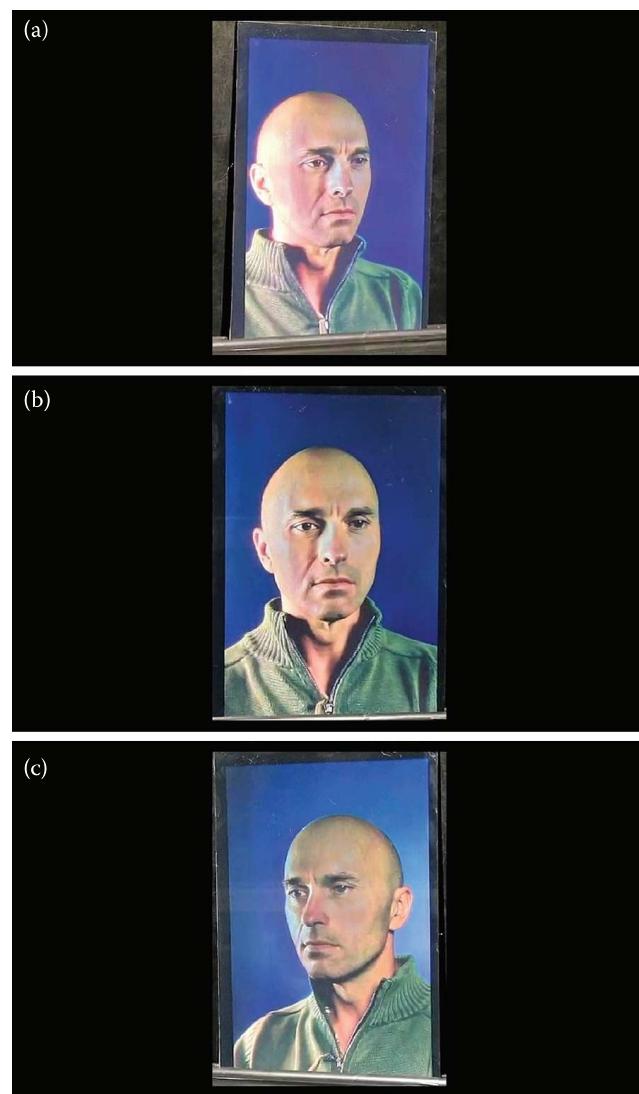


FIGURE 10.15 Example of an HPO DWDH full-colour reflection portrait (30 cm × 40 cm) created using a portable Geola holocam system. (a) View from the left, (b) view from the centre and (c) view from the right. (Copyright Julio Ruis, ITMA, Spain 2010.)

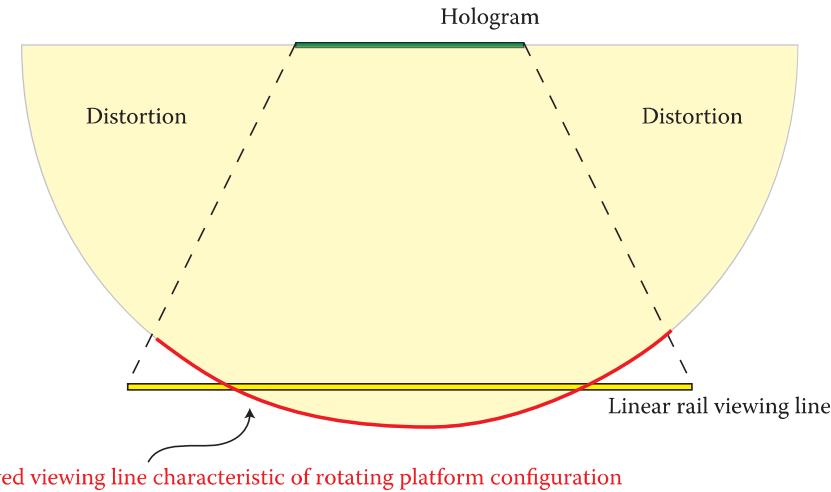


FIGURE 10.16 Comparison of the geometry of a linear camera track and that of the curved track characteristic of the technique of a static camera and rotating platform. At oblique angles, the curved track differs substantially from the linear track, leading to unacceptable distortion in the HPO hologram.

Figure 10.17 shows a plan view of the geometry. We start by digitising the effective camera position* in terms of the rotation angle of the platform

$$\theta_\xi = \frac{(k-1)}{(N_K-1)} \Phi_c \quad (10.43)$$

This means that we take pictures at even angle increments; if the rotation platform rotates with a uniform angular velocity, then this is equivalent to the camera taking pictures at a constant rate. Next, we digitise the horizontal coordinate of a paraxial projection of the SLM onto a plane at an arbitrary perpendicular distance h^* :

$$U' = 2h^* \tan\left(\frac{\Psi_{PH}}{2}\right) \frac{(\mu'-1)}{(N_M-1)} \quad (10.44)$$

As before, we digitise the hologram and the film frame using

$$X = \frac{(\alpha-1)}{(N_A-1)} D_x \quad (10.45)$$

and

$$x' = \frac{(i'-1)}{(N_I-1)} D_x \quad (10.46)$$

Because we are using a centred camera configuration, we put

$$\begin{aligned} D_x &= D_X \\ N_I &= N_A \end{aligned} \quad (10.47)$$

We now need one additional coordinate system. These are the world coordinates $(\mathcal{X}, \mathcal{Z})$ which we shall choose to be right-handed, two-dimensional orthonormal Cartesian coordinates with origin located at the centre of the hologram. A test ray emanating from a hogel described by the horizontal coordinate, X , and associated with the horizontal SLM coordinate, U' , will then intersect the camera circle at a

* For ease of analysis, we consider the equivalent case of a static platform and a targeted camera that rotates around the centre of the platform.

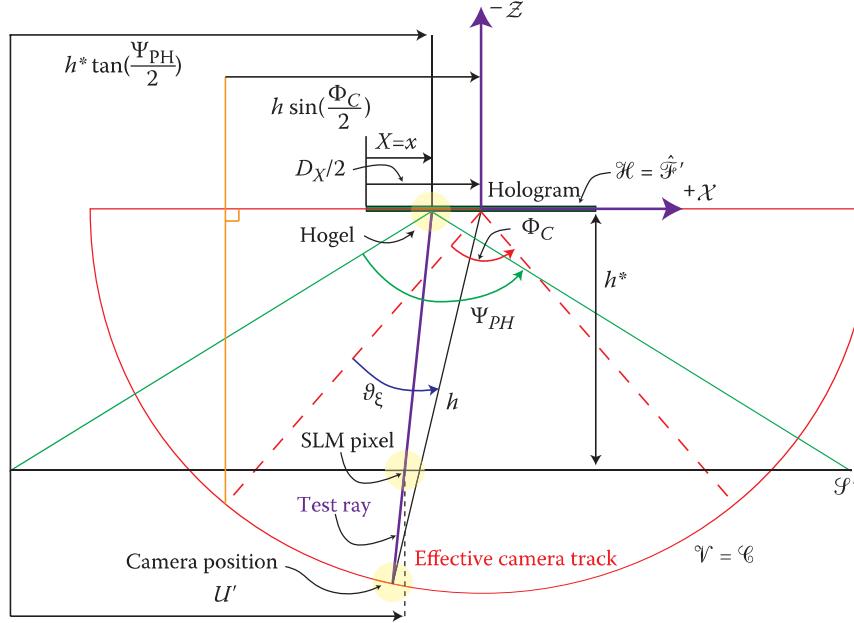


FIGURE 10.17 Plan view of the geometry characterising the static camera and rotating platform configuration. We use the equivalent circular camera track with centre-targeted camera. In particular, we consider the intersection of the purple test ray emanating from a given hogel with both the camera track and with a projection of the SLM plane at a perpendicular distance, h^* , from the hologram. This procedure leads us to the I-to-S transformation characteristic of the configuration.

horizontal world coordinate given by the following vector expression, which we obtain by simply extending the direction vector of the test ray by a distance, λ

$$\begin{pmatrix} \mathcal{X} \\ \sqrt{h^2 - \mathcal{X}^2} \end{pmatrix} = \begin{pmatrix} X - D_x/2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} U' - h^* \tan(\Psi_{PH}/2) \\ h^* \end{pmatrix} \quad (10.48)$$

The z component immediately defines λ

$$\lambda = \frac{\sqrt{h^2 - \mathcal{X}^2}}{h^*} \quad (10.49)$$

whereupon we can use the x component to solve for the ray intersection

$$\mathcal{X} = \frac{X - \frac{D_x}{2} \pm \sqrt{\left(X - \frac{D_x}{2}\right)^2 + \left[1 + \frac{1}{h^{*2}} \left\{U' - h^* \tan\left(\frac{\Psi_{PH}}{2}\right)\right\}^2\right] \left[\frac{h^2}{h^{*2}} \left\{U' - h^* \tan\left(\frac{\Psi_{PH}}{2}\right)\right\}^2 - \left(X - \frac{D_x}{2}\right)^2\right]}}{1 + \frac{1}{h^{*2}} \left\{U' - h^* \tan\left(\frac{\Psi_{PH}}{2}\right)\right\}^2} \quad (10.50)$$

This may now be written in terms of the integer coordinates

$$\mathcal{X} = \frac{\left[\frac{D_x}{2}\right] \left(\frac{2\alpha - 1 - N_A}{(N_A - 1)}\right) \pm \sqrt{\varpi}}{1 + \tan^2(\Psi_{PH}/2) \left\{\frac{2\mu' - 1 - N_M}{(N_M - 1)}\right\}^2} \quad (10.51)$$

where

$$\begin{aligned} \mathfrak{W} = & \left[\frac{D_X}{2} \right]^2 \left(\frac{2\alpha - 1 - N_A}{(N_A - 1)} \right)^2 \\ & + \left[h^2 \tan^2 \left(\frac{\Psi_{PH}}{2} \right) \left\{ \frac{2\mu' - 1 - N_M}{(N_M - 1)} \right\}^2 - \left[\frac{D_X}{2} \right]^2 \left(\frac{2\alpha - 1 - N_A}{N_A - 1} \right)^2 \right] \times \\ & \left[1 + \tan^2 \left(\frac{\Psi_{PH}}{2} \right) \left\{ \frac{2\mu' - 1 - N_M}{(N_M - 1)} \right\}^2 \right] \end{aligned} \quad (10.52)$$

The camera position, k , is then related to \mathcal{X} by the simple expression

$$k = \left\| \frac{(N_K + 1)}{2} + \frac{(N_K - 1)}{\Phi_C} \sin^{-1} \left[\frac{\mathcal{X}}{h} \right] \right\| \quad (10.53)$$

This may be used in place of Equation 8.63 in the calculation of the relevant printer data set, ${}^{\mu\nu}\mathbf{S}_{\alpha\beta}$. The full \mathbf{I} -to- \mathbf{S} transformation for an inverting camera and a conjugate SLM position may be written as*

$$\begin{aligned} {}^{\mu\nu}\mathbf{S}_{\alpha\beta} = & {}^k\hat{\mathbf{I}}_{\hat{i}\hat{j}} \quad \forall \alpha \in \mathbb{N} | \alpha \leq N_A \quad \forall \beta \in \mathbb{N} | \beta \leq N_B \\ & \forall \mu \in \mathbb{N} | N_{U1}(\alpha) \leq \mu \leq N_{U2}(\alpha) \quad \forall v \in \mathbb{N} | N_{V1}(\beta) \leq v \leq N_{V2}(\beta) \\ = & 0 \quad \text{otherwise} \end{aligned} \quad (10.54)$$

where

$$\begin{aligned} \hat{i} &= N_I - \alpha + 1 \\ \hat{j} &= N_J - \beta + 1 \\ k &= \left\| \frac{(N_K + 1)}{2} + \frac{(N_K - 1)}{\Phi_C} \sin^{-1} \left[\frac{\mathcal{X}}{h} \right] \right\| \end{aligned} \quad (10.55)$$

and

$$\begin{aligned} N_{V1}(\beta) = & \left\| \frac{N_V + 1}{2} + \frac{(N_V - 1)}{2h \tan \left(\frac{\Psi_{PV}}{2} \right)} \left\{ D_Y \left(\frac{N_B - \beta}{N_B - 1} \right) - \frac{(D_Y + H)}{2} \right\} \right\| \\ N_{V2}(\beta) = & \left\| \frac{N_V + 1}{2} + \frac{(N_V - 1)}{2h \tan \left(\frac{\Psi_{PV}}{2} \right)} \left\{ D_Y \left\{ \frac{(N_B - \beta)}{N_B - 1} + \frac{H}{D_Y} \right\} - \frac{(D_Y + H)}{2} \right\} \right\| \end{aligned} \quad (10.56)$$

* Note that in Equation 10.56 we have chosen to define the vertical window focus at a distance of h from the hologram surface. Another popular choice is to set this focus to follow the camera track. One does this by simply replacing the two h variables in Equation 10.56 by $\sqrt{h^2 - \mathcal{X}^2}$. Note also that the variable \mathcal{X} in Equation 10.55 must be calculated using Equations 10.51 and 10.52.

The limit expressions for $N_{U1}(\alpha)$ and $N_{U2}(\alpha)$ may be obtained by setting, respectively, $k = 1$ and $k = N_K$ in Equations 10.51 through 10.53:

$$N_{U\gamma} = \frac{N_M + 1}{2} + \frac{1}{2} \left\{ \mathcal{X}_\gamma + \frac{D_X}{2} \frac{(1 + N_A - 2\alpha)}{N_A - 1} \right\} \frac{\cot\left(\frac{\Psi_{PH}}{2}\right)(N_M - 1)}{\sqrt{h^2 - \mathcal{X}_\gamma^2}} \quad (10.57)$$

where

$$\begin{aligned} \mathcal{X}_1 &= -h \sin\left(\frac{\Phi_C}{2}\right) \\ \mathcal{X}_2 &= h \sin\left(\frac{\Phi_C}{2}\right) \end{aligned} \quad (10.58)$$

10.3 Image Acquisition from a Physical Scene: Full-Parallax Digital Holograms

Perhaps the greatest difference between the problems of image data acquisition for HPO holograms and full-parallax holograms is that a scanning camera cannot usually be used in the latter case when the object is able to move. In the HPO case, a small amount of movement is often desired as this adds realism to the final hologram. For example, a human portrait may be made of a subject breaking into a smile as the camera traverses the rail from left to right. Then, on replay, the observer will observe, on traversing the viewing zone of the hologram, an exact replay of this animation. As long as the animation is “slow”, then the horizontal parallax is not affected.* In full-parallax holography, the only way to possibly include animation is to use a plurality of vertically arranged cameras that track simultaneously along a rail. However, given the discussions of Section 10.2.4, it will be immediately obvious that this solution is going to be rather difficult for any but the simplest display holograms. Nevertheless, we should not discount this possibility because, for small depth holograms, the vertical parallax does not require a lot of sampling; as we shall see in the next chapter, chromatic blurring in reflection holograms operates principally in the vertical direction, blurring the vertical parallax. For wideband illumination, a large stacked array of compact cameras can therefore sometimes provide a solution when it is necessary to include moving objects. Clearly, however, in the context of ultra-realistic holographic images of large depth and wide angle of view, not only would one require typically thousands of cameras in a stack, but also each camera would either need to be extremely small or the stack extremely large.

Rail-mounted camera systems for full-parallax image data acquisition are therefore nearly always used with static objects. Applications particularly suited to rail-mounted camera acquisition are museum artefact archival and biological archival. Often, large magnifications are desirable (either up or down). The techniques presented in the last section transfer over directly to the full-parallax case with the camera now being mounted on a two-dimensional translating stage in place of a linear rail. In principle, each technique we discussed in the context of HPO holograms can be combined with another for the second dimension. Nine principal ways therefore exist to control the camera using a two-dimensional stage. In addition, a rotation platform may be used for either dimension leading to a

* Fast animation leads to coupling of time and parallax leading to unexpected effects in the 3D properties of the image.

total of 16 ways in which we might possibly contemplate recording a full-parallax data set. It is worth pointing out that the previous restrictions of the rotating platform technique that operate in the HPO case no longer apply to full-parallax applications. Here, due to Huygens' principle, we see no distortion at oblique angles. Accordingly, a rotating platform becomes an extremely useful tool for many full-parallax applications. We shall study only one case here for reasons of economy of space, but the method we describe can be mechanically applied to any of the 16 possible configurations. The case we will study is that of the horizontally positioned rotating platform combined with a simple vertically translating camera.

10.3.1 Horizontal Rotating Platform with Vertically Linear-Translating Camera

As in Section 10.2.7, a discussion of this geometry is naturally divided into two parts. The first part, sometimes referred to as “keystoning”, is about transforming the raw camera data into a centred camera data set. The second part is about deriving the I-to-S transformation characteristic of the configuration. We shall start with the first problem and develop the mathematics required to define a centred camera data set.

10.3.1.1 Calculation of a Centred Camera Data Set

Figure 10.18 shows vertical and horizontal sections through the relevant geometry. We define 3D world coordinates as usual but with the difference now that their origin is at the centre of the hologram frame. A general point on the plane (\hat{x}, \hat{y}) may therefore be written in world coordinates as

$$\mathbf{Q}(\hat{x}, \hat{y}) = \hat{\mathbf{e}}_x \left(\hat{x} - \frac{D_x}{2} \right) + \hat{\mathbf{e}}_y \left(\hat{y} - \frac{D_y}{2} \right) \quad (10.59)$$

where $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_y$ and $\hat{\mathbf{e}}_z$ are the world coordinate unit vectors. Likewise, a general camera position may be expressed as

$$\mathbf{C} = \hat{\mathbf{e}}_x h \sin(\vartheta_\xi - \Phi_c/2) + \hat{\mathbf{e}}_y (\zeta - D_\zeta/2) + \hat{\mathbf{e}}_z h \cos(\vartheta_\xi - \Phi_c/2) \quad (10.60)$$

Accordingly, the direction vector of a ray passing through both the camera point and an arbitrary point on the (\hat{x}, \hat{y}) plane can be written as

$$\mathbf{k} = \left\{ h \sin\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) - \hat{x} + \frac{D_x}{2} \right\} \hat{\mathbf{e}}_x + \left\{ \zeta - \frac{D_\zeta}{2} - \hat{y} + \frac{D_y}{2} \right\} \hat{\mathbf{e}}_y + \left\{ h \cos\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) \right\} \hat{\mathbf{e}}_z \quad (10.61)$$

The unit vector $\hat{\mathbf{k}}_o$, representing the direction between the centre of the raw camera frame and the vertically central camera point, is then given by

$$\hat{\mathbf{k}}_o = \sin\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) \hat{\mathbf{e}}_x + \cos\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) \hat{\mathbf{e}}_z \quad (10.62)$$

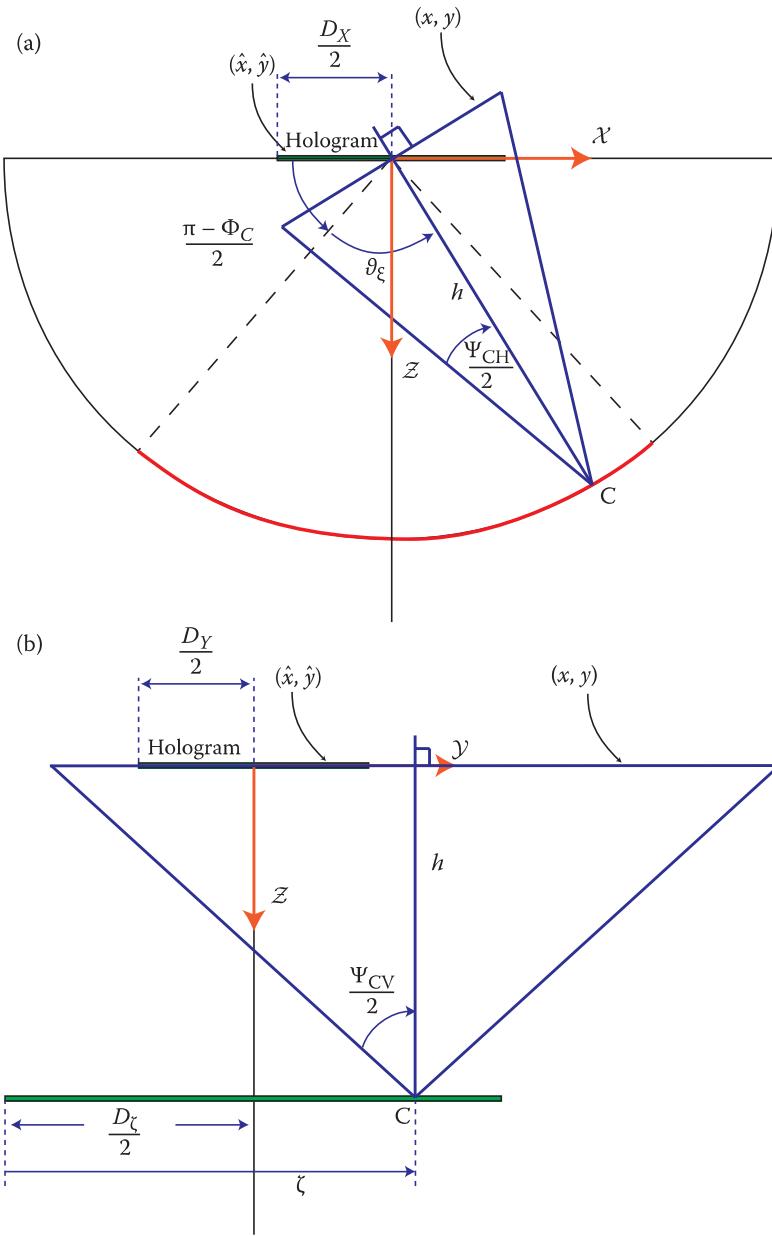


FIGURE 10.18 Full-parallax image data acquisition: geometry for a horizontal rotating platform with vertically forward-facing linear-translating camera. (a) Plan view and (b) side view.

We can use this vector to define the unit vectors of the (x,y) system as

$$\hat{\mathbf{e}}_x = -\frac{\hat{\mathbf{k}}_o \times \hat{\mathbf{e}}_y}{|\hat{\mathbf{k}}_o \times \hat{\mathbf{e}}_y|} = -\frac{1}{|\hat{\mathbf{k}}_o \times \hat{\mathbf{e}}_y|} \begin{vmatrix} \hat{\mathbf{e}}_x & \hat{\mathbf{e}}_y & \hat{\mathbf{e}}_z \\ \sin\left(\vartheta_\xi - \frac{\Phi_C}{2}\right) & 0 & \cos\left(\vartheta_\xi - \frac{\Phi_C}{2}\right) \\ 0 & 1 & 0 \end{vmatrix} \quad (10.63)$$

$$= \hat{\mathbf{e}}_x \cos\left(\vartheta_\xi - \frac{\Phi_C}{2}\right) - \sin\left(\vartheta_\xi - \frac{\Phi_C}{2}\right) \hat{\mathbf{e}}_z$$

and

$$\hat{\mathbf{e}}_y = \frac{\hat{\mathbf{k}}_o \times \hat{\mathbf{e}}_x}{|\hat{\mathbf{k}}_o \times \hat{\mathbf{e}}_x|} = \frac{1}{|\hat{\mathbf{k}}_o \times \hat{\mathbf{e}}_x|} \begin{vmatrix} \hat{\mathbf{e}}_x & \hat{\mathbf{e}}_y & \hat{\mathbf{e}}_z \\ \sin\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) & 0 & \cos\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) \\ 1 & 0 & 0 \end{vmatrix} = \hat{\mathbf{e}}_y \quad (10.64)$$

A general point on the plane (x,y) may then be written as

$$\begin{aligned} \mathbf{P}(x, y) = \Theta + \hat{\mathbf{e}}_x x + \hat{\mathbf{e}}_y y &= \hat{\mathbf{e}}_x \left\{ x \cos\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) - h \tan\left(\frac{\Psi_{CH}}{2}\right) \cos\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) \right\} \\ &\quad + \hat{\mathbf{e}}_y \left\{ \zeta - D_\zeta/2 - h \tan\left(\frac{\Psi_{CV}}{2}\right) + y \right\} \\ &\quad + \hat{\mathbf{e}}_z \left\{ h \tan\left(\frac{\Psi_{CH}}{2}\right) \sin\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) - x \sin\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) \right\} \end{aligned} \quad (10.65)$$

where Θ is a world vector describing the origin of the raw camera frame

$$\begin{aligned} \Theta &= -\hat{\mathbf{e}}_x h \tan\left(\frac{\Psi_{CH}}{2}\right) - \hat{\mathbf{e}}_y h \tan\left(\frac{\Psi_{CH}}{2}\right) + (\zeta - D_\zeta/2) \hat{\mathbf{e}}_y \\ &= -\hat{\mathbf{e}}_x \left\{ h \tan\left(\frac{\Psi_{CH}}{2}\right) \cos\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) \right\} + \hat{\mathbf{e}}_y \left\{ \zeta - D_\zeta/2 - h \tan\left(\frac{\Psi_{CV}}{2}\right) \right\} \\ &\quad + \hat{\mathbf{e}}_z \left\{ h \tan\left(\frac{\Psi_{CH}}{2}\right) \sin\left(\vartheta_\xi - \frac{\Phi_c}{2}\right) \right\} \end{aligned} \quad (10.66)$$

We can now define the relationship between the (x,y) and (\hat{x}, \hat{y}) coordinates by considering the intersection of a test ray passing through the camera point and general points on both coordinate planes:

$$\mathbf{P}(x, y) + \lambda \mathbf{k} = \mathbf{Q}(\hat{x}, \hat{y}) \quad (10.67)$$

These three equations yield λ , the ray extension parameter, and the required coordinate transformation equations:

$$x = h \sec\left(\frac{\Psi_{CH}}{2}\right) \frac{\left(2h \sin\left(\frac{\Psi_{CH}}{2}\right) - (D_x - 2\hat{x}) \cos\left\{\frac{1}{2}(\Phi_c - \Psi_{CH}) - \vartheta_\xi\right\}\right)}{2h - (D_x - 2\hat{x}) \sin\left\{\frac{\Phi_c}{2} - \vartheta_\xi\right\}} \quad (10.68)$$

$$y = h \tan\left(\frac{\Psi_{CV}}{2}\right) + \frac{h(2\hat{y} - D_y + D_\zeta - 2\zeta)}{2h - (D_x - 2\hat{x}) \sin\left\{\frac{\Phi_c}{2} - \vartheta_\xi\right\}}$$

To calculate the centred camera data set (${}^k\hat{\mathbf{I}}_{ij}$) from the raw camera data set (${}^k\mathbf{I}_{ij}$), we must now use the appropriate integer coordinates:

$$\begin{aligned} x &= 2h \tan\left(\frac{\Psi_{CH}}{2}\right) \left(\frac{i-1}{N_I-1} \right) \quad i = 1, 2, 3, \dots, N_I \\ y &= 2h \tan\left(\frac{\Psi_{CV}}{2}\right) \left(\frac{j-1}{N_J-1} \right) \quad j = 1, 2, 3, \dots, N_J \end{aligned} \quad (10.69)$$

and

$$\begin{aligned} \theta_\xi &= \frac{(k-1)}{(N_K-1)} \Phi_C \quad k = 1, 2, 3, \dots, N_K \\ \zeta &= \frac{(g-1)}{(N_G-1)} D_\zeta \quad g = 1, 2, 3, \dots, N_G \end{aligned} \quad (10.70)$$

As in Section 10.2.2, we can then use Equations 10.68 through 10.70 with 10.9 to define the two real functions, $\Omega(\hat{i}, k)$ and $\Upsilon(\hat{i}, \hat{j}, k, g)$, which constitute real estimates of the integers i and j :

$$\begin{aligned} \Omega(\hat{i}, k) &= \frac{1}{2} + (N_I + 1) \\ &+ \frac{D_X(N_I - 1)(2\hat{i} - \hat{N}_I - 1) \cos\left[\frac{\Phi_C}{2} \left\{ \frac{(1-2k+N_K)}{1-N_K} \right\}\right] \cot\left(\frac{\Psi_{CH}}{2}\right)}{4h(\hat{N}_I - 1) + 2D_X(1 - 2\hat{i} + \hat{N}_I) \sin\left[\frac{\Phi_C}{2} \left\{ \frac{(1-2k+N_K)}{1-N_K} \right\}\right]} \end{aligned} \quad (10.71)$$

$$\begin{aligned} \Upsilon(\hat{i}, \hat{j}, k, g) &= \frac{1}{2}(N_J + 1) \\ &+ \frac{(\hat{N}_I - 1)(N_J - 1) \left\{ \frac{D_Y(N_G - 1)(2\hat{j} - 1 - \hat{N}_J) - D_\zeta(2g - 1 - N_G)}{(\hat{N}_J - 1)} \right\} \cot\left(\frac{\Psi_{CV}}{2}\right)}{2(N_G - 1) \left\{ 2h(\hat{N}_I - 1) + D_X(1 - 2\hat{i} + \hat{N}_I) \sin\left[\frac{\Phi_C}{2} \left\{ \frac{1-2k+N_K}{1-N_K} \right\}\right] \right\}} \end{aligned} \quad (10.72)$$

It is then a simple matter to define the centred data set via the transformation

$$\begin{aligned} {}^{kg}\hat{\mathbf{I}}_{\hat{i}\hat{j}} &= {}^{kg}\mathbf{I}_{ij} \quad \forall \hat{i} \in \mathbb{N} \mid \hat{i} \leq \hat{N}_I \quad \forall \hat{j} \in \mathbb{N} \mid \hat{j} \leq \hat{N}_J \\ &\quad \forall k \in \mathbb{N} \mid k \leq N_K \quad \forall g \in \mathbb{N} \mid g \leq N_G \end{aligned} \quad (10.73)$$

where, if we use nearest integer interpolation,

$$\begin{aligned} i &\equiv \|\Omega(\hat{i}, k)\| \\ j &\equiv \|\Upsilon(\hat{i}, \hat{j}, k, g)\| \end{aligned} \quad (10.74)$$

Or using bilinear interpolation,*

$$\begin{aligned}
 {}^{kg}\mathbf{I}_{ij} = & \frac{{}^{kg}\mathbf{I}_{\lfloor\Omega\rfloor\lfloor\gamma\rfloor}}{(\lceil\Omega\rceil - \lfloor\Omega\rfloor)(\lceil\gamma\rceil - \lfloor\gamma\rfloor)} (\lceil\Omega\rceil - \Omega)(\lceil\gamma\rceil - \gamma) \\
 & + \frac{{}^{kg}\mathbf{I}_{\lceil\Omega\rceil\lfloor\gamma\rfloor}}{(\lceil\Omega\rceil - \lfloor\Omega\rfloor)(\lceil\gamma\rceil - \lfloor\gamma\rfloor)} (\Omega - \lfloor\Omega\rfloor)(\lceil\gamma\rceil - \gamma) \\
 & + \frac{{}^{kg}\mathbf{I}_{\lfloor\Omega\rfloor\lceil\gamma\rceil}}{(\lceil\Omega\rceil - \lfloor\Omega\rfloor)(\lceil\gamma\rceil - \lfloor\gamma\rfloor)} (\lceil\Omega\rceil - \Omega)(\gamma - \lfloor\gamma\rfloor) \\
 & + \frac{{}^{kg}\mathbf{I}_{\lceil\Omega\rceil\lceil\gamma\rceil}}{(\lceil\Omega\rceil - \lfloor\Omega\rfloor)(\lceil\gamma\rceil - \lfloor\gamma\rfloor)} (\Omega - \lfloor\Omega\rfloor)(\gamma - \lfloor\gamma\rfloor)
 \end{aligned} \tag{10.75}$$

Following the optimisation procedure introduced in Sections 10.2.2 and 10.2.3, we can now define the minimal fixed FOV a camera must have to just fit all of the image data required:

$$\begin{aligned}
 \Psi_{CH} = 2\tan^{-1} & \left\{ \frac{D_X \sin\left(\frac{\Phi_C}{2} + \sin^{-1}\left[\frac{\sqrt{2}\sqrt{h^2(4h^2-D_X^2)\{1+\cos(\Phi_C)\}} - 2D_X h \sin\left(\frac{\Phi_C}{2}\right)}{4h^2}\right]\right)}{2h - D_X \cos\left(\frac{\Phi_C}{2} + \sin^{-1}\left[\frac{\sqrt{2}\sqrt{h^2(4h^2-D_X^2)\{1+\cos(\Phi_C)\}} - 2D_X h \sin\left(\frac{\Phi_C}{2}\right)}{4h^2}\right]\right)} \right\} \\
 \Psi_{CV} = 2\tan^{-1} & \left\{ \frac{D_Y + D_\zeta}{2h - D_X \sin\left(\frac{\Phi_C}{2}\right)} \right\}
 \end{aligned} \tag{10.76}$$

10.3.1.2 Derivation of the I-to-S Transformation

The derivation of the I-to-S transformation very closely follows the procedure we developed in Section 10.2.7. There are two obvious differences: the first is that we need to extend the world coordinates to the full three dimensional system ($\mathcal{X}, \mathcal{Y}, \mathcal{Z}$) as used in the last section. The second is the introduction of the vertical analogue of Equation 10.44:

$$V' = 2h^* \tan\left(\frac{\Psi_{PV}}{2}\right) \frac{(v' - 1)}{(N_v - 1)} \tag{10.77}$$

With these changes in hand, we may rewrite Equation 10.48 as follows

$$\begin{pmatrix} \mathcal{X} \\ \mathcal{Y} \\ \sqrt{h^2 - \mathcal{X}^2} \end{pmatrix} = \begin{pmatrix} X - D_X/2 \\ Y - D_Y/2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} U' - h^* \tan(\Psi_{PH}/2) \\ V' - h^* \tan(\Psi_{PV}/2) \\ h^* \end{pmatrix} \tag{10.78}$$

* See Appendix 7 for a more detailed description of bilinear and bicubic interpolation methods.

from which we derive

$$\mathcal{X} = \frac{\left[\frac{D_X}{2} \right] \left(\frac{2\alpha - 1 - N_A}{(N_A - 1)} \right) \pm \sqrt{\varpi}}{\left\{ 1 + \tan^2 \left(\frac{\Psi_{PH}}{2} \right) \left(2 \frac{(\mu' - 1)}{(N_M - 1)} - 1 \right)^2 \right\}} \quad (10.79)$$

where

$$\begin{aligned} \varpi = & \left[\frac{D_X}{2} \right]^2 \left[\frac{2\alpha - 1 - N_A}{(N_A - 1)} \right]^2 + \left\{ 1 + \tan^2 \left(\frac{\Psi_{PH}}{2} \right) \left(2 \frac{(\mu' - 1)}{(N_M - 1)} - 1 \right)^2 \right\} \\ & \times \left\{ h^2 \tan^2 \left(\frac{\Psi_{PH}}{2} \right) \left(2 \frac{(\mu' - 1)}{(N_M - 1)} - 1 \right)^2 - \left[\frac{D_X}{2} \right]^2 \left[\frac{2\alpha - 1 - N_A}{(N_A - 1)} \right]^2 \right\} \end{aligned} \quad (10.80)$$

and

$$\mathcal{Y} = D_Y \frac{(2\beta - 1 - N_B)}{2(N_B - 1)} + \frac{\tan \left(\frac{\Psi_{PV}}{2} \right) (N_M - 1)(2v' - 1 - N_V)}{\tan \left(\frac{\Psi_{PH}}{2} \right) (N_V - 1)(2\mu' - 1 - N_M)} \left\{ \mathcal{X} - D_X \frac{(2\alpha - 1 - N_A)}{2(N_A - 1)} \right\} \quad (10.81)$$

The camera position (k, g) , is then related to $(\mathcal{X}, \mathcal{Y})$ by the expressions

$$\begin{aligned} k &= \left\| \frac{(N_K + 1)}{2} + \frac{(N_K - 1)}{\Phi_C} \sin^{-1} \left[\frac{\mathcal{X}}{h} \right] \right\| \\ g &= \left\| 1 + \mathcal{Y} \frac{(N_G - 1)}{D_\zeta} \right\| \end{aligned} \quad (10.82)$$

The full **I-to-S** transformation for an inverting camera and a conjugate SLM position may finally be written

$$\begin{aligned} {}^{\mu v} \mathbf{S}_{\alpha\beta} &= {}^{kg} \hat{\mathbf{I}}_{\hat{i}\hat{j}} \quad \forall \alpha \left\{ \alpha \in \mathbb{N} \mid \alpha \leq N_A \right\} \quad \forall \beta \left\{ \beta \in \mathbb{N} \mid \beta \leq N_B \right\} \\ &\quad \forall \mu \left\{ \mu \in \mathbb{N} \mid N_{U1}(\alpha) \leq \mu \leq N_{U2}(\alpha) \right\} \quad \forall v \left\{ v \in \mathbb{N} \mid N_{V1}(\beta) \leq v \leq N_{V2}(\beta) \right\} \\ &= 0 \quad \text{otherwise} \end{aligned} \quad (10.83)$$

where

$$\begin{aligned} \hat{i} &= N_I - \alpha + 1 \\ \hat{j} &= N_J - \beta + 1 \\ k &= \left\| \frac{(N_K + 1)}{2} + \frac{(N_K - 1)}{\Phi_C} \sin^{-1} \left[\frac{\mathcal{X}}{h} \right] \right\| \\ g &= \left\| 1 + \mathcal{Y} \frac{(N_G - 1)}{D_\zeta} \right\| \end{aligned} \quad (10.84)$$

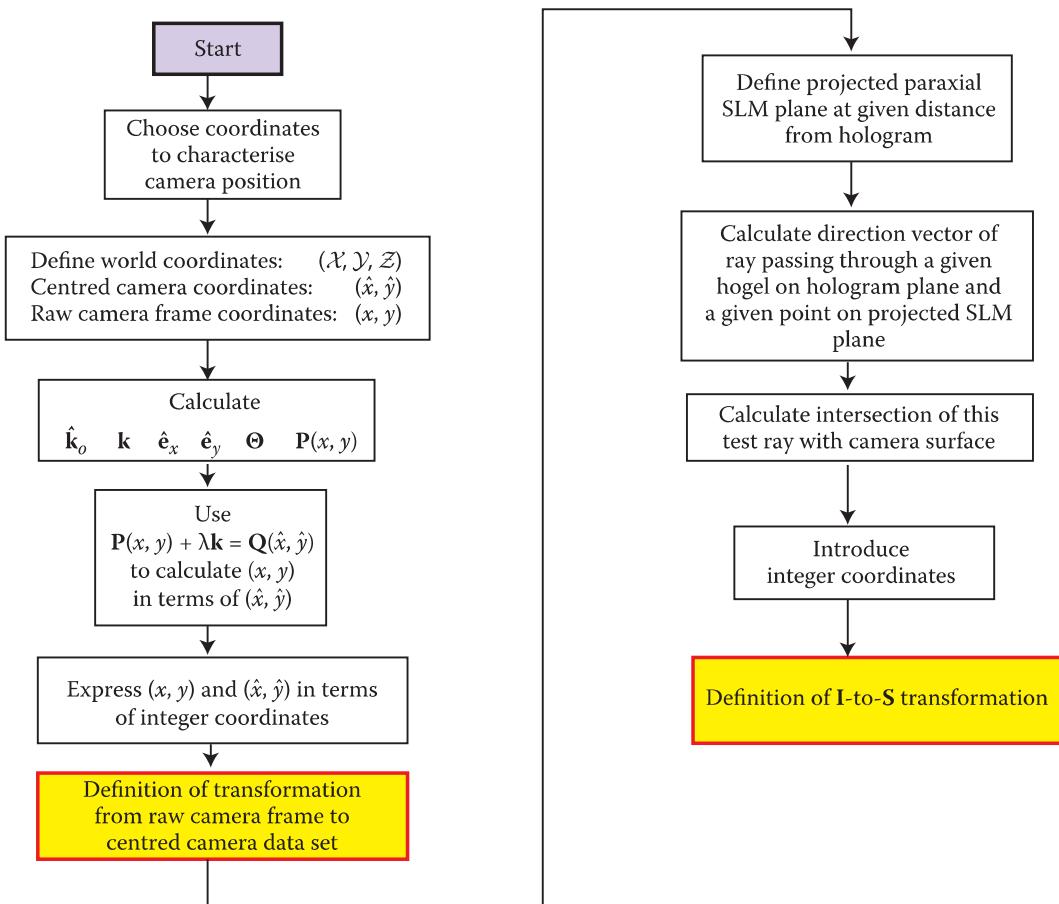


FIGURE 10.19 Flow chart showing the general procedure for defining the algorithms required for a general double-parallax data acquisition (holocam) system. In total, there are 16 basic variants. In each case, one must define an algorithm for the calculation of the centred camera data set and an **I-to-S** transformation.

The limit expressions for $N_{U1}(\alpha)$, $N_{U2}(\alpha)$, $N_{V1}(\beta)$ and $N_{V2}(\beta)$ may be obtained by setting, respectively, $k = 1$ and $k = N_K$ in Equations 10.79, 10.80 and 10.82 and $g = 1$ and $g = N_G$ in Equations 10.81 and 10.82.

The procedure that we have followed to analyse the present camera configuration is summarised in Figure 10.19. This method may be mechanically applied to any of the 16 possible camera configurations to define the algorithm for both the generation of the centred camera frame in addition to the required **I-to-S** transformation.

10.3.2 3D Structured-Light Scanners

For HPO applications, the holocam is really the instrument of choice. It provides a high-resolution data set ideal for the creation of this type of hologram and can be fabricated for a reasonable price. With double-parallax scanning, however, the situation becomes a little less clear. If the object is static, then a two-dimensional holocam may be used and indeed this technique will always provide the highest quality solution. This is because a holocam generates exactly the information required to produce the hologram. Applications, such as cultural heritage or specimen archival, will as such nearly always seek to use a two-dimensional holocam. For certain applications, however, there is sometimes an advantage in using an alternative method that seeks to create an intrinsically 3D model of the scene that is to be reproduced as a hologram. Although a rendering engine must then be used to generate the multiple photographs needed to create the hologram, this approach does allow real-world data to be seamlessly mixed with virtual computer models, something that can be tricky with holocam data. In fact, if the integration of



FIGURE 10.20 Photograph of the low-cost Pico Scan structured-light scanner from 4DDynamics. (Photograph supplied courtesy of 4DDynamics BVBA.)

virtual and real objects is a vital requirement, then such alternative methods may even be considered for HPO holograms.

A data acquisition system that is able to produce such 3D model files directly is called the “structured-light scanner”. At the time of writing, the price-point for such scanners is dropping quickly and useful commercial devices are starting to appear. However, the time of data acquisition is still rather high for all but the highest-range models—usually around the 20 s mark, making such systems more suitable for static objects rather than for human subjects. An example of a low-cost system is the PicoScan product from the Belgian company 4DDynamics, which was priced at the time of writing at just under €3000 (Figure 10.20).

Structured-light scanners use a projector to project a pattern of lines onto the object to be scanned; a camera then photographs the distorted lines and uses this distortion data to infer the geometry through a simple mathematical transform. Problems occur when the structured light disappears down a hole and the camera or cameras cannot see what is happening in this region; reflective and translucent materials* provide additional problems. Equally, lines can cross-over and it can be difficult to tell which line is which. In such cases, the mathematical transformation becomes singular and it can be difficult or impossible to reconstruct the exact geometry. However, by using more complex patterns of the structured light, time modulation, many cameras and more complex image-processing algorithms, it is possible to address the most frequently encountered of these problems.

An application that has driven fast high-end, structured-light scanners is the film industry, where full-body scanning of human subjects is being increasingly called for. Systems here often integrate many cameras and projectors and a computer then seamlessly integrates all the data. Frequently, algorithms are used to correct for slight movements during the data acquisition period such as heartbeats or breathing. Scanning times can be reduced in these systems to well under a second. This, however, comes at a cost. Such systems are extremely expensive. For example, a minimum full-body scanner setup from 4DDynamics consists of four CX-PRO scanners (Figures 10.21 and 10.22) each costing approximately €15,000 with a typical integrated system including installation and software, coming in at just under the €100,000 mark. Although the output can be in a point-cloud format, most 4DDynamics users choose to output a mesh including textures. Postprocessing software, which is included by the company, permits the user to opt for vertex colour output or (UV) texture map to a watertight mesh. Depending on the settings, a merging of four raw body scans on a standard laptop takes approximately 6 min.

* Human skin can, in particular, present problems due to the phenomenon of subsurface scattering.



FIGURE 10.21 Photograph of the high-end CX-PRO 4DDynamics structured-light scanner suitable for body-scanning applications. (Photograph supplied courtesy of 4DDynamics BVBA.)



FIGURE 10.22 Rendered images produced from the 3D watertight mesh and texture maps generated by a CX-PRO 4DDynamics structured-light scanner. (Photograph supplied courtesy of 4DDynamics BVBA.)

10.4 Images Derived from a Virtual Computer Model: HPO Holograms

In this section, we will cover how the popular 3D design program 3D StudioMax version 8 can be programmed to produce a rendered image sequence suitable for the production of an HPO reflection hologram. Many other commercial programs are capable of generating the type of image data required for digital display holograms, but space prevents us from discussing each programming environment.

Before proceeding, however, we shall mention a few programs that are particularly suited to generating the image data for direct-write and master-write digital holograms (DWDH and MWDH). First and foremost is the open-source program Blender, which uses a scripting language based on Python. This is an excellent place to start with 3D modelling for digital holography. Other commercial programs to note at the time of writing are SoftImage, Cinema 4D, Maya and LightWave.

10.4.1 Data Preparation

Each digital holographic printer will have its own unique recommended parameter set. This will be available from the manufacturer or from the printer operator in the case of a service bureau. For the purposes of this section, we shall use information published by Geola Digital UAB, which operates

TABLE 10.1

Recommended Printing Parameters for HPO Reflection Holograms

Hologram Height (mm)	Hologram Width (mm)	Horizontal Camera FOV (°)	Camera Track (mm)	Camera Distance (mm)	Frames Rendered	Frames Rendered	Image Size (0.8 mm Hogel)		Image Size (1.6 mm Hogel)	
					(High Angular Resolution)	(Low Angular Resolution)	Pixels	Pixels	Pixels	Pixels
200	x300	82.79	793.6	619.0	497	—	1366	x250	683	x125
300	x400	82.72	1060.8	828.7	664	640	1826	x374	913	x187
490	x640	82.83	1704.0	1327.9	1066	640	2930	x612	1465	x306
640	x900	82.81	2395.2	1867.0	1498	640	4118	x800	2059	x400
999	x1250	82.83	3331.2	2595.8	2083	640	5726	x1248	2863	x624
999	x1500	82.85	4000.0	3115.0	2501	640	6874	x1248	3437	x624
300	x200	77.41	793.6	619.0	497	—	1242	x374	621	x187
490	x330	77.48	1300.8	1015.0	814	640	2038	x612	1019	x306
640	x490	79.07	1704.0	1327.9	1066	640	2742	x800	1371	x400
999	x750	78.83	2660.8	2073.3	1664	640	4262	x1248	2131	x624

Source: Courtesy of Geola Digital UAB.

a DWDH reflection holography printing service from its offices in Vilnius, Lithuania. Table 10.1 lists Geola's recommended parameter set. Here, the supported formats are listed as well as a series of parameters for each such format. These include the camera distance, h , the required camera track length, $D_\xi = W$, the horizontal camera FOV, Ψ_{CH} , the number of rendered frames, N_K and the pixel dimensions of the individual photographs, N_i and N_j . Note that Geola offers two hogel diameters, 0.8 and 1.6 mm. It also offers a high and low angular resolution for scenes with large or small depths.

To illustrate how a file is prepared in 3D StudioMax, we will treat in detail the case of a 20 cm × 30 cm landscape format high-resolution hologram having a hogel diameter of 0.8 mm. Table 10.1 may then be used to create any of the other formats by simply replacing the relevant parameters in the description given below.

10.4.2 Creating a 20 cm × 30 cm Landscape Hologram

We will review here the simplest technique in 3D StudioMax to create a data set for the *simply translating camera*. In Section 10.5, we will cover centred camera configurations in the context of double parallax. These techniques, which use the scripting capabilities of Max, can easily be applied to the HPO case. For now, however, we will use the animation facility in Max to generate the required HPO data set.

We start by launching 3D StudioMax with the scene we wish to use. Note that the model we use here (a Ducati motorbike) has been scaled to a larger dimension than required for the final holographic image. We shall therefore use a scale of 1:10 to reduce the image to an appropriate size for our present application. Note that the world coordinates (0,0,0) represent the exact centre of the hologram.

First, we must go to the time configuration panel to set the number of animation frames, N_K (right click on any animation control at the bottom right—i.e., the “▶” button). The animation end time will need to be changed from 100 to 496. Press “OK” when you have done this. This will give a total of 497 frames in the animation sequence, which is what we require from Table 10.1 (Figure 10.23).

Now we create a “Free” camera that faces (0,0,0) and is positioned at a distance of 619 mm in front of the hologram.* You can do this by selecting the camera panel from the top right-hand graphical user interface (GUI) menu, selecting **Free** and then clicking anywhere in the “Front” window. Then, ensuring that the free movement icon is selected (the crossed arrows in the top row of large GUI icons) enter the (X,Y,Z) coordinates at the bottom centre of the screen as -396.8, -619.0 and 0.0.

* Using the 10:1 scale, we therefore set a distance of 619 cm in Max.

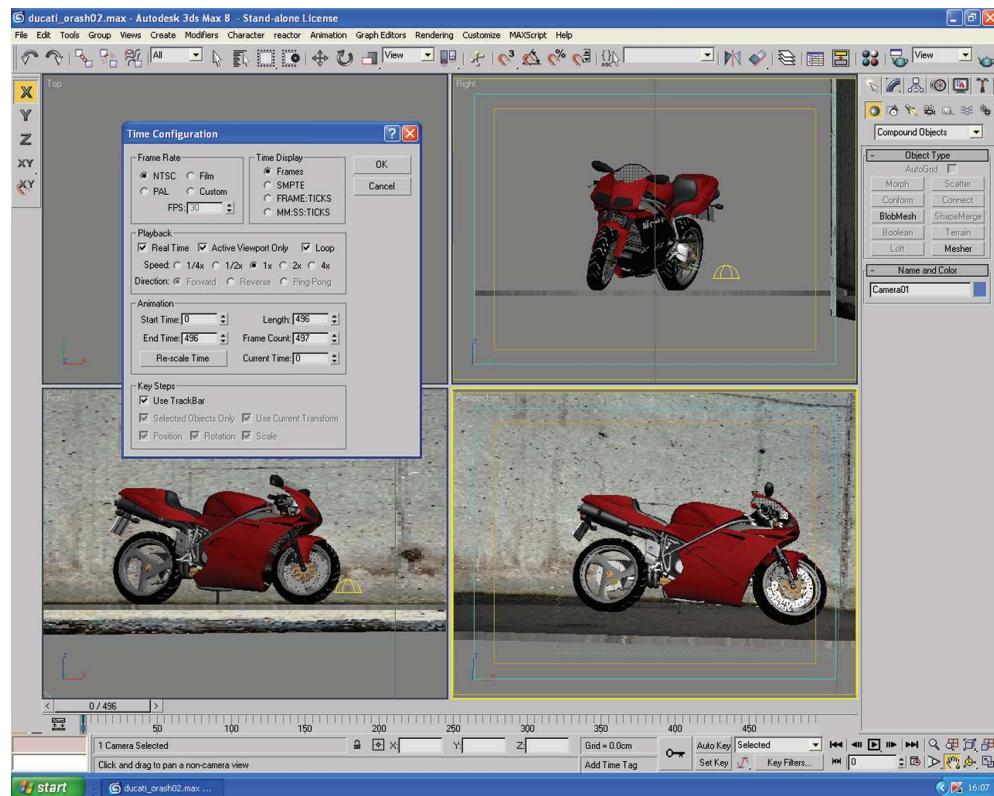


FIGURE 10.23 To set the number of frames in the animation sequence, open the time configuration dialogue in Max by right-clicking the “▶” button at the bottom right-hand corner of the GUI.

We can enter the starting *X* value for the camera animation: To do this, activate the **Set Key** toggle button* at the bottom left of the screen, type -396.8^{\dagger} in the central *X* box and click on the key icon to the immediate left of the **Set Key** toggle control. The default frame is 0, and so this sets frame 0 to an *X* value of -396.8 . Move the slider now to go to frame 496; this time, enter a value of $+396.8$ in the *X* box and click on the key icon. Press the **Set Key** button again to exit the key framing mode. At this point, you should have a camera travelling from left to right from frame 0 to 496.[‡]

The next thing we need to do is to make the path of the camera linear. Right click on the camera and select the curve editor. In the curve editor, select World > Objects > Camera01 > Transform > Position *X* position. The screen should then look like Figure 10.24. Make the curve linear by selecting the two end points in turn and clicking the **Set Tangents to Linear** icon[§] in the top menu. The graph should then look like that of Figure 10.25.

We now need to set up the FOV of the camera. To do this, select the camera and click on the modify panel.[¶] Change the **FOV** value to 82.79° and make sure the little icon to the left looks like two horizontal arrows (Figure 10.26). You can also tick the **Show Cone** option and select a **Near Range** of 0 cm and a **Far Range** of 619 cm. This will draw in each of the design windows the camera’s cone of vision just up to the hologram surface.

* The animation slider bar will now turn red.

[†] Again, this is actually -396.8 cm given the 10:1 scale we are using.

[‡] You can verify this by moving the slider and watching the camera animate.

[§] This is the icon with a 45° red line.

[¶] The modify panel is the icon that looks like a quarter tyre at the top right of the screen (second to the left, second row down).

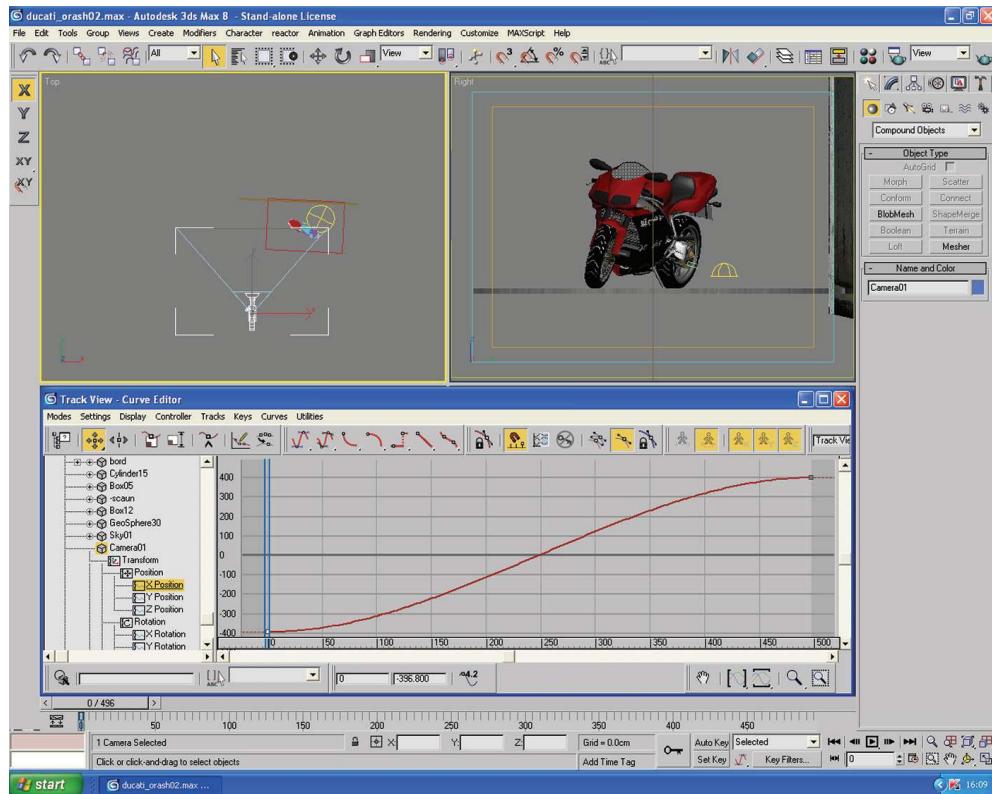


FIGURE 10.24 Setting the camera animation through the Curve Editor in Max, which can be accessed by right-clicking on the camera and selecting **Curve Editor** from the drop-down menu.

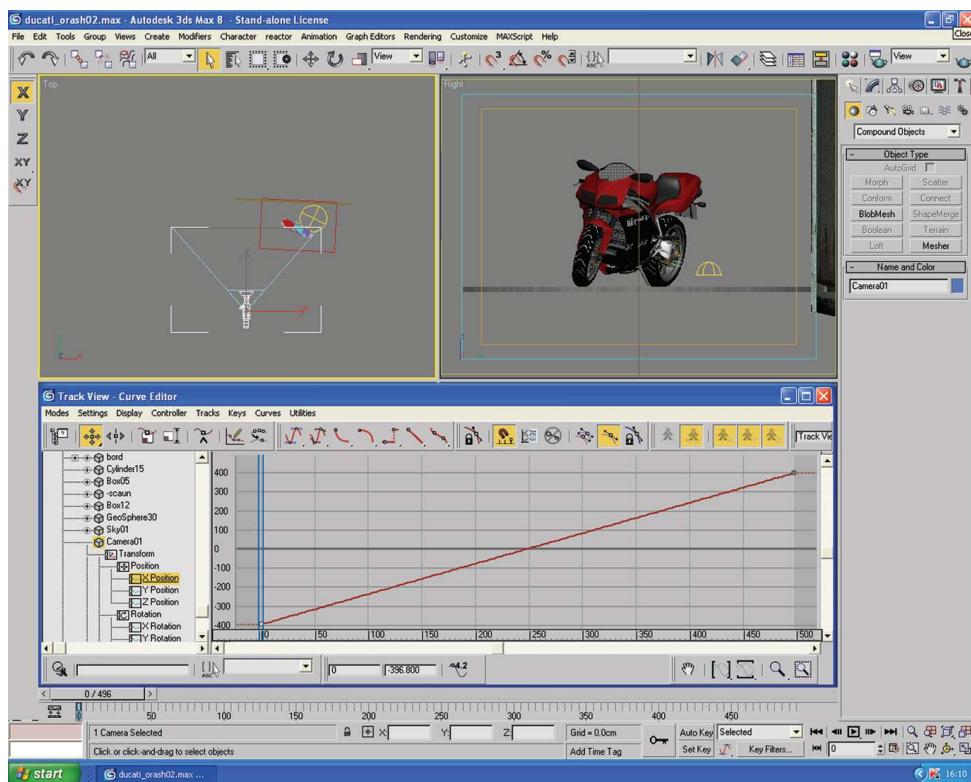


FIGURE 10.25 By selecting the two end-points in turn and using the **Set Tangents to Linear** button, the position animation of the camera becomes linear.

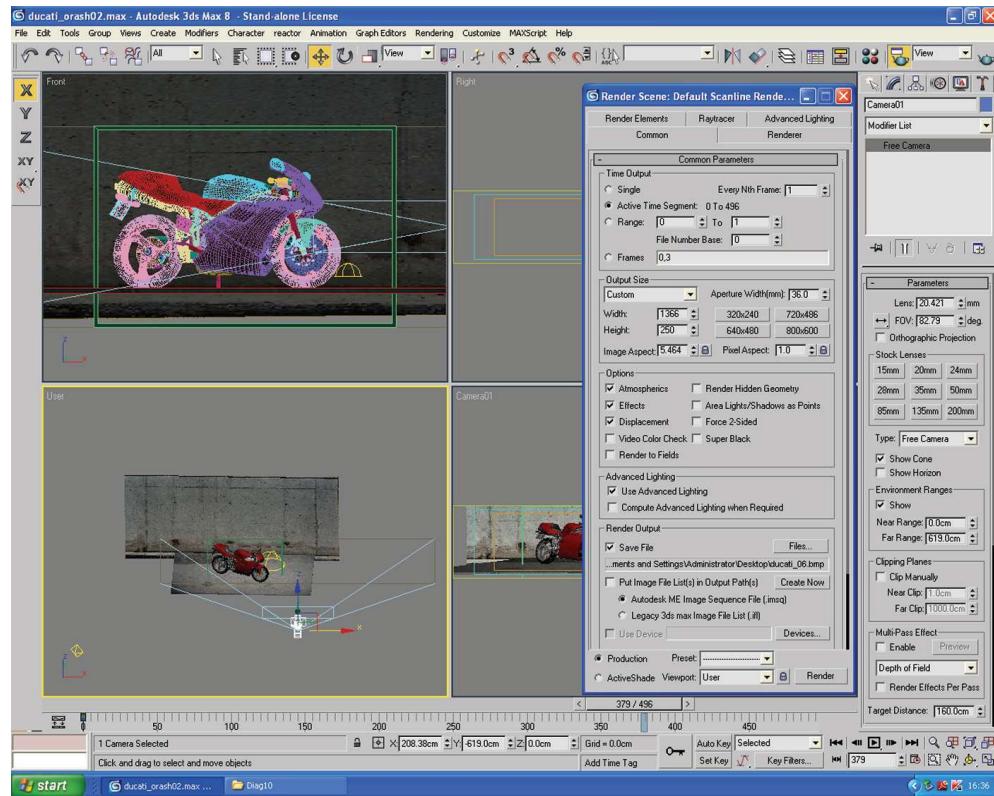


FIGURE 10.26 Setting the render parameters in Max from the render dialogue.

To set the camera resolution (Figure 10.26), open the render menu (F10 in windows or the teapot icon at the top of the GUI) and enter the **Output Size** panel **Width: 1366** and **Height: 250**. Choose the **Active Time Segment** option in the **Common Parameters** panel to tell Max to render all your files. Tick the **Save File** option in the render Output panel and then input a rendering directory by pressing the **Files** button. Make sure that the renderer is set to output bitmaps in 24 bits (16 M colours, no alpha channels). Also, check that the **ViewPort** at the bottom of the render panel is set to the active camera. You will need to define one of the viewports as the camera port (e.g., **Camera01**) before the drop-down list will allow you to select the camera. You do this by right-clicking a viewport name (at the top-left of each viewport) and selecting the camera view from the **Views** entry.

Finally, it is a good idea to create a rectangular hologram frame to make sure everything is aligned properly. To do this, create a 300×200 rectangle centred at (0,0,0) with normal pointing in the Y direction.* This is used as a reference to check how the hologram will look. Everything that is in front of this rectangle (closer to the camera) will appear to “come out” of the hologram and everything behind it will give the impression of being “inside” the hologram. Obviously, you do not want to render this “hologram frame” on the final output so you can delete it when you are happy with the scene. You should check that the edges of the rectangle exactly match the camera frame at the extreme camera positions; this means that everything has been entered properly. When objects are placed in the scene, if something is clipped at the edges, it will also be clipped in the final hologram. Make sure to centre your objects accordingly. By moving the animation slider from frame 0 to 496, you should now obtain views similar to those of Figure 10.27.

When you are happy with the setup, simply press render on the render menu, and this will generate the required series of 496 24-bit BMP files each having dimensions of 1366×250 . Figure 10.28 shows a selection of the 497 files generated. Figure 10.29 shows the actual 20×30 cm DWDH reflection hologram together with a $1 \text{ m} \times 1 \text{ m}$ format made using the same Max file and the corresponding parameters listed in Table 10.1.

* You might want to extrude the rectangle and then use a Boolean operation under the compound object menu to create a more visible picture frame.

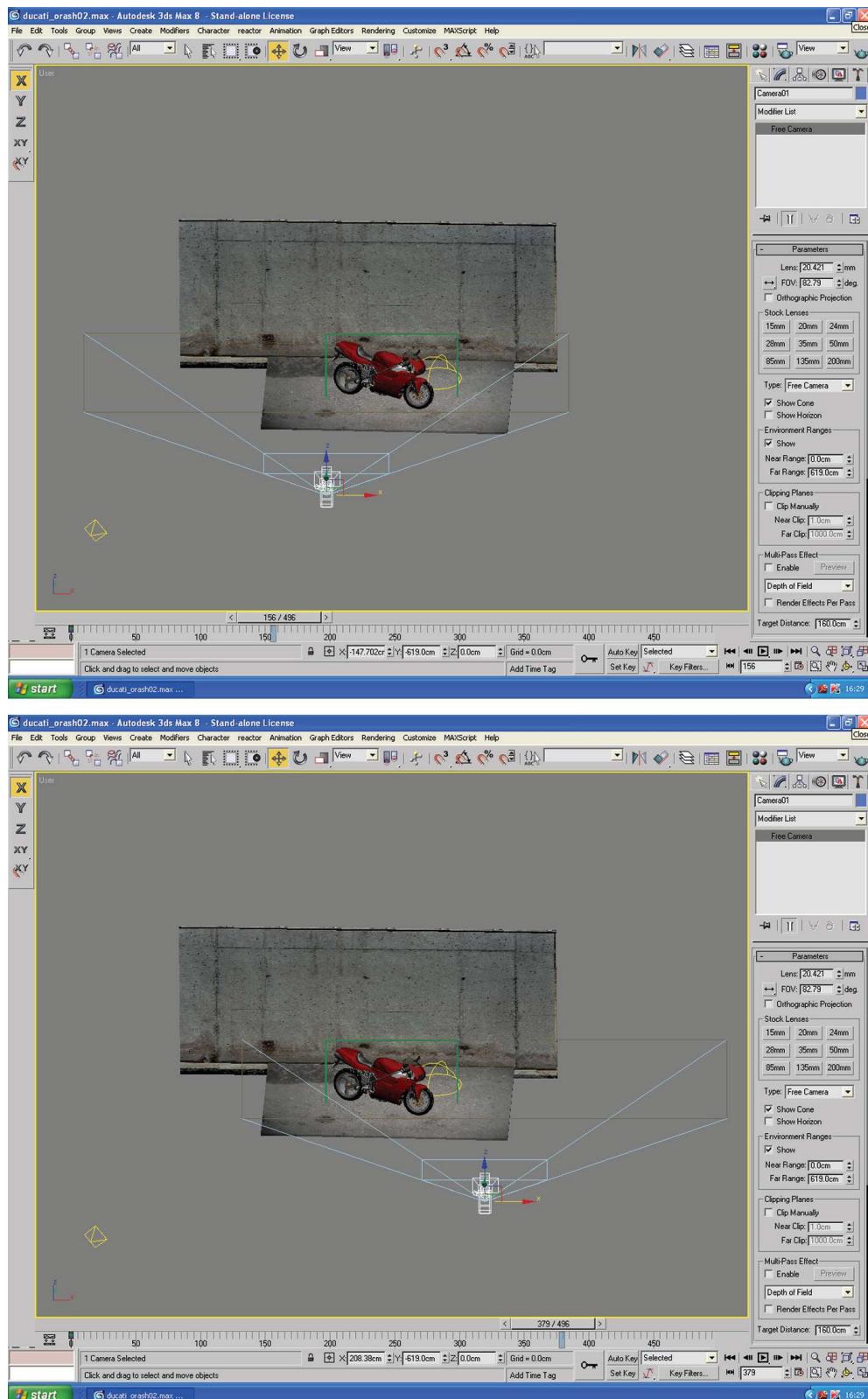


FIGURE 10.27 This is how the scene should look like in Max before rendering. The figure shows two diagrams, one with the animation slider at frame number 156 and the other at frame number 379. The first frame is frame 0 and the last frame is 496.

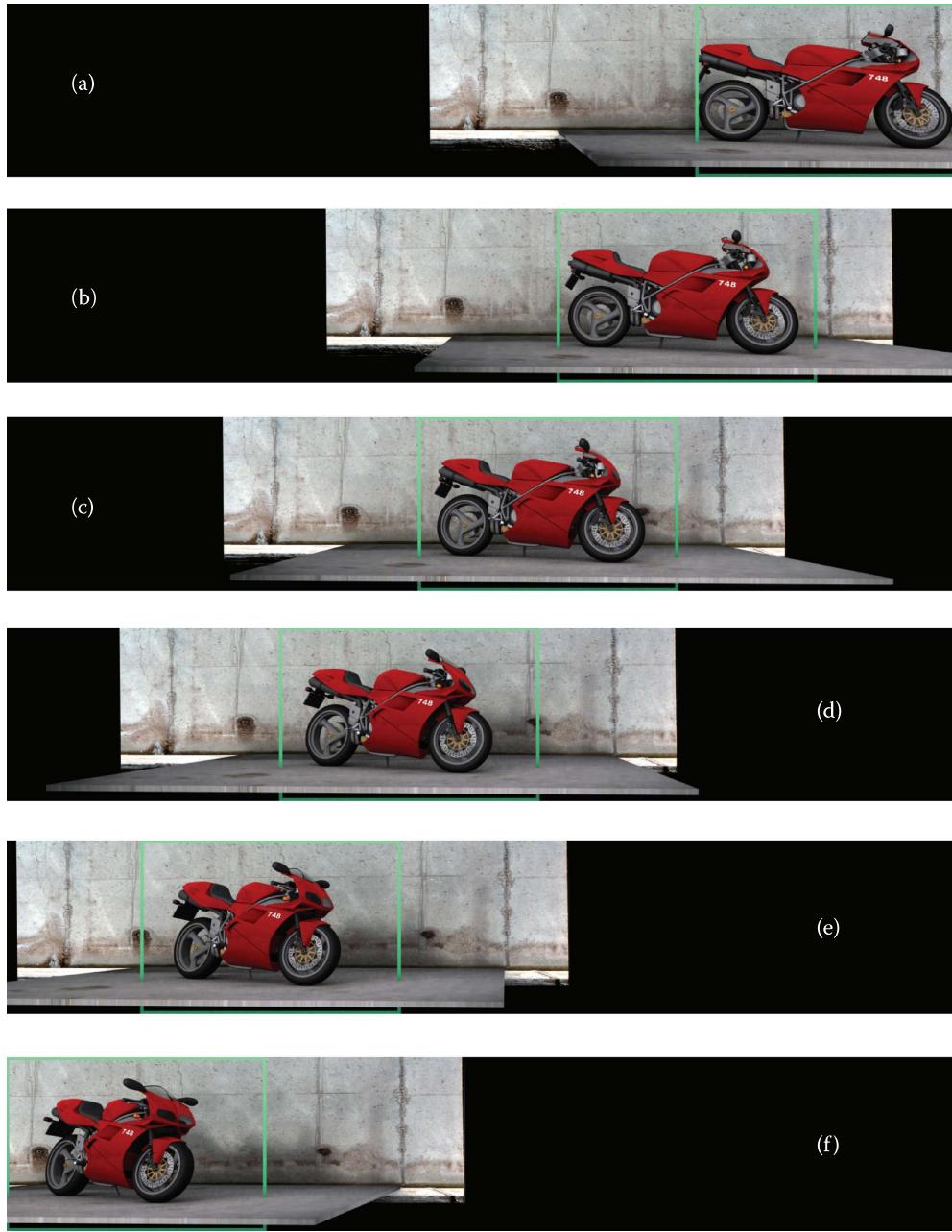


FIGURE 10.28 Six of the 497 rendered images of the Ducati motorbike produced by the animation sequence described in the text: (a) image 1, (b) image 100, (c) image 200, (d) image 300, (e) image 400 and (f) image 497. The camera geometry is a *simply translating camera*. Note the green frame that has been included to emphasise the hologram location within the data set. The data inside the green frame corresponds to a centred camera data set.

A good rule of thumb for any hologram is the following: for an optimum volume effect, generally one should aim to have one-third of the image in front of the hologram and two-thirds behind. Of course, this depends on the type of scene that is being generated to a greater or lesser extent. For a 30 cm × 20 cm hologram, the user should usually limit objects to a maximum of approximately 6 to 7 cm in front of the film; this limit not only ensures the sharpness of the object (particularly when using silver halide emulsions and non-diode illumination) but also prevents too much clipping.

For animation in an HPO hologram, it must be slow and subtle. It must not be treated like video (fast panning, transitions, and zooms are a definite no). Good examples are a cross-dissolve of an object to reveal the interior (given that the dissolve envelope is very long and linear), a character slowly waving hands, etc. Bad



FIGURE 10.29 Photograph of (a) final $20\text{ cm} \times 30\text{ cm}$ HPO DWDH refection hologram and (b) photograph of a larger format ($1\text{ m} \times 1\text{ m}$) HPO DWDH reflection hologram produced from the same Max file. Both holograms have a hogel diameter of 0.8 mm. (Image created in 3D StudioMax by Razvan Maftei, 2005. Photographs courtesy of Geola Technologies Ltd.)

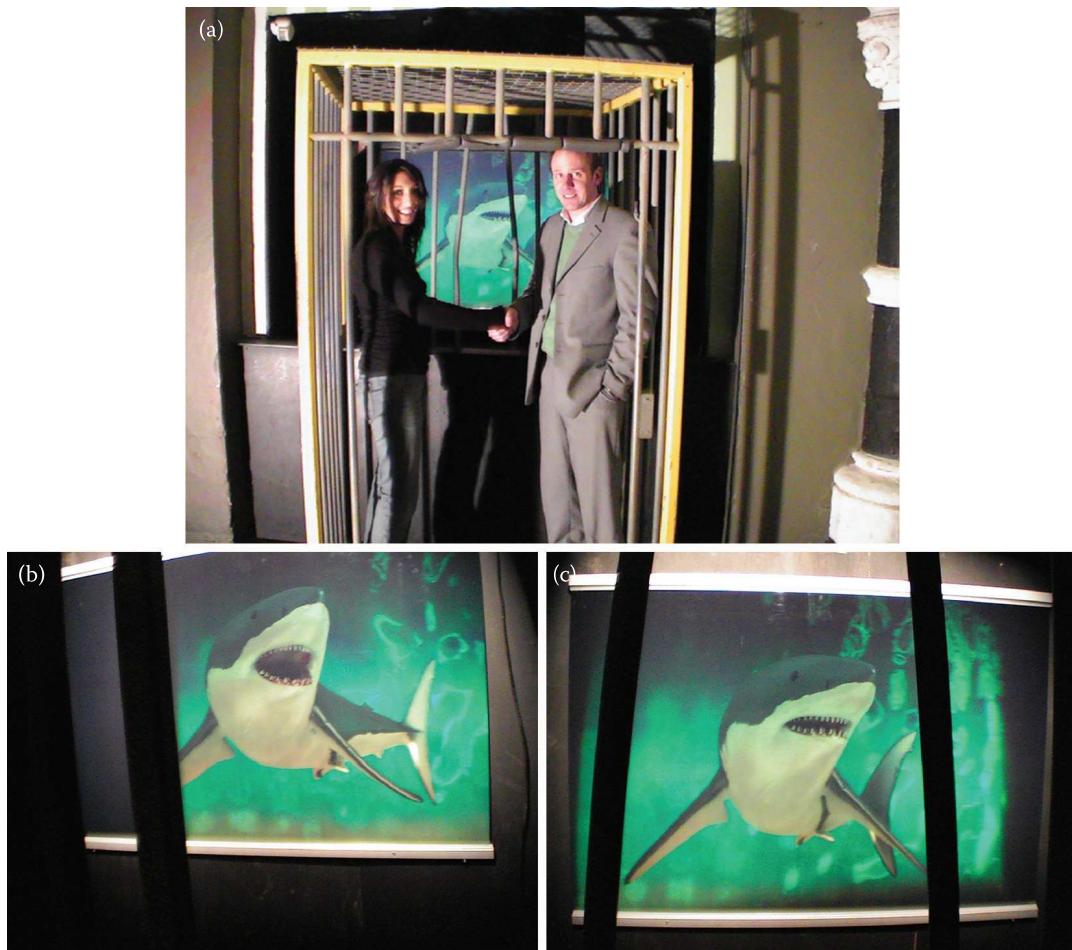


FIGURE 10.30 Three photographs showing a large-format ($1\text{ m} \times 1.5\text{ m}$) HPO DWDH hologram of a computer-created and completely life-like great white shark. The hologram was installed behind a metal cage in the SeaLife centre in Brighton, UK, to simulate viewing from an underwater cage. The hologram featured a slow animation as is clear from the pictures. (a) Installation of the hologram in 2005, (b) view from the right and (c) view from the left. (Image created in 3D StudioMax by Razvan Maftei, 2005. Photographs courtesy of Geola Technologies Ltd.)

examples are systems of movement that act to mimic the parallax; these will change the perceived depth of the hologram—sometimes radically! Also, every animation key frame should be linear; never use splines!

Holograms are replayed using light; they will reflect or transmit light information according to the data files used to print them; making a scene in darkness like a scene from *Batman* at night is unlikely to be as good as the cliché “bowl of fruits” given the performance of current silver halide emulsions. Scenes should therefore be bright and contrasted. Avoid using ambient lighting to compensate for the lack of overall brightness. This can kill the contrast of the scene. A good hologram always has a good lighting rig in the 3D creation software.

One further point to make about HPO holograms is that “streaking” across the physical surface of the hologram can appear. This phenomenon is caused by bright spots close to the hologram surface. Although such bright spots stand still as the observer moves from right to left, if a bright spot is situated slightly in front of the hologram surface, a corresponding high-intensity spot will appear to track from right to left on the hologram surface. This causes a line of high-intensity light at the hologram surface that exceeds the normal



FIGURE 10.31 Two HPO DWDH holograms created from computer models. The top image shows a hologram installed by Geola in the NIKE store in Oxford St., London in 2005. The bottom image is of an early HPO DWDH hologram created by XYZ Imaging Inc., circa 2003/2004, seen here with the company's founding president Eric Bosco. Note that the bottom image is made up from nine “tiles” which have been printed separately.

linear response of the emulsion, causing a “burn”. As well as being unsightly, the hologram replay characteristics within the burnt area are modified and the colour balance and brightness of objects behind the burn can change when viewed through the overexposed area. This effect is similar to the effect seen in analogue holograms, in which light is focussed too much at a given point on the surface of the hologram leading to a burn pattern. The difference with HPO holograms is that the burn pattern is nearly always elongated into a line due to the HPO nature of the recording. As such, one needs to be extremely careful when using highly reflective surfaces in HPO holograms if employing recording materials with a limited dynamic recording range.

10.5 Images Derived from a Virtual Computer Model: Full-Parallax Holograms

We have seen in previous sections how a linear tracking camera can be programmed in 3D StudioMax to generate all the virtual camera frames required for an HPO hologram. In this section, we will discuss how to program a virtual camera that tracks in two dimensions to generate the camera frames required for a full-parallax digital hologram. To do this, we will use the scripting language of Max known as MAXScript. This approach is in fact often used with HPO holograms as well. In particular, MAXScript allows one to directly produce a *centred camera* output rather than a *simply translating camera* output. Most 3D design programs offer scripting languages and this is often the most convenient technique.

10.5.1 Creating a MAXScript GUI

We will now discuss how to program using MAXScript, a basic script that automatically generates the rendered camera images (using a centred camera configuration) required for the printing of a double parallax three-colour DWDH reflection hologram. To do this, we will use the results from Chapter 9 and, in particular, we will use a first quantised camera distance to achieve a DWDH data set free from interpolation errors.

First of all, Max must be started and the chosen 3D model file opened. Once this has been done, click on the **MAXScript** menu in the top bar of the Max GUI and select **New Script** from the drop-down menu. A blank MAXScript editing window will now appear. Select **Edit** from the top bar menu of the MAXScript edit window and choose **Edit Rollout**. This will bring up the Visual MAXScript editing window (Figure 10.32). This is where you will be able to design the GUI for the MAXScript program.

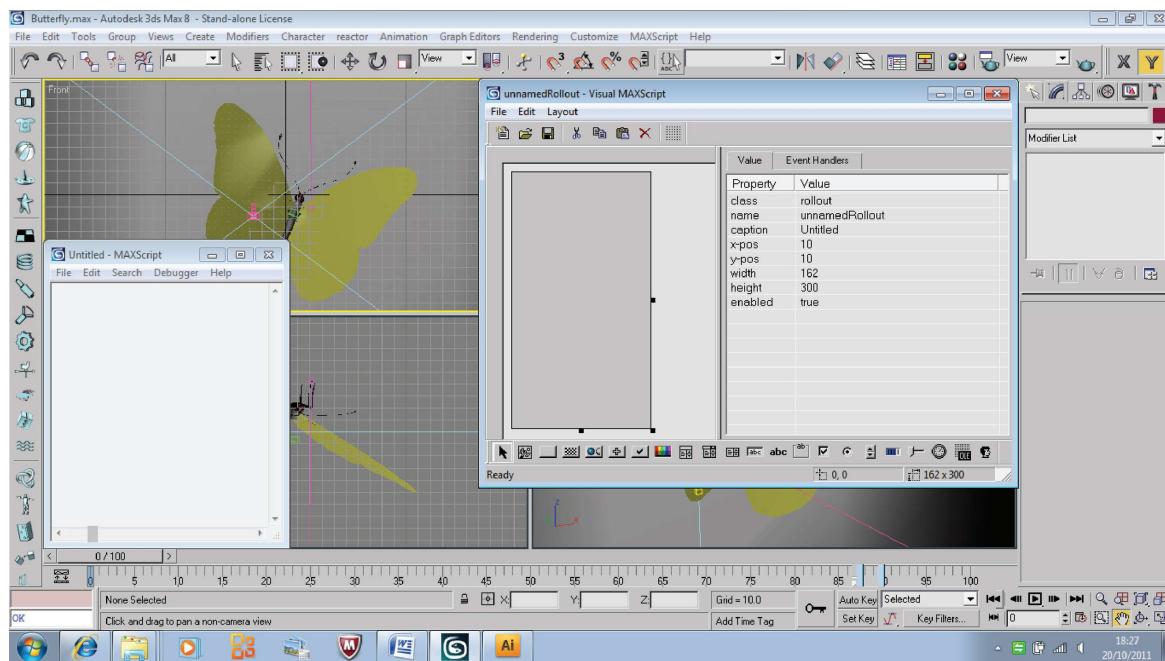


FIGURE 10.32 Visual MAXScript editing window.

The visual MAXScript window, by default, contains a blank panel and, at the bottom, a list of possible buttons and functions that can be added to this panel. The panel to the right lets you set the names and values of these controls. It also lets you tell Max what code to generate when an event is generated.

The first thing we shall do is to add a button. We will use this button to start the animation sequence of the rendering process, so we shall call it “Start Render”. To add the button to the panel, simply click once on the blank square icon on the bottom menu row of the Visual MAXScript editing window. Then click, hold and drag out the shape of the button desired on the panel. Now go to the right-hand windowpane and select the **Value** tab. Here, you can fill out a name for the button—we shall use BRen—and the caption “Start Render” (Figure 10.33).

By clicking on the panel itself, you will bring up in the right-hand pane the details of the panel itself. We will change the name and caption to “DoubleParallaxRender”.

Because it is good practice to save your work from time to time, click on the disk symbol in the top icon menu bar of the visual MAXScript editing window now or, alternatively, select **Save** from the drop-down **File** menu. You will see that this immediately generates the following code in the MAXScript text editing window:

```
Rollout DoubleParallaxRender "DoubleParallaxRender" width:162 height:300
(
    Button BRen "Start Render" pos:[11,24] width:132 height:44
)
```

This code simply defines the button you have created. Braces indicate sections of self-contained code. To prompt Max to include code for an event handler, simply select the button and then choose the **Event Handler** tab in the right-hand windowpane. Tick the **pressed** option and then save again using the disk symbol. The text edit window should now look like this:

```
rollout DoubleParallaxRender "DoubleParallaxRender" width:162 height:300
(
    button BRen "Start Render" pos:[11,24] width:132 height:44
    on BRen pressed do
    (
    )
)
```

By placing your own code between the two new braces, you will be able to tell Max what to do when the button is pressed. For now, however, we will continue with populating the panel by adding three more buttons: “Make Directories”, “Make Camera” and “Calculate”. We will name these buttons “BMD”, “BCam” and “BCal”. For each, we must create event handlers (tick the **pressed** option). We will then

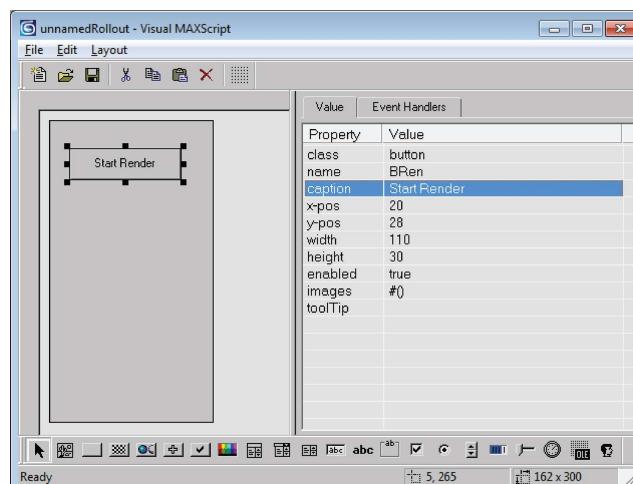


FIGURE 10.33 Programming the “Start Render” button in the Visual MAXScript editing window.

use the “Group Box” control to lay things out sensibly. This is the MAXScript control that looks like a little square with an “ab” at its centre top. Click on this control and then click, hold and drag until a box appears on the panel. The group box is useful for keeping bits of information together—you can change the caption in the **value** tab to the right of the visual MAXScript editor. We will want to make six group boxes which we will entitle “SLM Definition”, “Printer Optics Definition”, “Window Parameters”, “Hologram Definition”, “Calculated Parameters” and “Job Definition”. We will then populate these group boxes with three types of control. These are the Spinner, the Edit Box and the Label. The spinner, which looks like an up-facing and down-facing arrowhead stacked one on top of the other, allows the controlled input of numeric values. You can specify in the **value** tab the name, caption, range, default value and the type of input. We shall always use the integer input type. The range format is (low value, high value, default value at time of execution).

The Edit Box is very similar to the spinner control except that we will be using it for text input. We will just need the name and caption property here. Finally, the Label control is useful as an indicator so that Max can display information on chosen variables while it is working. With two of the spinner controls, we will choose to restrict the valid input to even numbers. This is the case for the hologram size expressed in hogels (N_A and N_B). Although we could choose to accept odd numbers here, certain of the integer formulae change and therefore it makes sense in the context of this simple example to restrict ourselves to the even cases. To do this, we select the spinner control and then click on the event handler tab in the right-hand pane and tick **changed**. Upon saving, this then generates the following code:

```
on NA changed val do
(
)
```

We can then insert a command within the brackets to select only even input:

```
on NA changed val do
(
    NA.value=2.0*floor(NA.value/2.0)
)
```

With these instructions in mind, you should be able to create the MAXScript panel shown in Figure 10.34. Table 10.2 lists the properties set for each control and Figure 10.35 lists the code Max generates to define the panel.

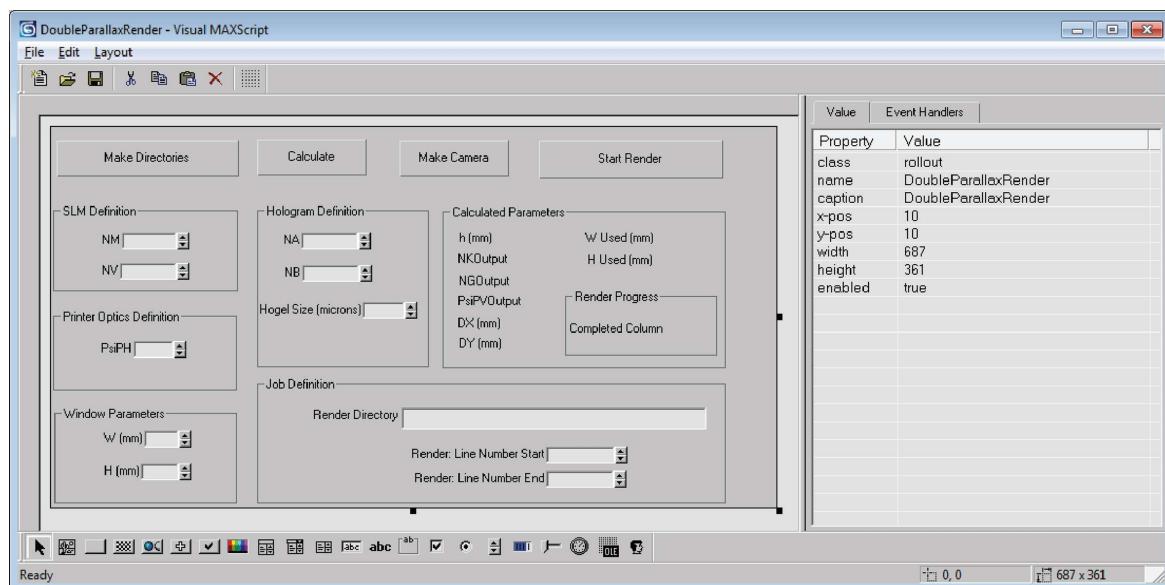


FIGURE 10.34 Finished MAXScript panel.

TABLE 10.2

MAXScript Parameters and Settings for the MAXScript Panel of Figure 10.34

<i>Buttons</i>				
Name	Caption	Event Handler		
BMD	Make Directories	Pressed ticked		
BCal	Calculate	Pressed ticked		
BCam	Make Camera	Pressed ticked		
BRen	Start Render	Pressed ticked		
<i>Spinners</i>				
Name	Caption	Range	Type	Event Handler
NM	NM	[100, 5000, 1280]	Integer	Change ticked
NV	NV	[100, 5000, 1024]	Integer	Change ticked
NA	NA	[100, 3000, 600]	Integer	Change ticked
NB	NB	[100, 3000, 400]	Integer	Change ticked
PsiPH	PsiPH	[10, 120, 89]	Integer	
Win	W (mm)	[10, 5000, 600]	Integer	
Hin	H (mm)	[10, 5000, 400]	Integer	
Delta	Hogel size (μm)	[10, 3000, 800]	Integer	
NStartLine	Render: line number start	[1, 5000, 1]	Integer	
NstopLine	Render: line number stop	[1, 5000, 1]	Integer	
<i>Edit Boxes</i>				
Name	Caption			
RenderDir	Render directory			
<i>Labels</i>				
Name	Caption			
hOutput	h (mm)			
NKOutput	NKOutput			
NGOutput	NGOutput			
PsiPVOOutput	PsiPVOOutput			
DXOutput	DX (mm)			
DYOutput	DY (mm)			
W_Output	W Used (mm)			
H_Output	H Used (mm)			

10.5.2 Function of Script

The MAXScript GUI window described in the previous section forms the basis of a script that will generate the rendered camera images required to make a double parallax three-colour DWDH reflection hologram. The script has four buttons, each of which, through its individual event handler, controls a separate function; these functions will be described by code, which we will write shortly. The buttons are designed to be pressed in sequence from left to right. The first button is the “Make Directories” button. This will take the render director we specify and will either make sure that it exists or create it. It will also create subdirectories for each line of the hologram.

The next button is the “Calculate” button which takes the various inputs and calculates some useful parameters including the first quantised camera distance, h , the actual physical size of the hologram, $D_x \times D_y$, the actual window size which is subject to an adjustment, $W \times H$, the number of camera views required, $N_K \times N_G$, and the vertical printer FOV, Ψ_{PV} . The reason for displaying these values is to let the

```

G DP.ms - MAXScript
File Edit Search Debugger Help
rollout DoubleParallaxRender "DoubleParallaxRender" width:687 height:361
{
    button BRen "Start Render" pos:[462,13] width:175 height:36
    edittext RenderDir "Render Directory" pos:[248,266] width:373 height:21
    label hOutput "h (mm)" pos:[386,98] width:108 height:18
    spinner Wm "W (mm)" pos:[59,288] width:7 height:16 range:[0,5000,600] type:#integer
    spinner HM "H (mm)" pos:[59,319] width:7 height:16 range:[0,5000,400] type:#integer
    spinner NA "NA" pos:[229,100] width:74 height:16 range:[100,3000,600] type:#integer
    spinner NB "NB" pos:[230,131] width:74 height:16 range:[0,3000,400] type:#integer
    GroupBox grp1 "Calculated Parameters" pos:[371,74] width:268 height:155
    GroupBox grp2 "Hologram Definition" pos:[195,73] width:161 height:155
    GroupBox grp3 "Window Parameters" pos:[3,265] width:173 height:92
    GroupBox grp4 "SLM Definition" pos:[2,73] width:175 height:83
    spinner NM "NM" pos:[57,100] width:74 height:16 range:[100,5000,1280] type:#integer
    spinner NV "NV" pos:[57,129] width:74 height:16 range:[100,5000,1024] type:#integer
    label NKOutput "NKOutput" pos:[385,118] width:66 height:18
    label NGOutput "NGOutput" pos:[386,139] width:74 height:18
    GroupBox grp5 "Printer Optics Definition" pos:[2,173] width:174 height:77
    spinner PsiPH "PsiPH" pos:[56,203] width:72 height:16 range:[0,120,89] type:#integer scale:0.1
    spinner NStopLine "Render: Line Number End" pos:[353,326] width:192 height:16 range:[1,5000,1] type:#integer
    GroupBox grp6 "Job Definition" pos:[195,237] width:444 height:120
    spinner NStartLine "Render: Line Number Start" pos:[350,303] width:196 height:16 range:[1,5000,1] type:#integer
    button BMD "Make Directories" pos:[5,13] width:173 height:34
    button BCAL "Calculate" pos:[195,12] width:104 height:34
    label PsiFVOutput "PsiFVOutput" pos:[385,158] width:82 height:18
    label DXOutput "DX (mm)" pos:[385,179] width:87 height:18
    label DYOutput "DY (mm)" pos:[386,198] width:86 height:18
    label W_Output "W Used (mm)" pos:[506,98] width:118 height:18
    label H_Output "H Used (mm)" pos:[506,119] width:124 height:18
    label CompleteColOutput "Completed Column" pos:[491,184] width:135 height:18
    GroupBox grp19 "Render Progress" pos:[487,154] width:144 height:63
    spinner Delta "Holog Size (microns)" pos:[279,162] width:70 height:16 range:[100,3000,800] type:#integer
}

G DP2.ms - MAXScript
File Edit Search Debugger Help
button BCam "Make Camera" pos:[330,13] width:104 height:34
on BRen pressed do
(
)
on BMD pressed do
(
)
on BCAL pressed do
(
)
on BCam pressed do
(
)

```

FIGURE 10.35 MAXScript code corresponding to the panel in Figure 10.34.

operator decide if he or she wants to re-input certain values before continuing. The “Make Camera” button then creates a camera and assigns the active view to this camera.

The last button is the “Start Render” button, which starts the rendering process. The job definition window allows a region of lines to be assigned to the script. This facility is useful because in most cases a farm of many computers will be calculating the whole data set. In this way, one can easily set the lines that a given computer is responsible for rendering.

10.5.3 Global Variables

For different parts of the MAXScript program to communicate with each other, we must define variables with a global scope. This is done using the “global” command. We will need the following variables:

```
global W,h1,H,DX,DY,NK,NG,TopDirName
```

This line may be inserted anywhere in the main program, but it is good practice to put it at the top of the code.

10.5.4 Writing the “Make Directories” Event Handler

The code for the “Make Directories” event handler should be inserted between the two braces immediately following the statement “on BMDpressed do” in the MAXScript text editor window. To insert the code, you must first save and close the visual MAXScript editor window. The first code we will insert is the following:

```
TopDirName= "C;/" +RenderDir.Text
makedir TopDirName
```

The first line creates a text variable TopDirName and sets it to the string “C:/”. It then takes the text data entered into the RenderDir Edit Box control on the panel at execution time and adds this to the string. The second line then uses the MAXScript control “makedir” to create a new directory with the name contained in TopDirName. If the directory already exists, nothing happens. The MAXScript commands are well documented—just click **Help** on the menu bar of the MAXScript text editor and then choose **Help** in the drop-down menu. This will take you directly to the MAXScript reference. Before being able to run the code, you will need to check at the very bottom of the file if the following code line is present:

```
"CreateDialog DoubleParallaxRender".
```

Sometimes, the visual editor forgets to add this or misses a letter or two at the end.

To run the code, click on the **File** menu at the top of the text editor and choose **Evaluate All** from the drop-down menu. The script will now run and the window we designed will appear. Type in the word “TestDir1” as the Render Directory and press **Make Directories**. Now go to C:/ and check that a new empty directory called TestDir1 has been created.

We will now create a subdirectory for each camera line rendered. In the panel we specified two variables—**NStartLine** and **NStopLine**—are the values of the index g for which we want to calculate render data. In each line directory, we will calculate N_k files. The following code uses a simple MAXScript “for” loop to accomplish the task of creating the subdirectories:

```
for i=NStartLine.value to NStopLine.value do
(
    LineText=i as string
    DirName=TopDirName+"/Line"+LineText
    makedir DirName
)
```

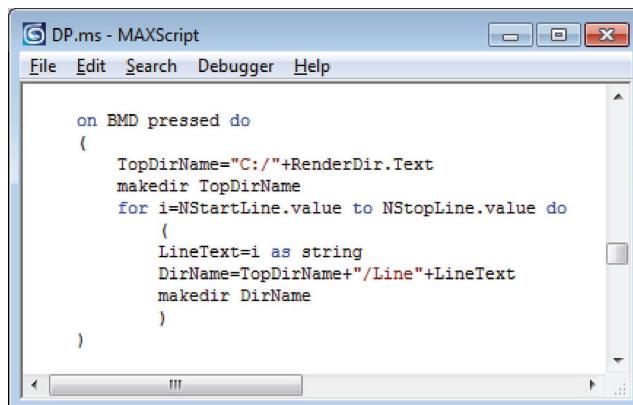


FIGURE 10.36 MAXScript code of the “Make Directories” event handler.

This code then creates subdirectories of the form “Line1”, “Line2”, “Line3”, etc. in the parent directory specified. The full code of the “Make Directories” event handler is shown in Figure 10.36.

10.5.5 Writing the “Calculate” Event Handler

We now come to the next event handler. This code simply calculates a number of useful parameters that the user should check before initiating the render process. The formulae on which the calculations are based are all listed in Chapter 9, Section 2.

```
h1=Delta.value/2.0/1000.0*(NM.value-1)/tan(PsiPH.value/2.0)
hOutput.caption="h (mm) =" +h1
```

This calculates the first quantised camera distance, h , through Equation 9.3. The values of N_K , N_G , W and H are likewise defined by Equations 9.5 and 9.6:

```
NK=2*floor(Win.value*1000.0/2.0/Delta.value) as integer
NG=2*floor(Hin.value*1000.0/2.0/Delta.value) as integer
NKText=NK as string
NGText=NG as string
NKOutput.caption="NK = " +NKText
NGOutput.caption="NG = " +NGText
W=(NK-1)*Delta.value/1000.0
H=(NG-1)*Delta.value/1000.0
WText=W as string
HText=H as string
W_Output.caption="W (mm) = " +WText
H_Output.caption="H (mm) = " +HText
```

Note that we have restricted N_K and N_G to be even as N_A and N_B have been restricted to even numbers. The vertical FOV is determined by the SLM aspect ratio and the given horizontal FOV:

$$\Psi_{cv} = 2 \tan^{-1} \left\{ \frac{(N_v - 1)}{(N_M - 1)} \tan \left(\frac{\Psi_{ch}}{2} \right) \right\} \quad (10.85)$$

```
PsiPV=2.0*atan(float(NV.value-1)/float(NM.value-1)*tan(PsiPH.value/2.0))
PsiPVText=PsiPV as string
PsiPVOutput.caption="PsiPV = " +PsiPVText
```

Finally, D_X and D_Y are defined by

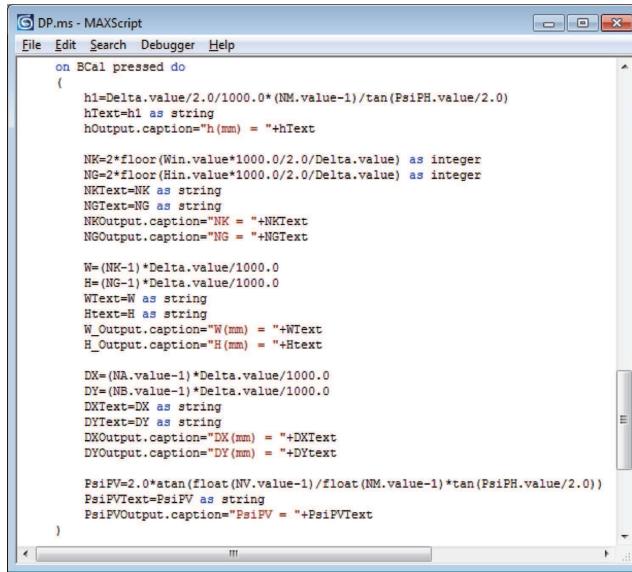
$$D_X = (N_A - 1)\delta \quad (10.86)$$

$$D_Y = (N_B - 1)\delta$$

where δ is the hogel diameter, giving:

```
DX=(NA.value-1)*Delta.value/1000.0
DY=(NB.value-1)*Delta.value/1000.0
DXText=DX as string
DYText=DY as string
DXOutput.caption="DX (mm) = " +DXText
DYOutput.caption="DY (mm) = " +DYtext
```

The full code of the “Calculate” event handler is shown in Figure 10.37.



```

DP.ms - MAXScript
File Edit Search Debugger Help
on BCal pressed do
(
    h1=Delta.value/2.0/1000.0*(NM.value-1)/tan(PsiPH.value/2.0)
    hText=h1 as string
    NKOutput.caption="h(mm) = "+hText

    NK=2*floor(Win.value*1000.0/2.0/Delta.value) as integer
    NG=2*floor(Hin.value*1000.0/2.0/Delta.value) as integer
    NKText=NK as string
    NGText=NG as string
    NKOutput.caption="NK = "+NKText
    NGOutput.caption="NG = "+NGText

    W=(NK-1)*Delta.value/1000.0
    H=(NG-1)*Delta.value/1000.0
    WItext=W as string
    Htext=H as string
    W_Output.caption="W(mm) = "+WItext
    H_Output.caption="H(mm) = "+Htext

    DX=(NA.value-1)*Delta.value/1000.0
    DY=(NB.value-1)*Delta.value/1000.0
    DXText=DX as string
    DYText=DY as string
    DXOutput.caption="DX(mm) = "+DXText
    DYOutput.caption="DY(mm) = "+DYText

    PsiPV=2.0*atan(float(NV.value-1)/float(NM.value-1)*tan(PsiPH.value/2.0))
    PsiPVItext=PsiPV as string
    PsiPVOutput.caption="PsiPV = "+PsiPVItext
)

```

FIGURE 10.37 MAXScript code of the “Calculate” event handler.

10.5.6 Writing the “Make Camera” Event Handler

This event handler creates the camera. Its horizontal FOV is defined by the formula

$$\Psi_{CH} = 2 \tan^{-1} \left\{ \frac{W + D_X}{2h} \right\} \quad (10.87)$$

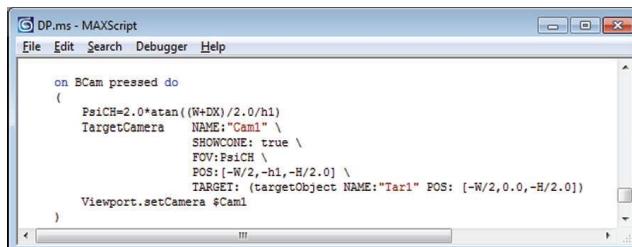
The camera can be created in MAXScript using the following commands:

```

PsiCH=2.0*atan((W+DX)/2.0/h1)
TargetCamera  NAME: "Cam1" \
              SHOWCONE: true \
              FOV:PsiCH \
              POS: [-W/2,-h1,-H/2.0] \
              TARGET: (targetObject NAME: "Tar1" POS: [-W/2,0.0,-H/2.0])

```

This creates a camera, Cam1, positioned at the starting location. A target point is also defined by the variable Tar1 and is set such that the camera faces straight ahead. Note the use of the backslash characters to spread the command out over several lines.



```

DP.ms - MAXScript
File Edit Search Debugger Help
on BCam pressed do
(
    PsiCH=2.0*atan((W+DX)/2.0/h1)
    TargetCamera  NAME:"Cam1" \
                  SHOWCONE: true \
                  FOV:PsiCH \
                  POS:[-W/2,-h1,-H/2.0] \
                  TARGET: (targetObject NAME:"Tar1" POS: [-W/2,0.0,-H/2.0])
    Viewport.setCamera $Cam1
)

```

FIGURE 10.38 MAXScript code of the “Make Camera” event handler.

Because we will be using render commands that only work in the active viewport, we will also need to set the active viewport to the camera that we have just created:

```
Viewport.SetCamera $Cam1
```

The full code of the “Make Camera” event handler is shown in Figure 10.38.

10.5.7 Writing the “Start Render” Event Handler

This is the last and most complicated event handler. It must perform a two-dimensional loop over the lines, g , and the columns, k , of the window using the created camera to render each frame. The line loop is limited by the variables NStartLine and NStopLine. The column loop will, however, be the full range from 1 to N_G . At each value of k and g , only a region of the camera frame must be rendered. This is accomplished using the MAXScript command “render”:

```
Render\
    OUTPUTFILE: FrameFileName \
    OUTPUTWIDTH: NI \
    OUTPUTHEIGHT: NJ \
    RENDERTYPE: #RegionCrop \
    REGION: #(N1x,N1y,N2x,N2y) \
    VFB: OFF
```

The size of the full camera frame is given by the parameters N_I and N_J , which are defined by

$$\begin{aligned} N_I &= 1 + \frac{D_x + W}{\delta} \\ N_J &= 1 + \frac{D_y + H}{\delta} \end{aligned} \tag{10.88}$$

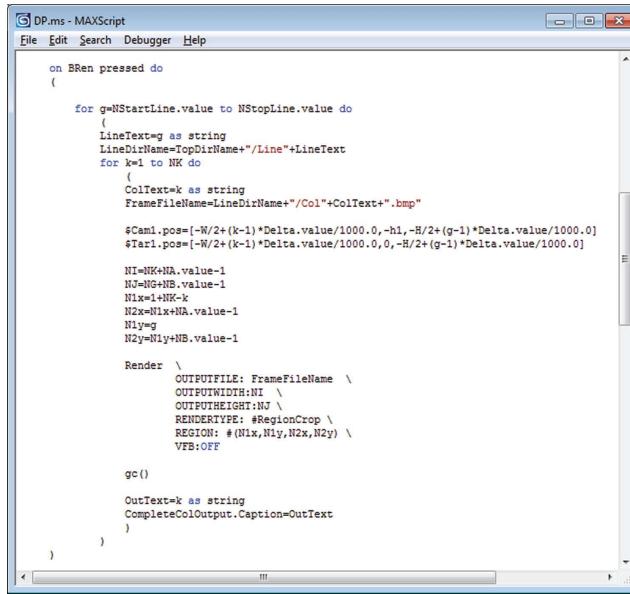
Together with Equation 10.87, this means that the full camera frame will always contain the hologram frame as a complete subset, irrespective of the position of the camera in the window. However, at each camera position, we will only need to render a small part of the full camera frame to create a “centred camera” data set. The subsection of the full camera frame that we must render is therefore a function of the camera indices k and g , and is given by the simple formula:

$$\begin{aligned} N_{1X} &= N_K - k + 1 \\ N_{2X} &= N_{1X} + N_A - 1 \\ N_{1Y} &= g \\ N_{2Y} &= N_{1Y} + N_B - 1 \end{aligned} \tag{10.89}$$

The camera must be moved before each render operation by the commands:

```
$Cam1.pos=[-W/2+(k-1)*Delta.value/1000.0,-h1,-H/2+(g-1)*Delta.value/1000.0]
$Tar1.pos=[-W/2+(k-1)*Delta.value/1000.0,0,-H/2+(g-1)*Delta.value/1000.0]
```

Figure 10.39 shows a full listing of the “Start Render” event handler. There are several things in particular that we should mention. The first is that the render command can fill up memory quickly and as a result you need to periodically use the command `gc()`. Second, the render command, when it uses `RenderCrop`, contains a bug in that the last column and last row of pixels are incorrectly set to zero,



```

DP.ms - MAXScript
File Edit Search Debugger Help
on BRen pressed do
{
    for gNStartLine.value to NStopLine.value do
    {
        LineDirName=TopDirName+"/Line"+LineText
        for k=1 to NK do
        {
            ColText=k as string
            FrameFileName=LineDirName+"/Col"+ColText+".bmp"
            $Cami.pose=[-W/2+(k-1)*Delta.value/1000.0,-h1,-H/2+(g-1)*Delta.value/1000.0]
            $Tari.pose=[-W/2+(k-1)*Delta.value/1000.0,0,-H/2+(g-1)*Delta.value/1000.0]

            N1=NK+NA.value-1
            N2=NG+NB.value-1
            N1x=1+NK-k
            N2x=N1x+NA.value-1
            N1y=y
            N2y=N1y+NB.value-1

            Render \
                OUTPUTFILE: FrameFileName \
                OUTPUTWIDTH:N1 \
                OUTPUTHEIGHT:N2 \
                RENDERTYPE: #RegionCrop \
                REGION: #(N1x,N1y,N2x,N2y) \
                VFB:OFF

            gc()

            OutText=k as string
            CompleteColOutput.Caption=OutText
        }
    }
}

```

FIGURE 10.39 MAXScript code of the “Start Render” event handler.

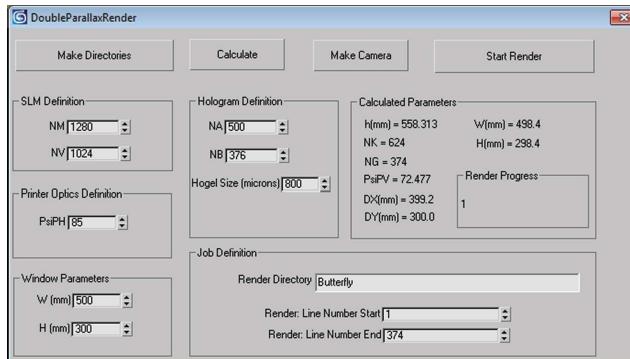


FIGURE 10.40 Final MAXScript panel filled out to generate the complete image set required for a typical 30×40 cm DWDH reflection hologram.

leading to one-pixel-thick black lines at the extreme edges of the camera frame.* For most applications, this has a minimal effect and practically one just specifies N_A and N_B to be two hogels larger than necessary.

In the code presented here, we have opted to write out the render data in the form of bitmap files. It is worth mentioning that Max supports the jpeg format. To write jpgs instead of bitmaps, you just need to change the code line

```
FrameFileName=LineDirName+"/Col"+ColText+".bmp"
```

to

```
FrameFileName=LineDirName+"/Col"+ColText+".jpg"
```

By selecting a reasonable compression setting, you can save a lot of space without much effect on the final hologram. You can change various options including compression in the Max render dialogue.

The script has been designed for system units of millimetres. Be sure to check that you have chosen this as your default preference in Max! Figure 10.40 shows a screenshot of the panel for the case of a

* This error is also present in all later versions of Max up until the time of writing.

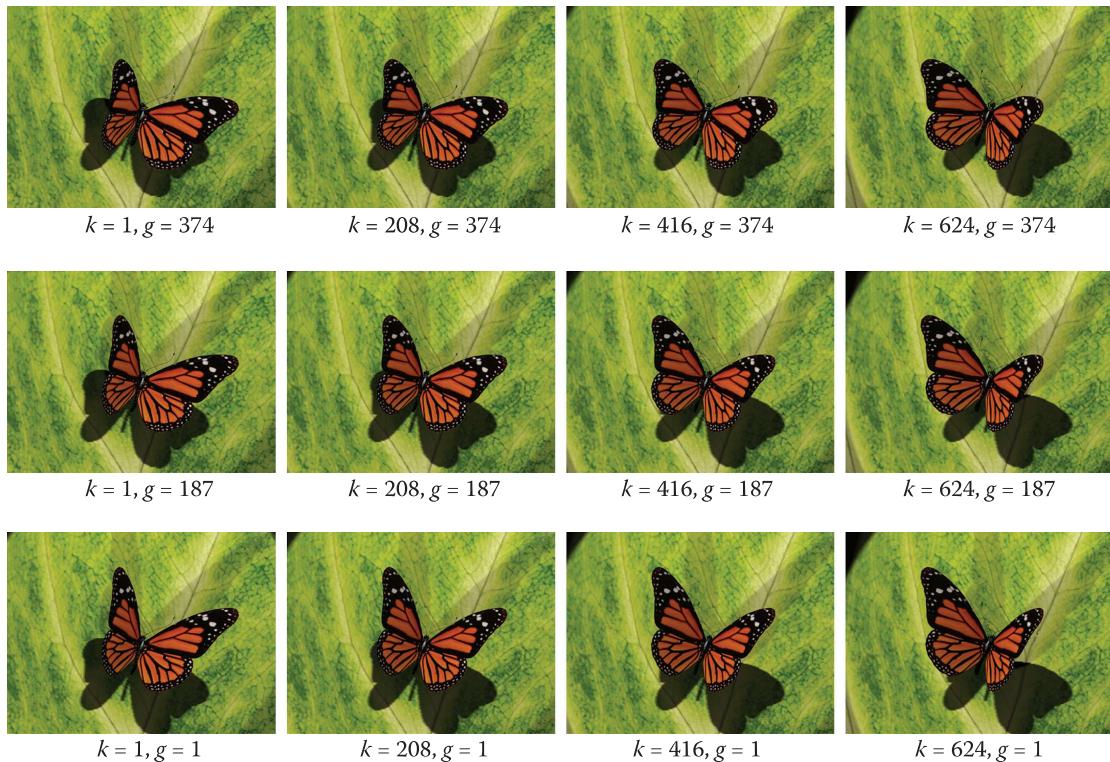


FIGURE 10.41 Selected camera frames generated by the MAXScript program with the settings of the panel shown in Figure 10.40. The total data set for the $30 \text{ cm} \times 40 \text{ cm}$ hologram comprises 624×374 frames requiring 125 GB of disk space using the .bmp format.

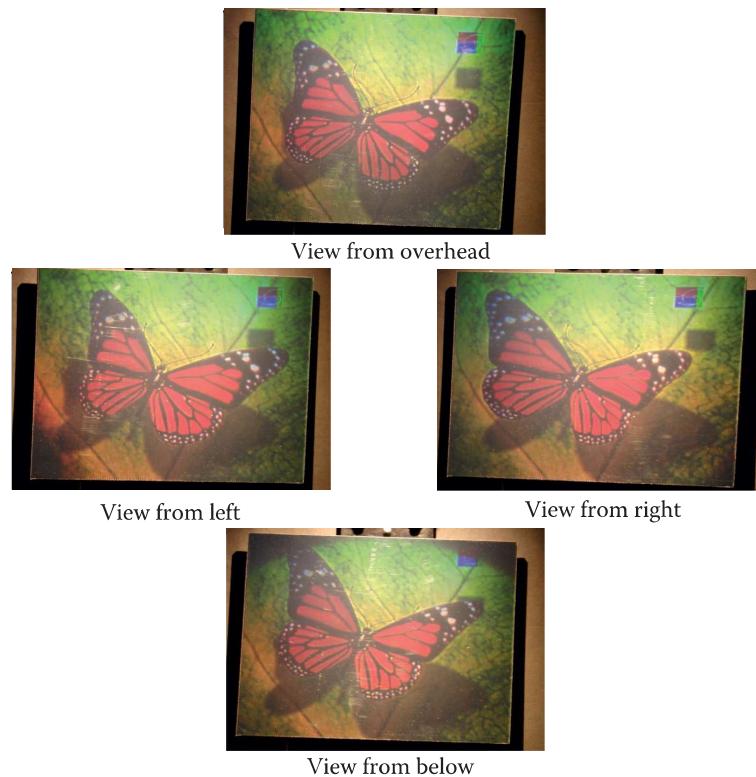


FIGURE 10.42 Several pictures of the final DWDH reflection hologram of Figures 10.40 and 10.41 printed on PFG3CN silver halide emulsion and illuminated by a 50 W MR-16 halogen lamp. (Photograph courtesy of Geola Digital UAB.)

typical 30×40 cm hologram. Figure 10.41 shows selected camera frames generated by the script, and Figure 10.42 shows several pictures of an actual DWDH reflection hologram printed using this data. A full listing of the complete MAXScript program is given in Appendix 5. The data set for the hologram comprises 624×374 individual frames requiring a total of 125 GB of disk space. By using jpegs with a compression setting of 75% instead of bitmaps, this can be reduced to just under 6 GB. Rendering with either format on a standard 2011 laptop takes approximately 38 h.

REFERENCE

1. D. Brotherton-Ratcliffe, A. Nikolskij, S. Zacharovas, J. Pileckas and R. Bakanas, *Image capture system for a digital holographic printer*, US Patent 8,154,584 (filed 2006, granted 2012).

11

Theoretical Basis for High-Fidelity Display Holograms

11.1 Introduction

This book is primarily about the creation of ultra-realistic images through holography. In recent years, the quality of images has progressed rapidly in the field, and we are optimistic that this process will continue. But why should we be confident of this? This chapter is partly a concise answer to this question. However, it also fulfils another vital role. It introduces, in some detail, the basic physical models that mathematically describe how a high-fidelity hologram actually functions. A clear understanding of these models is invaluable for any worker in the field. Besides providing formulae for vital processes, such as diffractive efficiency, image blurring, and aberration correction, a comprehensive understanding of the underlying physics allows the worker to develop a clear intuition of what is likely to work best.

We shall start our discussion of display holograms with the thin transmission hologram. This is the simplest place to start and the most productive. A simple first-order mathematical analysis shows why any hologram, thick or thin—reflective or transmissive—is subject to source size and chromatic blurring. Formulae for these effects and for dispersion in the transmission hologram are derived. This analysis forms the foundation for rainbow and achromatic holography.

In Section 11.4, we extend the discussion to volume phase holograms. Here, we derive the Bragg condition from a perturbation analysis of the phase function using spherical waves. We also derive both paraxial and fully non-paraxial equations describing the general reconstruction of images in both transmission and reflection geometries when different wavelengths and angles are used at recording and playback. The non-paraxial equations are of fundamental use in predistortion and in distortion correction algorithms employed in digital display holography. In Section 11.6, we describe the thick volume phase hologram in terms of a set of parallel-stacked mirrors (PSM). We develop the same non-paraxial equations as in Section 11.4 using this method, demonstrating the equivalence of the two pictures. We use a very basic discussion of the PSM model* to introduce the important differences between transmission and reflection holograms, and we estimate diffractive efficiencies as well as angular and wavelength sensitivities for both types.

In Section 11.10, we describe a more rigorous treatment of the problem of diffraction from volume phase holograms. This is Kogelnik's one-dimensional coupled wave theory [1]. We present the main results of this theory using the simpler σ -polarisation and a two-dimensional grating, although we later extend this somewhat briefly to the π -polarisation. This leads us to a discussion of the physical characteristics of suitable materials for ultra-realistic imaging. We also present a discussion of the N -coupled wave theory [2–4] because this analytical extension of Kogelnik's theory nicely demonstrates how holograms of diffuse images can be expected to attain almost perfect diffractive efficiency in the limit of a large amplitude reference recording beam. It also provides some intuition that multicolour volume phase holograms may be designed, at least in principle, to have high diffractive efficiencies—and indeed this will be confirmed in Chapter 12 when we study polychromatic gratings and holograms more rigorously using the PSM model.

In Section 11.11, we return to the intrinsic image blurring in analogue and digital holograms in the light of what we have learnt. We end the chapter with a brief discussion in Section 11.12 of computational methods for calculating the diffraction efficiency from planar gratings.

* We return to the parallel-stacked mirror model in Chapter 12, in which we present a detailed discussion of this theory.

11.2 Three-Dimensional Paraxial Theory of the Thin Transmission Hologram

We consider two infinitely thin holograms* each produced by the interference of a pair of spherical waves, one originating from an object point and the other originating from a reference point (Figure 11.1). In the first case, we take the object point to have coordinates (x_o, y_o, z_o) and the reference source to be located at (x_r, y_r, z_r) . In the second case, the object point is located at (x_i, y_i, z_i) and the reference source is at (x_c, y_c, z_c) . The hologram is, in both cases, located on the $x = 0$ plane.

Now, our two holograms will be essentially identical if the surface distribution of the time-averaged square of the electric field amplitude is the same for each. With this in mind, we will examine a zone near the coordinate origin without loss of generality. The time-averaged squared electric field distribution at a point on the first hologram $(0, y_h, z_h)$ is then given by

$$\langle \mathbf{E}^2(y_h, z_h) \rangle = \frac{1}{T} \int \left\{ \Re \left(\mathbf{E}_r \frac{e^{i[\omega t - \frac{2\pi}{\lambda_r} \rho_r(y_h, z_h)]}}{\rho_r(y_h, z_h)} + \mathbf{E}_o \frac{e^{i[\omega t - \frac{2\pi}{\lambda_r} \rho_o(y_h, z_h)]}}{\rho_o(y_h, z_h)} \right) \right\}^2 dt \quad (11.1)$$

Likewise, the distribution on the second hologram may be written as

$$\langle \mathbf{E}^2(y_h, z_h) \rangle = \frac{1}{T} \int \left\{ \Re \left(\mathbf{E}_c \frac{e^{i[\omega t - \frac{2\pi}{\lambda_c} \rho_c(y_h, z_h)]}}{\rho_c(y_h, z_h)} + \mathbf{E}_i \frac{e^{i[\omega t - \frac{2\pi}{\lambda_c} \rho_i(y_h, z_h)]}}{\rho_i(y_h, z_h)} \right) \right\}^2 dt \quad (11.2)$$

In both cases

$$\rho_\gamma(y_h, z_h) = \sqrt{x_\gamma^2 + (y_\gamma - y_h)^2 + (z_\gamma - z_h)^2} \quad \forall \gamma \in \{r, o, c, i\} \quad (11.3)$$

The E variables control the relative amplitude of each wave. Now, Equation (11.1) can be rewritten assuming $T \rightarrow \infty$ as

$$\begin{aligned} \langle \mathbf{E}^2(y_h, z_h) \rangle &= \frac{1}{2} \left\{ \frac{|E_r|^2}{\rho_r^2(y_h, z_h)} + \frac{|E_o|^2}{\rho_o^2(y_h, z_h)} \right\} \\ &\quad + \frac{\kappa_{ro}(y_h, z_h) \sqrt{E_r E_o E_r^* E_o^*}}{\rho_o(y_h, z_h) \rho_r(y_h, z_h)} \cos \left[\frac{2\pi}{\lambda_r} \rho_o(y_h, z_h) - \frac{2\pi}{\lambda_r} \rho_r(y_h, z_h) + \zeta_{ro}(E_r, E_o) \right] \end{aligned} \quad (11.4)$$

where $\kappa_{ro}(x_h, z_h)$ is a function related to the (approximately linear) polarisation of the sources. Both the first term and the multiplier of the second term are slowly varying functions of y_h and z_h . The trigonometric term, on the other hand, varies extremely rapidly with y_h and z_h as its argument is very large. As a consequence, our two holograms can be considered identical if they share an identical trigonometric argument or if

$$\frac{2\pi}{\lambda_r} \{\rho_r - \rho_o\} = \pm \frac{2\pi}{\lambda_c} \{\rho_c - \rho_i\} \pm \{\zeta_{ro} - \zeta_{ci}\} \pm 2n\pi \quad (11.5)$$

Here, the term $\zeta_{ro} - \zeta_{ci}$ is a constant phase term that depends only on the magnitudes of the reference and object waves. The term $2n\pi$, where n is any integer, simply reflects the periodic nature of the trigonometric function. Figure 11.2 shows several illustrations of typical fringe patterns calculated from Equation 11.4.

* In this chapter we will use the words “grating” and “hologram” interchangeably. In Chapter 12 we will make a distinction between these two terms when we specifically treat the hologram as the large N limit of a spatially multiplexed grating.

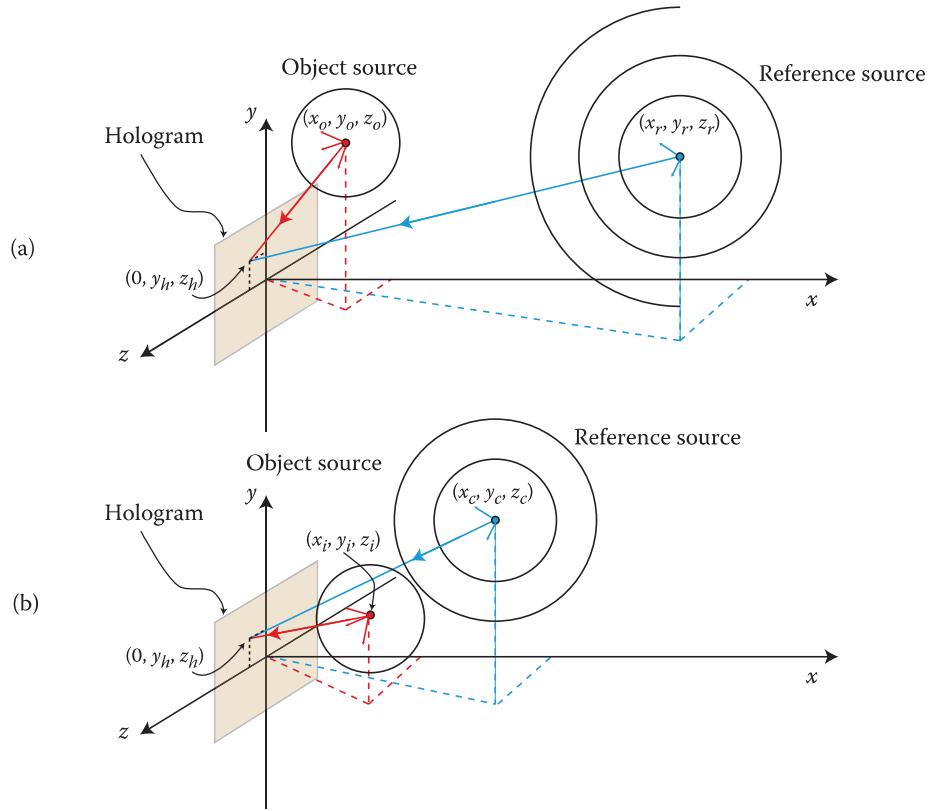


FIGURE 11.1 This diagram may be interpreted in two ways. In the first way, (a) and (b) show two different recording geometries for a thin transmission hologram. In each case, there is a single reference spherical wave and a single object spherical wave. The recording wavelength and position of the object and reference sources are different in each case. If we now assume that the fringe pattern within each transmission hologram produced by recording methods (a) and (b) are identical (i.e., if the fringe function $\langle E^2(y_h, z_h) \rangle$ is the same for both scenarios), then we can interpret the two diagrams in a different fashion: (a) now represents the recording of a hologram; (b) now represents how it can be expected to replay at another wavelength and using another position for the reference source. Under this interpretation and as drawn, (b) represents the reconstruction of a virtual image situated at (x_i, y_i, z_i) when the hologram recorded in (a) is illuminated by a spherical source whose origin is (x_c, y_c, z_c) . This dual interpretation requires the use of the fundamental holography theorem: that an ideal hologram reconstructed with exactly the same wavelength and geometry as used in recording will faithfully reproduce the original image wave.

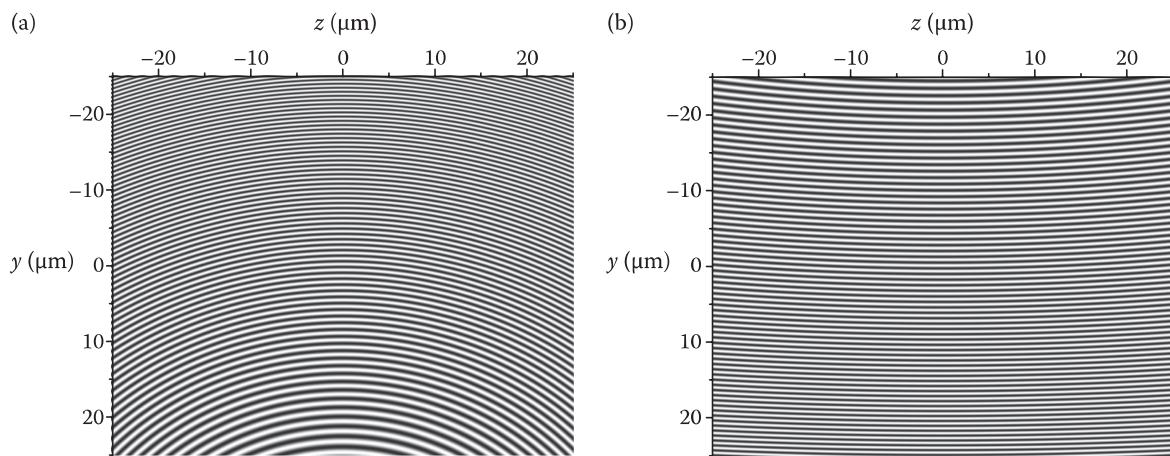


FIGURE 11.2 Some example pictures of $\langle E^2 \rangle$ for transmission gratings as calculated from Equation 11.4. Recording wavelength is 532 nm. (a) Object sink located at $x_o = -100 \mu\text{m}$, $y_o = 0 \mu\text{m}$, $z_o = 0 \mu\text{m}$; reference source located at $x_r = 100 \mu\text{m}$, $y_r = 100 \mu\text{m}$, $z_r = 0 \mu\text{m}$. Object source located at $x_o = 100 \mu\text{m}$, $y_o = 0 \mu\text{m}$, $z_o = 0 \mu\text{m}$; (b) reference source located at $x_r = 100 \mu\text{m}$, $y_r = 100 \mu\text{m}$, $z_r = 0 \mu\text{m}$. Note the really quite small curvature of the fringes. For larger object and reference distances, the fringes will appear essentially straight on the scale of the recording wavelength.

We can use Equation 11.5 to develop a useful theory of the thin transmission hologram by expanding the general function ρ_γ in terms of a two-dimensional power series in the small quantities y_h and z_h :

$$\rho_\gamma = \sum_{m=0}^M \sum_{n=0}^N \rho_{pmn} z_h^m y_h^n \quad \forall \gamma \in \{r, o, c, i\} \quad (11.6)$$

Using Equation 11.3, we see that the lowest-order terms can be written as follows:

$$\begin{aligned} \rho_{\gamma 00} &= \sqrt{x_\gamma^2 + y_\gamma^2 + z_\gamma^2}; & \rho_{\gamma 10} &= -\frac{z_\gamma}{\sqrt{x_\gamma^2 + y_\gamma^2 + z_\gamma^2}} \\ \rho_{\gamma 01} &= -\frac{y_\gamma}{\sqrt{x_\gamma^2 + y_\gamma^2 + z_\gamma^2}}; & \rho_{\gamma 11} &= -\frac{y_\gamma z_\gamma}{(x_\gamma^2 + y_\gamma^2 + z_\gamma^2)^{3/2}} \end{aligned} \quad (11.7)$$

Substitution of these expressions in Equation 11.5 leads to a set of non-linear equations for (x_i, y_i, z_i) in terms of (x_o, y_o, z_o) . By imposing a paraxial condition on both the “*i*” and “*o*” source points, however, we can linearise these equations. In practice, we do this by expanding the coordinate pairs (y_i, z_i) and (y_o, z_o) as Taylor series and then truncating Equation 11.5 at quadratic orders in these four variables. The expansion orders $(m = 1, n = 0)$, $(m = 0, n = 1)$ and $(m = 1, n = 1)$ then lead to the following paraxial equations:

$$\frac{1}{x_i} = \frac{y_c z_c}{(x_c^2 + y_c^2 + z_c^2)^{3/2}} \mp \frac{\lambda_c}{\lambda_r} \left\{ \frac{y_r z_r}{(x_r^2 + y_r^2 + z_r^2)^{3/2}} - \frac{1}{x_o} \right\} \quad (11.8)$$

$$\frac{y_i}{x_i} = \frac{y_c}{\sqrt{x_c^2 + y_c^2 + z_c^2}} \mp \frac{\lambda_c}{\lambda_r} \left\{ \frac{y_r}{\sqrt{x_r^2 + y_r^2 + z_r^2}} - \frac{y_o}{x_o} \right\} \quad (11.9)$$

$$\frac{z_i}{x_i} = \frac{z_c}{\sqrt{x_c^2 + y_c^2 + z_c^2}} \mp \frac{\lambda_c}{\lambda_r} \left\{ \frac{z_r}{\sqrt{x_r^2 + y_r^2 + z_r^2}} - \frac{z_o}{x_o} \right\} \quad (11.10)$$

where x_o , x_r and x_c are assumed to have positive values.* Note that the $(m = 0, n = 0)$ order defines the term $\{\zeta_{ro} - \zeta_{ci}\} + 2n\pi$ in Equation 11.5, and all higher orders simply restate the information contained in the first four orders. To recap, these expressions relate two identical thin transmission holograms. The first is recorded at a wavelength of λ_r using an object wave originating at (x_o, y_o, z_o) and a reference wave at (x_r, y_r, z_r) . The second is recorded at a wavelength of λ_c using an object wave originating at (x_i, y_i, z_i) and a reference wave at (x_c, y_c, z_c) . The holograms are identical because Equation 11.5 means that they both share identical interference patterns. Now, this is useful because a thin hologram recorded at a given wavelength with a given source and reference is known to replay an image, constituting a faithful reproduction, of the object source if illuminated by the same reference source at the same wavelength. We can demonstrate this easily enough for the case of an amplitude hologram (the case of the phase hologram is only a little more tedious [e.g., 5]).

Suppose we consider an arbitrary object wave of time dependence $e^{i\omega t}$

$$\mathbf{E}_o = \Re[\mathbf{A}(x, y, z)e^{-i\Phi(x, y, z)}] \quad (11.11)$$

and an arbitrary reference wave

$$\mathbf{E}_r = \Re[\mathbf{B}(x, y, z)e^{-i\psi(x, y, z)}] \quad (11.12)$$

* In fact, if we take x_o to be negative, then this represents either an object sink or alternatively an object source but recorded in “reflection” mode. The thin transmission hologram makes no distinction between these two cases.

Then, on the surface of the photographic plate, we can expect the following average squared electric field distribution

$$\langle \mathbf{E}^2 \rangle \sim \frac{1}{2} \left(|\mathbf{A}|^2 + |\mathbf{B}|^2 \right) + \frac{1}{2} \mathbf{A}^* \cdot \mathbf{B} e^{i(\phi-\psi)} + \frac{1}{2} \mathbf{A} \cdot \mathbf{B}^* e^{-i(\phi-\psi)} \quad (11.13)$$

If we assume that after development, the transmission of the plate is proportional to $\langle \mathbf{E}^2 \rangle$, then the complex light amplitude produced at the plate surface on illumination by the same reference wave will be

$$\mathbf{E}_i \propto \mathbf{B} e^{-i\psi} \langle \mathbf{E}^2 \rangle \sim \left\{ \frac{|\mathbf{A}|^2 + |\mathbf{B}|^2}{2} \right\} \mathbf{B} e^{-i\psi} + \left\{ \frac{1}{2} \mathbf{A}^* \cdot \mathbf{B} \right\} \mathbf{B} e^{i(\phi-2\psi)} + \left\{ \frac{1}{2} \mathbf{A} \cdot \mathbf{B}^* \right\} \mathbf{B} e^{-i\phi} \quad (11.14)$$

If $|\mathbf{B}|$ is roughly constant over the plate and the polarisations of the waves do not vary, then the third term faithfully reproduces the original object wave. By choosing an off-axis geometry, the first and second terms may be angularly displaced from this third wave.

Hence, we are naturally led to an alternative interpretation of Equations 11.8 through 11.10. They can be regarded as specifying how a first hologram recorded at a wavelength λ_r using an object wave originating at (x_o, y_o, z_o) and a reference wave at (x_r, y_r, z_r) reproduce an image at (x_i, y_i, z_i) when illuminated by a wave originating at the reference point (x_c, y_c, z_c) at wavelength λ_c .

The presence of the second term in Equation 11.14 is related in the \mp sign appearing in Equations 11.8 through 11.10. This sign is present in these equations because the interference pattern of Equation 11.4 is invariant under the transformation

$$\rho_o(y_h, z_h) - \rho_r(y_h, z_h) + \frac{\lambda_r}{2\pi} \zeta_{ro}(E_r, E_o) \rightarrow -\rho_o(y_h, z_h) + \rho_r(y_h, z_h) - \frac{\lambda_r}{2\pi} \zeta_{ro}(E_r, E_o) \quad (11.15)$$

This leads, in general, to two reconstruction images being present in a thin transmission hologram—the *virtual image* and the *real image*. Using the notation we have developed and with x_o , x_r and x_c positive and relating to sources rather than sinks, the negative sign in Equations 11.8 through 11.10 represents the virtual image (x_i positive) whereas a positive sign indicates the real image (x_i negative). In all cases, the negative sign represents the primary non-conjugate image and the positive sign the conjugate image.

The simultaneous reconstruction of both a virtual image and a real image is a property of the thin transmission hologram. When the reference beam is on-axis or nearly on-axis, both images will be present in the same line of view—however, as the reference beam angle is increased, the conjugate or parasitic image is quickly shifted to high angles where it either fails to reconstruct or suffers large aberration. The volume nature of most common transmission holograms, as we shall see shortly, also conspires to eliminate the conjugate image.

11.2.1 Collimated Reference and Object Beams

In many cases of interest, display holograms are written with essentially collimated reference and replay beams. The paraxial theory is enormously simplified under this approximation and it is therefore instructive to start any discussion of the thin transmission hologram here.

We will start by writing the two reference waves in Equations 11.8 through 11.10 in spherical polar coordinates (Figure 11.3):

$$\frac{1}{x_i} = \frac{\sin^2 \theta_c \cos \phi_c \sin \phi_c}{r_c} \mp \frac{\lambda_c}{\lambda_r} \left\{ \frac{\sin^2 \theta_r \cos \phi_r \sin \phi_r}{r_r} - \frac{1}{x_o} \right\} \quad (11.16)$$

$$\frac{y_i}{x_i} = \sin \theta_c \cos \phi_c \mp \frac{\lambda_c}{\lambda_r} \left\{ \sin \theta_r \cos \phi_r - \frac{y_o}{x_o} \right\} \quad (11.17)$$

$$\frac{z_i}{x_i} = \sin \theta_c \sin \phi_c \mp \frac{\lambda_c}{\lambda_r} \left\{ \sin \theta_r \sin \phi_r - \frac{z_o}{x_o} \right\} \quad (11.18)$$

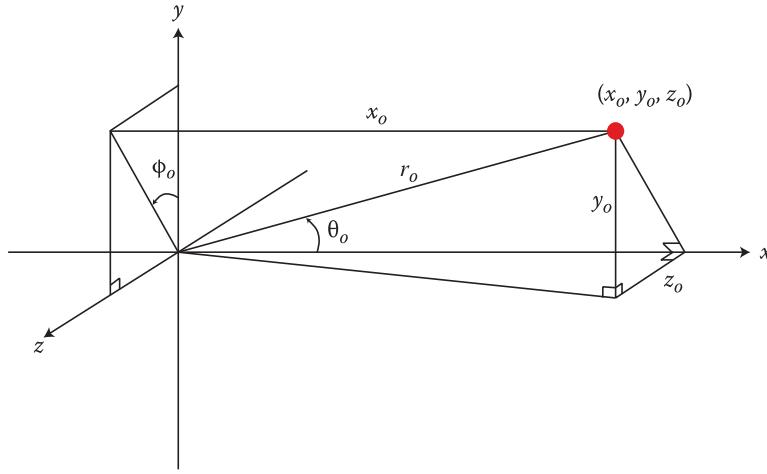


FIGURE 11.3 Recording geometry in spherical polar coordinates. The diagram shows how an object point (x_o, y_o, z_o) is defined by the equivalent right-handed spherical system (r_o, θ_o, ϕ_o) . The hologram is located on the (y, z) plane at $x = 0$.

Here, we have used the relations

$$\begin{aligned} x_r &= r_r \cos \theta_r; & y_r &= r_r \sin \theta_r \cos \phi_r; & z_r &= r_r \sin \theta_r \sin \phi_r \\ x_c &= r_c \cos \theta_c; & y_c &= r_c \sin \theta_c \cos \phi_c; & z_c &= r_c \sin \theta_c \sin \phi_c \end{aligned} \quad (11.19)$$

with $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$ and r positive.

In the case that $r_c = r_r \rightarrow \infty$ and $\phi_c = \phi_r = 0$ we can characterise the reference and replay waves by plane waves whose wave vectors have purely altitudinal angles of, respectively, θ_r and θ_c . Equations 11.16 through 11.18 then take the simple form*

$$x_i = x_o \frac{\lambda_r}{\lambda_c}; \quad y_i = x_o \left(\frac{\lambda_r}{\lambda_c} \sin \theta_c - \sin \theta_r \right) + y_o; \quad z_i = z_o \quad (11.20)$$

This tells us that if we record a single point (x_o, y_o, z_o) and then replay the hologram in white light (Figure 11.4), the virtual image of this point will be dispersed into a line defined by the equation

$$y_i = x_i \sin \theta_c - x_o \sin \theta_r + y_o \quad (11.21)$$

This line is orthogonal to the z direction; its gradient in the x, y plane is usually expressed as the tangent of the achromatic angle

$$\tan \Theta_A \equiv \sin \theta_c \quad (11.22)$$

Equations 11.20 through 11.22 form the basis of rainbow and achromatic holography as discovered by Steven Benton in 1969 [6]. They show that dispersion only occurs along the achromatic plane and that the horizontal (z) plane is dispersion-free. This of course means that one can construct transmission holograms whose parallax is in the horizontal dimension (horizontal parallax-only [HPO] holograms) and where the vertical dimension is dispersed. These results were used in Chapter 8, in which they were essential in letting us define the digital image data transformations required for the printing of direct-write digital holography (DWDH) rainbow and achromatic holograms.

We should note that if the hologram replay beam is not collimated—that is, it is a spherical wave with finite r_c —then Equations 11.16 through 11.18 show us that there is now dispersion in all dimensions; as such, rainbow and achromatic holograms can be expected to suffer from blurring due to the coupling of the horizontal parallax with horizontal dispersion if the illumination source is placed too close to the

* Here, we have used the minus sign in Equations 11.16 to 11.18 to treat the primary image.

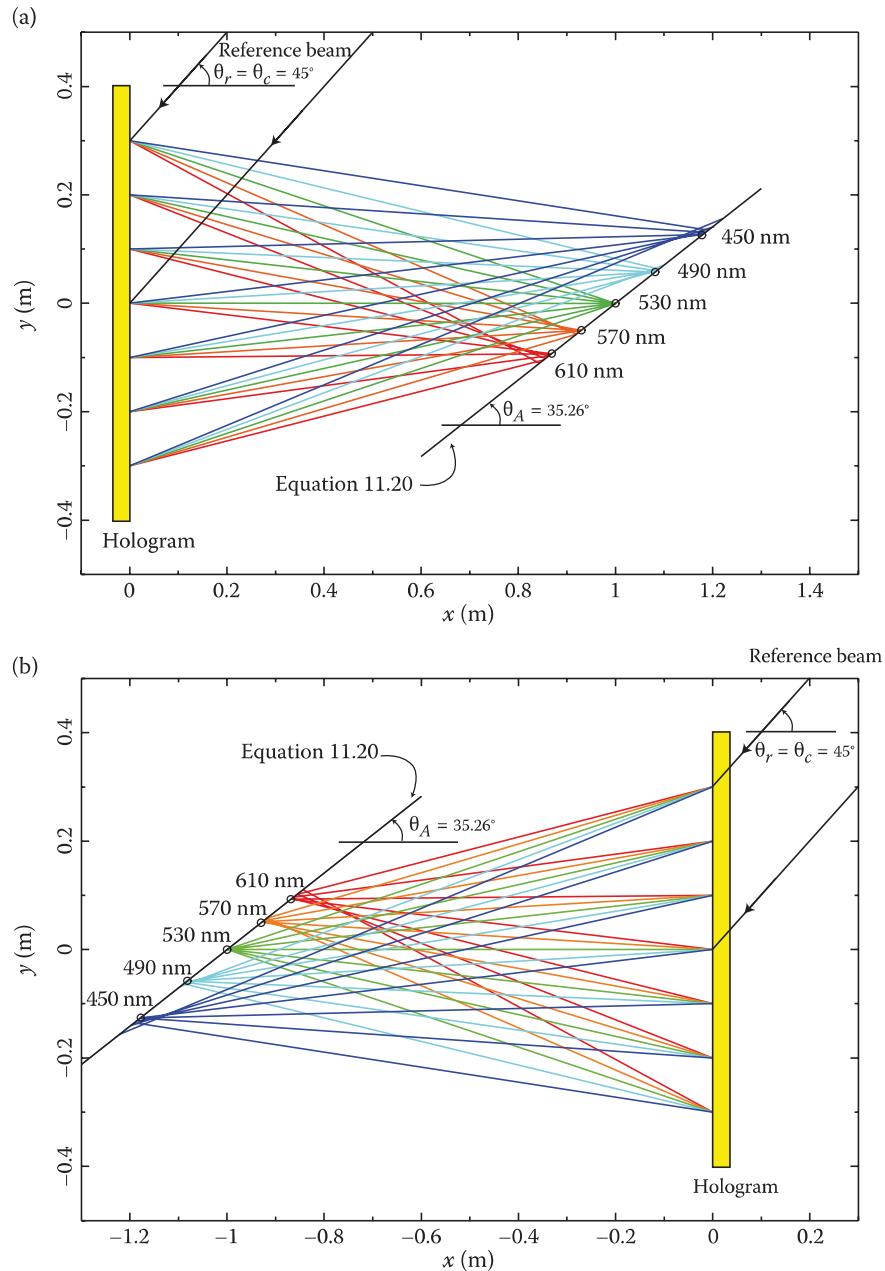


FIGURE 11.4 Reconstruction of (a) a virtual and (b) real image of a finite-aperture thin transmission hologram using identical geometry as at recording but at 5 different wavelengths. The reference beam is assumed to be collimated ($r_r = r_c \rightarrow \infty$) and axial ($\phi_r = \phi_c = 0$) and has an angle of incidence of $\theta_r = \theta_c = 45^\circ$. In (a) a single object source at $(x_o, y_o, z_o) = (1.0, 0.0, 1.1 \text{ m})$ is recorded at 530nm and then replayed at (450 nm, 490 nm, 530 nm, 570 nm, 610 nm). In (b) a corresponding object sink at $(x_o, y_o, z_o) = (1.0, 0.0, 1.1 \text{ m})$ is similarly recorded and replayed. The rays represent fully non-paraxial calculations. Note that away from the centre of the hologram the rays do not intersect at a defined point as they are expected to do under the paraxial theory. The solid line represents the paraxial calculation of equation 11.21 and the circles are the paraxial points as calculated from 11.20. Note that the reconstructed conjugate images (not shown) in the replay scenarios of both (a) and (b) which correspond to the choice of the plus sign in Equations 11.16 through 11.18 (rather than the minus sign which has been assumed in Equations 11.20 through 11.22) lead to highly shifted images with $y < 1 \text{ m}$.

hologram. Note, however, that if $\phi_c = 0$ and $r_c \rightarrow \infty$ but r_r and ϕ_r are finite, then the transverse z direction can once again be made dispersion-free. In this case, the image data for a rainbow or achromatic hologram will certainly need to be predistorted, but the important point is that the “mechanics” of these holograms depends more on the illumination rather than the recording geometry.

Equations 11.20 through 11.22 are equally valid for the case of positive x_o —in which case the image replays as a virtual image—and for negative x_o —in which case the image replays as a real image (see Figure 11.4).

11.2.2 Source-Size Blurring

We can use the paraxial theory to examine how a thin transmission hologram will blur due to an illumination source of finite size. Let us suppose that the illumination source has a finite extent which we define by the angle $\delta\theta_c$. We will make the approximation of an axial distant source on replay and recording:

$$\phi_c = \phi_r = 0; \quad r_c, r_r \rightarrow \infty \quad (11.23)$$

If the illumination source is circularly symmetric we may define the corresponding Cartesian uncertainties which describe the finite nature of the source as

$$\begin{aligned} \delta x &\equiv r\delta\theta_c \sin \langle \theta_c \rangle \\ \delta y &\equiv r\delta\theta_c \sin \langle \theta_c \rangle \\ \delta z &\equiv r\delta\theta_c \end{aligned} \quad (11.24)$$

where the operator $\langle \rangle$ indicates an average value for all rays. We may now use the Jacobian $\partial(x_i, y_i, z_i)/\partial(x_c, y_c, z_c)$, obtained by differentiating Equations (11.8) through (11.10), to map these uncertainties onto the image coordinates.

$$\begin{aligned} \delta x_i &= \frac{\partial x_i}{\partial x_c} \delta x_c + \frac{\partial x_i}{\partial y_c} \delta y_c + \frac{\partial x_i}{\partial z_c} \delta z_c = 0 \\ \delta y_i &= \frac{\partial y_i}{\partial x_c} \delta x_c + \frac{\partial y_i}{\partial y_c} \delta y_c + \frac{\partial y_i}{\partial z_c} \delta z_c = \left\{ \frac{\lambda_r}{\lambda_c} x_o \cos \langle \theta_c \rangle \right\} \delta\theta_c \\ \delta z_i &= \frac{\partial z_i}{\partial x_c} \delta x_c + \frac{\partial z_i}{\partial y_c} \delta y_c + \frac{\partial z_i}{\partial z_c} \delta z_c = \left\{ \frac{\lambda_r}{\lambda_c} x_o \right\} \delta\theta_c \end{aligned} \quad (11.25)$$

When the recording and replay wavelengths are the same this then leads to an image point uncertainty due to the finite source of

$$\begin{aligned} \delta x_i &= 0 \\ \delta y_i &= x_o \cos \langle \theta_c \rangle \delta\theta_c \\ \delta z_i &= x_o \delta\theta_c \end{aligned} \quad (11.26)$$

These Equations provide us with a simple estimation of the intrinsic blur in the transverse y and z directions caused by a finite illumination source size. Note that there is no longitudinal blur as x_i is independent of x_c , y_c , and z_c for distant axial sources (Equation 11.23).

As somewhat of an aside, it can be instructive to verify these equations using a computer. For simplicity, we assume the case of an amplitude transmission hologram in which the transmissivity, $\tau(y, z)$ is proportional to $\langle E^2 \rangle$. We can then easily calculate $\tau(y, z)$ for a given object point and a given reference point according to Equation 11.4. Finally, we numerically reconstruct the hologram with a spherical wave using a Kirchoff diffraction integral over the hologram surface:

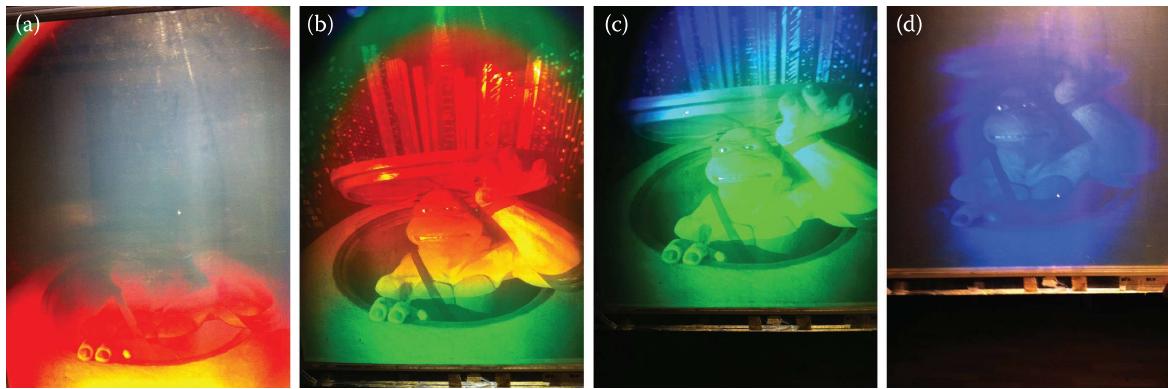


FIGURE 11.5 A single-slit analogue rainbow transmission hologram (60×80 cm) viewed from the front at various heights from (a) to (d) showing how the hologram changes colour due to the dispersion of Figure 11.4b. Note that each photograph shows more than a single colour for two reasons—the reference light is not precisely collimated and the photographs have not been taken precisely at the slit focus.

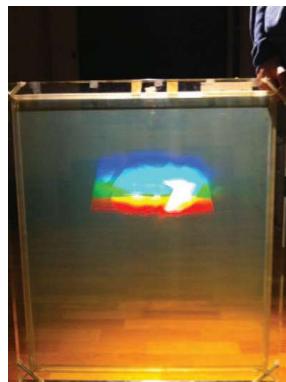


FIGURE 11.6 Same hologram as in Figure 11.5 but now illuminated from the other side, showing a dispersed virtual image of the master slit in accordance with Figure 11.4a.

$$E_i(x, y, z) = \frac{\lambda_c}{i} \int \tau(y, z) \frac{e^{i\frac{2\pi}{\lambda_c}(\rho_c + \rho_i)}}{\rho_c \rho_i} \mathcal{O}(\theta_c, \theta_i) dS \quad (11.27)$$

For small angles, we may approximate the obliquity factor to 1. Figure 11.7 shows one case of interest. The transmissivity* function of a small square hologram ($50 \mu\text{m} \times 50 \mu\text{m}$) whose centre is located at $(0,0,0)$ is calculated using a point reference source at $(x, y, z) = (10^{-2}, 10^{-2}, 0 \text{ m})$ and a point object sink at $(-10^{-4}, 0, 0 \text{ m})$. The hologram is then replayed using Equation 11.27 first for a point illumination source at $(10^{-2}, 10^{-2}, 0 \text{ m})$ and subsequently for a shifted source at $(10^{-2}, 1.1 \times 10^{-2}, 0 \text{ m})$ giving $\delta\theta \sim 0.0476$. The wavelength both at recording and at replay is assumed to be 532 nm. Figure 11.7a and b show plots of the calculated values of $\langle E^2 \rangle$ versus the vertical coordinate y at the plane in front of the hologram defined by $x = -10^{-4} \text{ m}$ and $z = 0$. Graph (a) shows the case of the unshifted reference beam and, as expected, shows a strong reconstruction peak at $y = 0$. Graph (b) shows the case of the shifted reference beam, and this time, we see that the peak reconstruction is shifted to $y = -3.4 \mu\text{m}$; this is very close to the result predicted by Equation 11.26. Note that the width of the reconstruction peaks and the presence of the small secondary peaks are determined by the fact that the size of the hologram is only $50 \mu\text{m}$.

* A 3% modulation of the transmissivity is assumed for the purposes of the calculation.

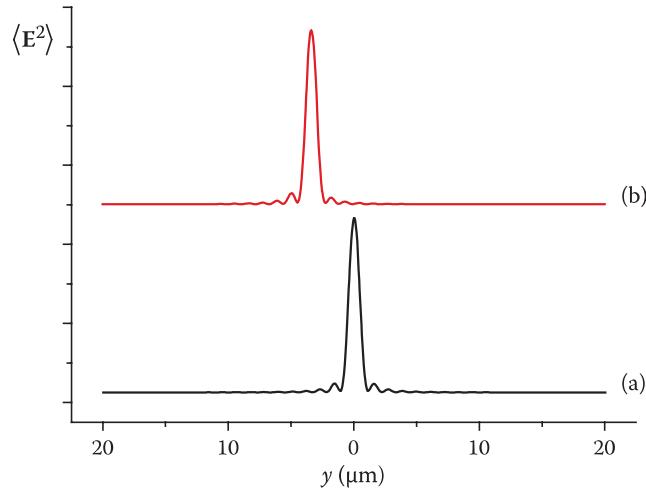


FIGURE 11.7 Diagram showing how a thin amplitude-modulated transmission hologram replays with a slightly shifted reference beam. The square hologram ($50 \times 50 \mu\text{m}$) is recorded using a point reference source at $(x, y, z) = (10^4, 10^4, 0 \mu\text{m})$ and a point object sink at $(-10, 0, 0 \mu\text{m})$. The average squared electric field is calculated using Equation 11.4. The amplitude transmissivity of the processed hologram is assumed to be proportional to this calculated distribution. A Kirchhoff integral is then computationally evaluated for two cases: (a) spherical replay wave originating at $(x, y, z) = (10^4, 10^4, 0 \mu\text{m})$; (b) spherical replay wave originating at $(x, y, z) = (10^4, 1.1 \times 10^4, 0 \mu\text{m})$. The average squared electric field of the diffracted wave is plotted against the y coordinate at $x = -10 \mu\text{m}$ and $z = 0$. The wavelength at recording and replay is 532 nm. As expected, when the replay and reference geometries are the same as in case (a), the reconstructed image point lies exactly at $(-10, 0, 0 \mu\text{m})$. With a shift in replay source as in case (b), the image point is also displaced. The amount of displacement is given by Equation 11.26.

To understand the significance of the relations in Equation 11.26 in terms of what a human observer actually “sees”, we need to define the practical resolution of the human eye. The average human eye can resolve two points separated by 1 mm at a distance of approximately 1 m.* This corresponds to an angle of approximately $\delta\theta_{\text{Eye}} = 0.06^\circ$.

Let us now consider the case of observing a transmission hologram from a distance h with the image being virtual and located behind the hologram surface. The source-size blurring of the image point in the vertical direction now corresponds to an angle of

$$\begin{aligned}\delta\theta_s &\sim \frac{\delta y_i}{h + x_i} \\ &\sim \delta\theta_c \frac{x_i \cos\langle\theta_c\rangle}{h + x_i}\end{aligned}\tag{11.28}$$

When $h = 0$ (in which case we are viewing the hologram with our eyes pressed to the plate) or $x_i \gg h$ (in which case we are considering very large image depths), this reduces simply to

$$\delta\theta_s \sim \delta\theta_c \cos\langle\theta_c\rangle\tag{11.29}$$

This tells us that the virtual image of a transmission hologram will appear blurred to a close observer or at large depths unless the illumination source used to reconstruct it is characterised by

$$\delta\theta_s \leq \delta\theta_{\text{Eye}}\tag{11.30}$$

For all values of $h \neq 0$, Equations 11.28 and 11.30 show that at $x_i = 0$, the virtual image is always perfectly sharp. In addition, close behind the hologram, the image will always appear sharp because for any $\delta\theta_s$, there will always be a zone in which $\delta\theta_s < \delta\theta_{\text{Eye}}$. However, at a critical image depth, blurring will overtake the eye’s resolution and deeper parts of the image will then always appear out of focus. This critical depth is given by

* This varies from person to person and also with the lighting conditions as the pupil size changes.

$$x_{\text{CRIT}} = \frac{h\delta\theta_{\text{Eye}}}{\delta\theta_c \cos\langle\theta_c\rangle - \delta\theta_{\text{Eye}}} \quad (11.31)$$

If we consider source-size blurring in the lateral (z) direction then the factor $\cos\langle\theta_c\rangle$ drops out in Equations 11.28, 11.29 and 11.31. For most light sources, lateral (z) blurring is therefore usually larger than the blurring in the vertical (y) dimension. Equation 11.28 shows that a real image in front of the hologram becomes blurred much more quickly than the virtual image behind the plate. Again, there is always a zone close to the hologram surface where the real image is sharp, but this zone will be shorter in the x dimension by a factor of

$$f = \frac{\delta\theta_c \cos\langle\theta_c\rangle - \delta\theta_{\text{Eye}}}{\delta\theta_c \cos\langle\theta_c\rangle + \delta\theta_{\text{Eye}}} \quad (11.32)$$

As before, if we consider lateral (z) blurring then the $\cos\langle\theta_c\rangle$ terms drop out. One can also note that at $h = x_i$, the blur becomes effectively infinite.

The above discussion is useful as it lets us begin to understand the “mechanics” of the thin transmission hologram. In particular, we see that any thin transmission hologram is highly sensitive to the physical size of the illumination source; as such, these holograms always require an illumination source of high spatial coherence. The discussion is also useful from a very practical point of view—as for digital rainbow and achromatic holograms, Equation 11.20 is precisely what is needed to formulate the required image data transformations.

Many transmission holograms are, however, not intrinsically thin and Bragg selection, which occurs due to the volume nature of any thick hologram, may in some cases somewhat mitigate the effects described by the source-size blurring formulae derived in this previous section. We shall see a little later that if we use a similar treatment to that discussed thus far in relation to the thin transmission hologram but this time for the case of an *infinitely* thick hologram (of almost vanishing permittivity modulation), then source-size blurring in this extreme case essentially disappears for monochromatic illumination. For most practical applications of display holography, however, current silver halide and photopolymer materials are just not thick enough to significantly alter source-size blurring. As a consequence, one of the fundamental constraints of ultra-realistic display holography is the spatial coherence of the illumination source.

11.2.3 Chromatic Blurring

When a thin transmission hologram is illuminated by a broadband source, each wavelength acts to form a virtual image of a given object point at a different location in space. This effect is known as chromatic blurring and is present in both transmission and reflection holograms.

We start by assuming that our illumination source is characterised by a spread of wavelengths $\delta\lambda$, that is,

$$\langle\lambda_c\rangle - \delta\lambda/2 \leq \lambda_c \leq \langle\lambda_c\rangle + \delta\lambda/2 \quad (11.33)$$

Equation 11.20 then defines the corresponding range inherited by x_i and y_i :*

$$\begin{aligned} \delta x_i &= x_o \frac{\lambda_r}{(\langle\lambda_c\rangle - \delta\lambda/2)} - x_o \frac{\lambda_r}{(\langle\lambda_c\rangle + \delta\lambda/2)} \sim x_o \lambda_r \frac{\delta\lambda}{\langle\lambda_c\rangle^2} \\ \delta y_i &= x_o \left(\frac{\lambda_r}{(\langle\lambda_c\rangle - \delta\lambda/2)} \sin\theta_c \right) - x_o \left(\frac{\lambda_r}{(\langle\lambda_c\rangle + \delta\lambda/2)} \sin\theta_c \right) \sim x_o \lambda_r \sin\theta_c \frac{\delta\lambda}{\langle\lambda_c\rangle^2} \\ \delta z_i &= 0 \end{aligned} \quad (11.34)$$

* Note that we could of course have proceeded as in Equation 11.25 by simply differentiating Equation 11.20.

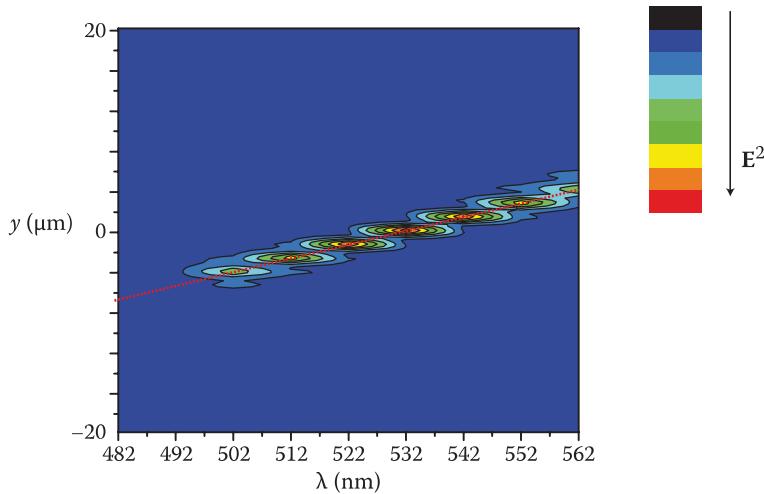


FIGURE 11.8 Contour plot of $\langle \mathbf{E}^2 \rangle$ for the hologram of Figure 11.7 illuminated by a discrete set of wavelengths. The dotted line is the theoretical prediction of the paraxial theory.

This is an important result because it shows that chromatic blurring only affects the vertical and depth dimensions of a thin transmission hologram under the paraxial plane wave approximation. Note, however, that as soon as spherical waves are used for the replay, this result no longer applies and chromatic blurring occurs in all three orthogonal directions.

Using the same methodology we employed with source-size blurring, we can analyse the consequences of chromatic blurring for the thin hologram. In fact, the behaviour is qualitatively identical. Blurring will always be greater for a real image in front of the hologram, and there will, in general, always exist a critical depth behind which all features in the hologram will appear blurred. This critical depth is

$$x_{CRIT} = \frac{\frac{h\delta\theta_{Eye}}{\delta\lambda} \sin\theta_c - \delta\theta_{Eye}}{\frac{\delta\lambda}{\lambda} \sin\theta_c - \delta\theta_{Eye}} \quad (11.35)$$

If $\delta\lambda/\lambda < \delta\theta_{Eye}$, then the entire hologram will appear unblurred. The corresponding formula for the front-projected real image is

$$x_{CRIT} = \frac{\frac{h\delta\theta_{Eye}}{\delta\lambda} \sin\theta_c + \delta\theta_{Eye}}{\frac{\delta\lambda}{\lambda} \sin\theta_c + \delta\theta_{Eye}}. \quad (11.36)$$

We can verify the relations in Equation 11.34 using a numerical evaluation of the Kirchhoff diffraction integral exactly as we did in Section 11.2.2. As an example, Figure 11.8 shows the numerically reconstructed signal of the hologram of Figure 11.7 for different reconstruction wavelengths at the object plane; the replay geometry is the same as that at recording. The dotted red line shows the theoretical fit produced by the paraxial theory, which is clearly in good agreement with the full diffractive simulation.

11.3 Laser Transmission Holograms and the Problem of Speckle

If a thin transmission hologram is illuminated by an effective point source of high temporal coherence at precisely its recording wavelength and recording geometry, the replayed image does not suffer from chromatic or source-size blurring. Large super-realistic monochromatic holograms have been made using this principle, often of very large size [7–9]. However, these types of holograms often suffer from speckle noise if the illumination source is too coherent. The major cause of speckle noise in laser

transmission holograms is subjective speckle—this type of speckle depends on the observer and, in particular, the diameter of the effective entrance pupil of the human eye for normal distance observation of the hologram. Practically, this leads to an angular speckle size of

$$\delta\theta \sim 2.4 \frac{\lambda}{D} \quad (11.37)$$

where D is the human eye diameter. The eye can resolve this and the result is an annoying dynamic random pattern superimposed on an otherwise clear hologram. From a theoretical point of view, the transmission hologram can record structures in objects down to the diffraction limit. However, in practice, speckle can severely limit this capability.

One effective way to reduce speckle is to use a lower temporal coherence of the replaying laser. Copper vapour lasers, for example, have been used very effectively in this regard. Of course, this introduces some chromatic blurring in the transmission hologram, which reduces image resolution, but a careful choice of the temporal coherence can produce stunning speckle-free images of great depth. The illumination source is of vital importance to ultra-realistic holographic imaging; we shall have more to say about this subject in Chapter 13.

11.4 Three-Dimensional Theory of the Thick Transmission Hologram

We can extend, in a simple fashion, our treatment of the thin transmission hologram to a thick hologram. To do this, we first assume that the hologram itself is surrounded by a thick dielectric having the same refractive index as the holographic emulsion and that all the source and reference points lie within this dielectric. Then, all we have to do is to use the following three-dimensional power series expansion of Equation 11.3 instead of our previous two-dimensional expansion:

$$\rho_\gamma = \sum_{m=0}^M \sum_{n=0}^N \sum_{p=0}^P \rho_{\gamma mnp} z_h^m y_h^n x_h^p \quad \forall \gamma \in \{r, o, c, i\} \quad (11.38)$$

where

$$\rho_\gamma(x_h, y_h, z_h) = \sqrt{(x_\gamma - x_h)^2 + (y_\gamma - y_h)^2 + (z_\gamma - z_h)^2} \quad \forall \gamma \in \{r, o, c, i\} \quad (11.39)$$

Mathematically, this is indeed simple—the only detail worth noting is that we must now regard all the Cartesian coordinates as representing an optical path rather than a pure distance—that is, $x \rightarrow nx$, $y \rightarrow ny$ and $z \rightarrow nz$ with n being the refractive index of the holographic emulsion; this is because we are now discussing the propagation of waves within a dielectric medium.

We should, however, be careful of what we are doing here physically. Indeed, we may still assert confidently that two holograms recorded each by a pair of different spherical waves are identical if and only if their time-averaged squared electric field distributions are identical. However, by including an expansion in x_h —or depth—it becomes more difficult to be confident of how a hologram will replay even under identical conditions in which it was recorded! This is because now that the hologram is thick, the replaying electromagnetic wave must propagate through a complex dielectric, and as it does so, its field distribution will inevitably change. Hence, inside the holographic emulsion, the replay wave will not be the same as the recording wave because the recording wave propagated through a uniform dielectric, not a hologram. To properly treat a thick hologram, we will therefore need to solve the wave equation within the holographic emulsion. We will do this later using the coupled wave theory. However, for now, we will continue with the simpler theory because not only will we see that it helps us understand the “mechanics” of the thick hologram, but we will also see that it produces formulae of significant practical use for digital holography.

Substitution of Equation 11.38 into Equation 11.5 defines the various ρ objects:

$$\begin{aligned}\rho_{\gamma 001} &= -\frac{x_\gamma}{\sqrt{x_\gamma^2 + y_\gamma^2 + z_\gamma^2}}; & \rho_{\gamma 010} &= -\frac{y_\gamma}{\sqrt{x_\gamma^2 + y_\gamma^2 + z_\gamma^2}} \\ \rho_{\gamma 100} &= -\frac{z_\gamma}{\sqrt{x_\gamma^2 + y_\gamma^2 + z_\gamma^2}}; & \rho_{\gamma 110} &= -\frac{y_\gamma z_\gamma}{(x_\gamma^2 + y_\gamma^2 + z_\gamma^2)^{3/2}}\end{aligned}\quad (11.40)$$

As before, we impose a paraxial condition on both the “*i*” and “*o*” source points by expanding the coordinate pairs (y_i, z_i) and (y_o, z_o) as Taylor series and then truncating Equation 11.5 at quadratic orders in these four variables. The expansion orders $(m = 1, n = 0, p = 0)$, $(m = 0, n = 1, p = 0)$ and $(m = 1, n = 1, p = 0)$ then lead directly to our previous equations (Equations 11.8 through 11.10). However, the order $(m = 0, n = 0, p = 1)$ leads to an additional equation:

$$\frac{\lambda_c}{\lambda_r} = \pm \left(\frac{x_c}{\sqrt{x_c^2 + y_c^2 + z_c^2}} - 1 \right) \left(\frac{x_r}{\sqrt{x_r^2 + y_r^2 + z_r^2}} - 1 \right)^{-1} \quad (11.41)$$

This equation is the Bragg condition of a thick transmission hologram. It asserts that there is a unique replay wavelength for both the primary reconstructed image and the conjugate image of the hologram, which is determined by the recording geometry, replay geometry and recording wavelength. This is very different from the case of the thin transmission hologram in which any wavelength can be used to reconstruct the hologram. The Bragg effect is, however, disproportionately (from a practical point of view) important for reflection holograms which we shall study in Section 11.5. We shall see shortly that for all current types of transmission holograms used for display applications today, the Bragg effect is relatively weak and as such, for most common transmission holograms, Equation 11.41 constitutes at best an estimation of the wavelength of maximum diffractive replay.

11.4.1 Snell's Law at the Air–Hologram Boundary

Before we continue with a discussion of the consequences of the Bragg condition, we should remember that the thick hologram equations (Equations 11.8 through 11.10 and 11.41) that we have just derived pertain to the case of an infinitely thick dielectric of refractive index n . As such, the points (x_o, y_o, z_o) , (x_r, y_r, z_r) , (x_i, y_i, z_i) and (x_c, y_c, z_c) are all assumed to lie within the dielectric. If we wish to examine the case of these points being located in the air, and the hologram alone being located within the dielectric, then we must use Snell's law at the air–hologram boundary. Snell's law may be written in the coordinate system of Figure 11.3 as

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (11.42)$$

where n_1 and θ_1 are the refractive index and angle of incidence in medium 1, and n_2 and θ_2 are the refractive index and angle of incidence in medium 2.

At this point, it is instructive to write down the solution of Equation 11.5 using the relations in Equation 11.40 and the coordinates of Figure 11.3 without making a paraxial approximation for the object and image points. In this more general case, we obtain the following three relations for the primary non-conjugate image:

$$\frac{\sin^2 \theta_i \cos \phi_i \sin \phi_i}{r_i} = \frac{\sin^2 \theta_c \cos \phi_c \sin \phi_c}{r_c} - \frac{\lambda_c}{\lambda_r} \left\{ \frac{\sin^2 \theta_r \cos \phi_r \sin \phi_r}{r_r} - \frac{\sin^2 \theta_o \cos \phi_o \sin \phi_o}{r_o} \right\} \quad (11.43)$$

$$\sin \theta_i \cos \phi_i = \sin \theta_c \cos \phi_c - \frac{\lambda_c}{\lambda_r} \{ \sin \theta_r \cos \phi_r - \sin \theta_o \cos \phi_o \} \quad (11.44)$$

$$\sin \theta_i \sin \phi_i = \sin \theta_c \sin \phi_c - \frac{\lambda_c}{\lambda_r} \{ \sin \theta_r \sin \phi_r - \sin \theta_o \sin \phi_o \} \quad (11.45)$$

In these equations, all angles are within the dielectric and the distances, r , are real distances rather than optical paths. Now, if we regard the hologram as being surrounded by air instead of a dielectric, then we must replace the angles and distances in Equations 11.43 through 11.45 using the transformation

$$\begin{aligned} \sin \theta_\gamma &\rightarrow \frac{1}{n} \sin \theta_\gamma \\ \phi_\gamma &\rightarrow \phi_\gamma \\ r_\gamma &\rightarrow r_\gamma/n \end{aligned} \quad (11.46)$$

If we apply these transformations to Equations 11.43 through 11.45, then we will arrive at a set of equations that describe a thick transmission hologram of finite refractive index surrounded by air with all angles and distances being those measured in the air. Of course, we implicitly assume a certain ordering here. The hologram is located at $x = 0$ and all source and reference points are regarded as being located in the air to the right of $x = 0$. To contemplate the replay of the hologram, we are obliged to regard its thickness as being small compared with all source and reference distances because an observer would need to be located to the left of the hologram. However, this is not a problem as long as the hologram thickness is assumed to be much greater than the recording and replay wavelengths.

It is immediately evident that Equations 11.43 through 11.45 are invariant under the action of Equation 11.46 and as such the paraxial forms Equations 11.8 through 11.10, or equivalently, Equations 11.16 through 11.18 are also invariant. As such, these equations may be read as describing the exterior ray angles of either a thick or a thin hologram surrounded by air. However, if we write the non-paraxial form of Equation 11.41 in spherical coordinates

$$\frac{\lambda_c}{\lambda_r} = \frac{(\cos \theta_c - \cos \theta_i)}{(\cos \theta_r - \cos \theta_o)} \quad (11.47)$$

it is clear that this equation is not invariant under the action of Equation 11.46. Indeed, it adopts the new paraxial form in terms of exterior angles:

$$\frac{\lambda_c}{\lambda_r} = \left(\sqrt{1 - \frac{1}{n^2} \sin^2 \theta_c} - 1 \right) \left(\sqrt{1 - \frac{1}{n^2} \sin^2 \theta_r} - 1 \right)^{-1} \quad (11.48)$$

or equivalently

$$\frac{\lambda_c}{\lambda_r} = \left(\sqrt{1 - \frac{y_c^2 + z_c^2}{n^2(x_c^2 + y_c^2 + z_c^2)}} - 1 \right) \left(\sqrt{1 - \frac{y_r^2 + z_r^2}{n^2(x_r^2 + y_r^2 + z_r^2)}} - 1 \right)^{-1} \quad (11.49)$$

11.4.2 Blurring in the Ultra-Thick Transmission Hologram

We mentioned in the previous section that the Bragg condition disproportionately affected the reflection hologram rather than the transmission hologram. Most display transmission holograms today are made on emulsions of less than 12 μm thickness, and as such, the Bragg condition does little to alter the spectral behaviour of these holograms over and above that of the standard thin transmission hologram model. The situation is, however, a little different if the emulsion is *much* thicker. In the ultra-thick transmission hologram (of almost vanishing permittivity modulation), Equation 11.48 defines a unique replay wavelength for the primary image at a given recording wavelength and at a given reference angle. As such, an

ultra-thick transmission hologram illuminated by a broadband source of finite angular extent will still show dispersion along the achromatic plane—as for each θ_c within the range of the source $\langle \theta_c \rangle + \delta\theta_c/2 \geq \theta_c \geq \langle \theta_c \rangle - \delta\theta_c/2$, there will always exist a reconstruction wavelength that satisfies Equation 11.48. This means that the dispersion in the vertical direction will generally be limited by the angular size of the source—and the angular extent of this dispersion will be governed by Equation 11.29.

11.5 Reflection Holograms

The preceding analysis may be applied in exactly the same form to the case of a reflection hologram. However, now the object wave propagates in the reverse direction, so

$$\mathbf{E}_o(x_h, y_h, z_h, t) = \Re \left(\mathbf{E}_o \frac{e^{i[\omega t + \frac{2\pi}{\lambda_r} \rho_o(x_h, y_h, z_h)]}}{\rho_o(x_h, y_h, z_h)} \right) \quad (11.50)$$

This then changes the three-dimensional hologram fringe pattern to

$$\begin{aligned} \langle \mathbf{E}^2(x_h, y_h, z_h) \rangle &= \frac{1}{2} \left\{ \frac{|E_r|^2}{\rho_r^2(x_h, y_h, z_h)} + \frac{|E_o|^2}{\rho_o^2(x_h, y_h, z_h)} \right\} \\ &+ \frac{\kappa_{ro}(y_h, z_h) \sqrt{E_r E_o E_r^* E_o^*}}{\rho_o(x_h, y_h, z_h) \rho_r(x_h, y_h, z_h)} \cos \left[\frac{2\pi}{\lambda_r} \rho_o(x_h, y_h, z_h) + \frac{2\pi}{\lambda_r} \rho_r(x_h, y_h, z_h) + \zeta_{ro}(E_r, E_o) \right] \end{aligned} \quad (11.51)$$

which, on application of the expansion in Equation 11.38, gives identical paraxial formulae to Equations 11.8 through 11.10 together with the paraxial Bragg condition, which differs from the transmission formula (Equation 11.49)

$$\frac{\lambda_c}{\lambda_r} = \left(\sqrt{1 - \frac{y_c^2 + z_c^2}{n^2(x_c^2 + y_c^2 + z_c^2)}} + 1 \right) \left(\sqrt{1 - \frac{y_r^2 + z_r^2}{n^2(x_r^2 + y_r^2 + z_r^2)}} + 1 \right)^{-1} \quad (11.52)$$

Note that the image and object points, (x_i, y_i, z_i) and (x_o, y_o, z_o) , are now characterised by having negative values of x . Using the spherical coordinates of Figure 11.3, the associated non-paraxial equations in the plane wave limit ($r \rightarrow \infty$) can be rewritten as

$$\sin \theta_i \cos \phi_i - \sin \theta_c \cos \phi_c = -\frac{\lambda_c}{\lambda_r} \{ \sin \theta_r \cos \phi_r - \sin \theta_o \cos \phi_o \} \quad (11.53)$$

$$\sin \theta_i \sin \phi_i - \sin \theta_c \sin \phi_c = -\frac{\lambda_c}{\lambda_r} \{ \sin \theta_r \sin \phi_r - \sin \theta_o \sin \phi_o \} \quad (11.54)$$

$$\frac{\lambda_c}{\lambda_r} = \left(\sqrt{1 - \frac{1}{n^2} \sin^2 \theta_c} + \sqrt{1 - \frac{1}{n^2} \sin^2 \theta_i} \right) \left(\sqrt{1 - \frac{1}{n^2} \sin^2 \theta_r} + \sqrt{1 - \frac{1}{n^2} \sin^2 \theta_o} \right)^{-1} \quad (11.55)$$

Equations 11.53 and 11.54 are identical to their transmission hologram counterparts, Equations 11.44 and 11.45, as we use the following convention:

- Transmission grating: $0 \leq \theta_o, \theta_i, \theta_r, \theta_c \leq \pi/2; 0 \leq \phi_o, \phi_i, \phi_r, \phi_c \leq 2\pi; r \geq 0$
- Reflection grating: $0 \leq \theta_r, \theta_c \leq \pi/2; \pi/2 < \theta_o, \theta_i \leq \pi; 0 \leq \phi_o, \phi_i, \phi_r, \phi_c \leq 2\pi; r \geq 0$

Equations 11.53 through 11.55 are particularly useful in digital display holography as they describe how a given recorded ray will replay in terms of angle and wavelength.

If we are considering only the two-dimensional case ($\phi = 0$), then we should use the following alternative definitions:

- Transmission grating: $-\pi/2 \leq \theta_o, \theta_i, \theta_r, \theta_c \leq \pi/2; r \geq 0$
- Reflection grating: $\pi/2 < \theta_o, \theta_i \leq 3\pi/2; r \geq 0$

The two-dimensional plane wave formulae for any hologram can now be written as

$$\sin \theta_i - \sin \theta_c = -\frac{\lambda_c}{\lambda_r} \{ \sin \theta_r - \sin \theta_o \} \quad (11.56)$$

$$\frac{\lambda_c}{\lambda_r} = \left(\sqrt{1 - \frac{1}{n^2} \sin^2 \theta_c} \pm \sqrt{1 - \frac{1}{n^2} \sin^2 \theta_i} \right) \left(\sqrt{1 - \frac{1}{n^2} \sin^2 \theta_r} \pm \sqrt{1 - \frac{1}{n^2} \sin^2 \theta_o} \right)^{-1} \quad (11.57)$$

where the plus sign (+) in Equation 11.57 indicates a reflection hologram and the minus sign (-) indicates a transmission hologram.

Despite the apparent similarities, there is an important difference between thick transmission and thick reflection holograms. This is the orientation of the fringe pattern within the grating. In a thick hologram, the fringes are intrinsically three-dimensional and locally form nested planes. We shall see in the next section that a useful model of the thick hologram simply treats these nested planes as mirrors. Constructive and destructive interference then occurs between the various reflections from each such mirror or “Bragg plane”. It is this interference process that determines the wavelength of replay, and it is the (two-dimensional) orientation of the planes that decides principally how a given ray will be “redirected” by the hologram on replay.

The orientation of the fringe planes in a transmission and reflection hologram is quite different (Figure 11.9). This effect is exacerbated by Snell’s law, which acts to steepen the angle of incidence of a light ray within the emulsion layer. The end result is that transmission holograms are usually characterised by steep fringe planes whose normal vector is almost within the hologram plane; reflection

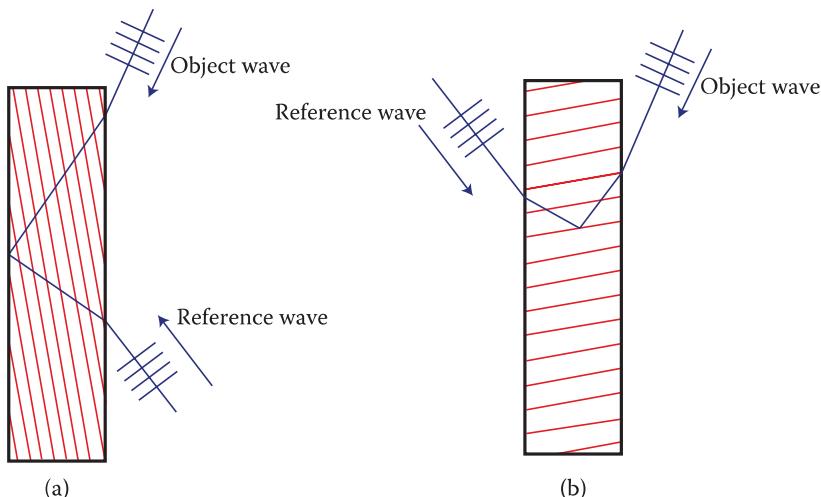


FIGURE 11.9 Typical fringe patterns for (a) a reflection grating and (b) a transmission grating.

holograms, on the other hand, are characterised by very shallow-angle fringe planes whose normal vector is more nearly aligned with the hologram normal. This has a profound effect on the importance of the Bragg selection process. The practical outcome is that common types of transmission volume holograms do not possess nearly as much (spectral) Bragg discrimination as their reflection counterparts. The result is that intermediate thickness transmission holograms lend themselves well to HPO rainbow and achromatic imaging, whereas their reflection counterparts lend themselves better to full-colour HPO imaging. In addition, full-aperture analogue (or full-parallax digital) reflection holograms can be expected to exhibit much deeper unblurred images when illuminated by white-light sources than their transmission counterparts. The higher (spectral) Bragg discrimination of the reflection hologram makes this hologram particularly interesting for large-format full-colour, full-parallax display imaging. Nonetheless, here, as we shall discuss in Chapter 13, very large depths can, from a practical point of view, only be attained by the use of special RGB illumination sources with narrow bands and relatively high spatial coherence.

11.6 A Simple Model of the Thick Hologram: Parallel Stacked Mirrors

The PSM model of a thick-phase hologram is a simple model that leads us to exactly the same results as those presented in the discussion of the previous section. We provide here a very basic and semi-heuristic discussion of this model. Such a discussion is primarily useful in terms of the simple but powerful picture it offers of what is actually going on in any volume transmission or reflection grating. In fact, the PSM model can be formulated in a much more rigorous way [10], but we shall wait until Chapter 12 before going into the extra detail required.

Figure 11.10 illustrates the interaction of two mutually coherent plane waves within a holographic emulsion. We can imagine these two waves as constituting an object and a reference wave. As drawn, the hologram is of the reflection type. The red lines indicate lines where the average square electric field will always be a minimum due to the interference of the two propagating waves. These are then the grating planes. On processing of the hologram, the square of the refractive index can be expected to exhibit a small sinusoidal modulation from plane to plane (this assumes that the permittivity of the medium varies as the time-averaged square of the electric field). On replay, one can expect an incident wave to be reflected from these grating planes. Of course, Maxwell's equations within a dielectric medium simply tell us that a wave will be partially reflected when it encounters any discontinuity in the refractive index [e.g., 11]. However, for the purpose of the present model, we shall make the plausible simplifying assumption that such reflection happens precisely and only at the indicated grating planes.

Analysing Figure 11.10, we see that the angle that the grating planes make to the object wave vector is given by

$$\psi = \frac{1}{2} \cos^{-1} (\hat{\mathbf{k}}_o \cdot \hat{\mathbf{k}}_r) \quad (11.58)$$

Here, $\hat{\mathbf{k}}_o$ and $\hat{\mathbf{k}}_r$ are the unit wave vectors of the object and reference waves. We can write them explicitly as Cartesian column vectors:

$$\hat{\mathbf{k}}_r = - \begin{pmatrix} \cos\theta_r \\ \sin\theta_r \cos\phi_r \\ \sin\theta_r \sin\phi_r \end{pmatrix}; \quad \hat{\mathbf{k}}_o = - \begin{pmatrix} \cos\theta_o \\ \sin\theta_o \cos\phi_o \\ \sin\theta_o \sin\phi_o \end{pmatrix} \quad (11.59)$$

As previously, we use the following convention (see Figure 11.3)

- Transmission grating: $0 \leq \theta_o, \theta_i, \theta_r, \theta_c \leq \pi/2; 0 \leq \phi_o, \phi_i, \phi_r, \phi_c \leq 2\pi; r \geq 0$
- Reflection grating: $0 \leq \theta_r, \theta_c \leq \pi/2; \pi/2 < \theta_o, \theta_i \leq \pi; 0 \leq \phi_o, \phi_i, \phi_r, \phi_c \leq 2\pi; r \geq 0$

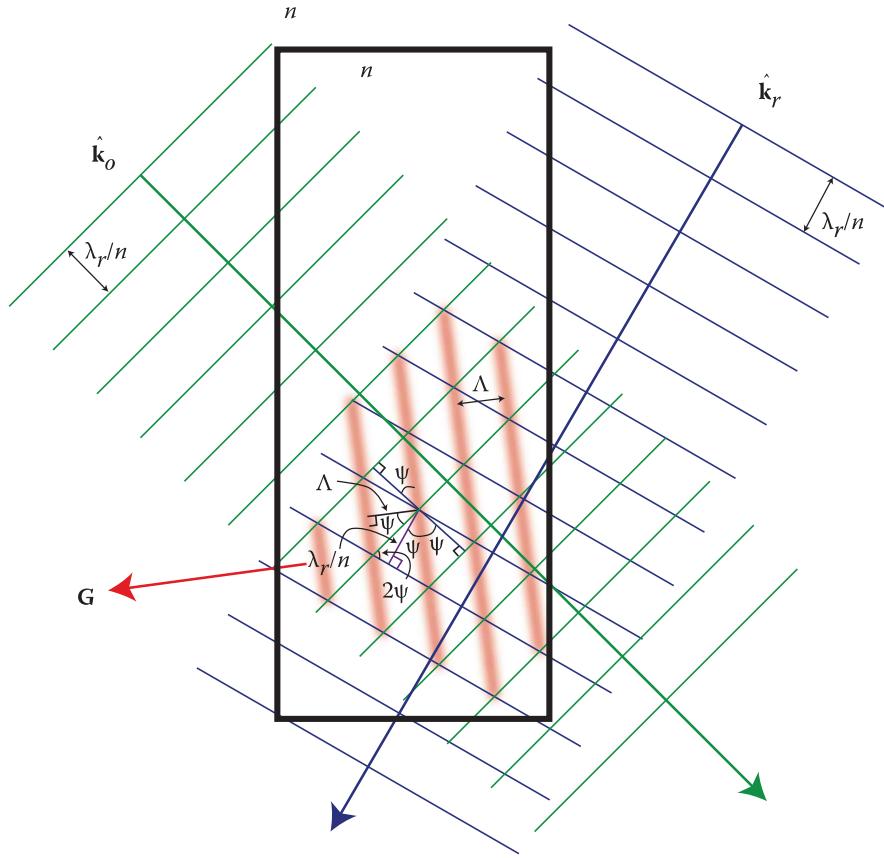


FIGURE 11.10 The interaction of two mutually coherent plane waves within a holographic emulsion with unit wave vectors $\hat{\mathbf{k}}_r$ and $\hat{\mathbf{k}}_o$. We can imagine these two waves as constituting, respectively, a reference and an object wave. As drawn, the hologram is of the reflection type. The red lines indicate lines of the grating planes in which the average square electric field will always be a minimum due to the interference of the two propagating waves.

From Figure 11.10 the grating planes can be seen to exactly bisect the angle between the two wave vectors. The distance between the grating planes is given by

$$\Lambda = \frac{\lambda_r}{n} \left| \frac{\cos \psi}{\sin(2\psi)} \right| = \frac{\lambda_r}{2n |\sin(\psi)|} = \frac{\lambda_r}{\sqrt{2n} \sqrt{1 - \hat{\mathbf{k}}_o \cdot \hat{\mathbf{k}}_r}} \quad (11.60)$$

The unit vector normal to the grating planes is given by

$$\hat{\mathbf{G}} = \frac{(\hat{\mathbf{k}}_r - \hat{\mathbf{k}}_o)}{|\hat{\mathbf{k}}_r - \hat{\mathbf{k}}_o|} \quad (11.61)$$

It is also useful to define the grating vector as

$$\mathbf{G} \equiv \frac{2\pi}{\Lambda} \hat{\mathbf{G}} = \frac{2\pi n}{\lambda_r} (\hat{\mathbf{k}}_r - \hat{\mathbf{k}}_o) \quad (11.62)$$

Figure 11.11 shows how we can imagine the hologram replay process. A plane wave with unit wave vector $\hat{\mathbf{k}}_c$ illuminates the hologram and is reflected from the grating planes producing another plane

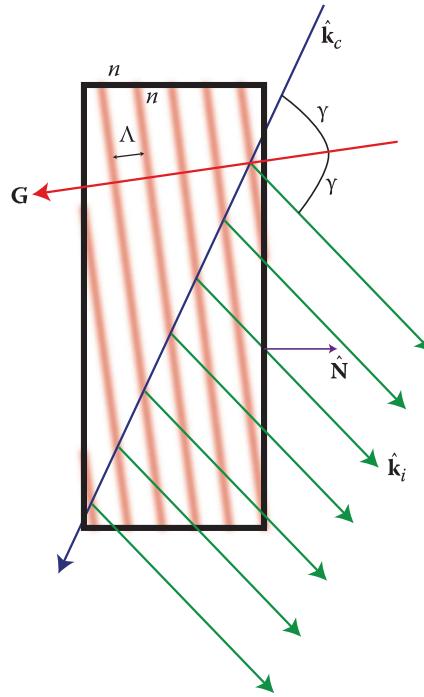


FIGURE 11.11 The hologram replay process. A plane wave with unit wave vector $\hat{\mathbf{k}}_c$ illuminates the grating of Figure 11.10 and is reflected from the grating planes producing another plane wave, this time with unit wave vector $\hat{\mathbf{k}}_i$.

wave, this time with unit wave vector $\hat{\mathbf{k}}_i$. The reflection process follows the normal law of classical reflection with angle of incidence equal to angle of reflection. This is most simply written as

$$\hat{\mathbf{k}}_i \times \hat{\mathbf{G}} = \hat{\mathbf{k}}_c \times \hat{\mathbf{G}} \quad (11.63)$$

This is a natural assumption to make because Equation 11.63 just comes from a plane wave solution of Maxwell's equations across a planar discontinuity in refractive index. Such an analysis [e.g., 11] shows that the amplitude reflection coefficient at near normal incidence for either π or σ -polarisations of the wave is given by

$$R \equiv \delta n / n \quad (11.64)$$

where $\delta n \ll n$ is the change in refractive index across the boundary.

Equation 11.63* can be solved for $\hat{\mathbf{k}}_i$

$$\begin{aligned} \hat{\mathbf{k}}_i &= \hat{\mathbf{k}}_c - 2\hat{\mathbf{G}}(\hat{\mathbf{k}}_c \cdot \hat{\mathbf{G}}) \\ &= \hat{\mathbf{k}}_c + \frac{\hat{\mathbf{k}}_c \cdot (\hat{\mathbf{k}}_r - \hat{\mathbf{k}}_o)}{\hat{\mathbf{k}}_o \cdot \hat{\mathbf{k}}_r - 1}(\hat{\mathbf{k}}_r - \hat{\mathbf{k}}_o) \end{aligned} \quad (11.65)$$

One can immediately see that the angle (in azimuth and altitude) a reconstruction ray reflects off a single grating plane is solely determined, under the PSM model, by the classical law of reflection and is in no

* Note that we discard the trivial solution $\hat{\mathbf{k}}_i = \hat{\mathbf{k}}_c$.

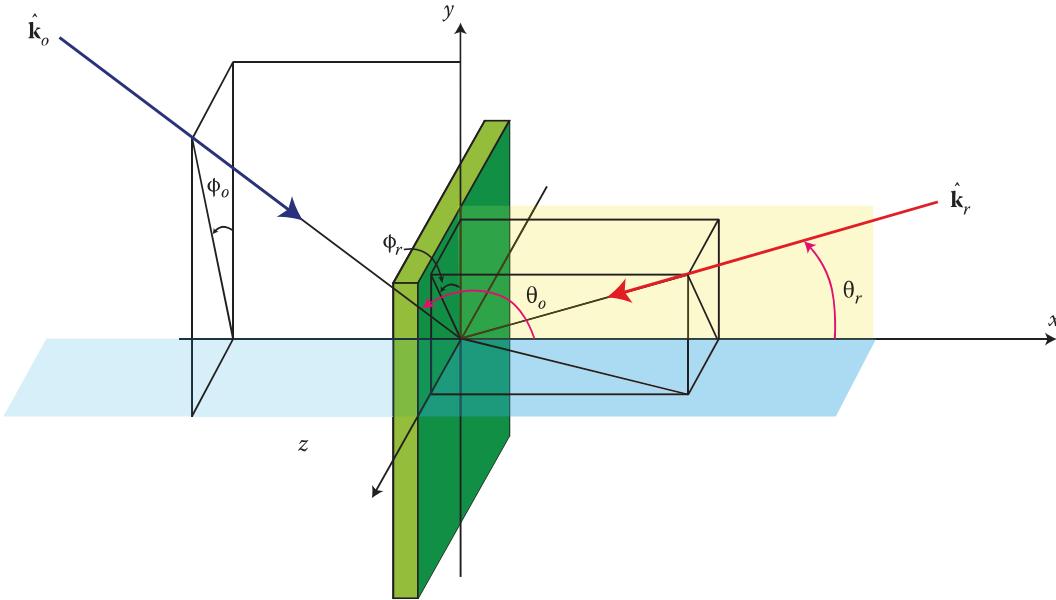


FIGURE 11.12 Altitudinal and azimuthal angles of incidence of the wave vectors of Figures 11.10 and 11.11.

way determined by the wavelength of the reconstructing wave. Of course, if reflections from successive grating planes produce waves that are in phase, then clearly the total reflected energy will be high; conversely, if the waves do not add up, they will tend to cancel. In fact, we can expect that if the parameter δn tends to zero with $(\delta n)d/(\Lambda \hat{\mathbf{G}} \cdot \hat{\mathbf{N}}) \gg 1$ (the thick hologram limit) then there will be an infinite sum of infinitesimal reflections from the grating, giving perfect discrimination and perfect diffraction.* The Bragg condition for constructive reflection can then be written simply as

$$\begin{aligned} \lambda_c/n &= 2\Lambda \cos \gamma \\ &= 2\Lambda \hat{\mathbf{k}}_c \cdot \hat{\mathbf{G}} \\ &= \frac{\lambda_r}{n(1 - \hat{\mathbf{k}}_o \cdot \hat{\mathbf{k}}_r)} \hat{\mathbf{k}}_c \cdot (\hat{\mathbf{k}}_r - \hat{\mathbf{k}}_o) \end{aligned} \quad (11.66)$$

If we now combine Equations 11.65 and 11.66, we arrive at the equation

$$\frac{1}{\lambda_r} (\hat{\mathbf{k}}_r - \hat{\mathbf{k}}_o) = \frac{1}{\lambda_c} (\hat{\mathbf{k}}_c - \hat{\mathbf{k}}_i) \quad (11.67)$$

or defining $\mathbf{k} \equiv \frac{2\pi n}{\lambda} \hat{\mathbf{k}}$

$$\mathbf{k}_c - \mathbf{k}_i = \mathbf{k}_r - \mathbf{k}_o = \mathbf{G} \quad (11.68)$$

Equations 11.66 through 11.68 tell us how any thick hologram (either reflection or transmission) recorded at one wavelength with a given object and reference ray will replay. Equation 11.68 is particularly interesting because it shows us that if we illuminate a thick grating with a plane wave of wave vector \mathbf{k}_c , the diffracted response will be a plane wave of wave vector $\mathbf{k}_c - \mathbf{G}$. Using the expressions in

* At least for the σ -polarisation. The π -polarisation is slightly more complicated as here the electric field vectors of both the illumination wave and the diffracted signal wave are in the plane of incidence and so under certain circumstances may be mutually orthogonal. This effect is discussed in Section 11.10.6.1.

Equation 11.59 for the unit wave vectors, we can rewrite these equations in terms of the altitudinal and azimuthal angles of incidence of the rays* as illustrated in Figure 11.12.

$$\frac{1}{\lambda_r}(\sin\phi_o \sin\theta_o - \sin\phi_r \sin\theta_r) = \frac{1}{\lambda_c}(\sin\phi_i \sin\theta_i - \sin\phi_c \sin\theta_c) \quad (11.69)$$

$$\frac{1}{\lambda_r}(\cos\phi_o \sin\theta_o - \cos\phi_r \sin\theta_r) = \frac{1}{\lambda_c}(\cos\phi_i \sin\theta_i - \cos\phi_c \sin\theta_c) \quad (11.70)$$

$$\frac{1}{\lambda_r}(\cos\theta_o - \cos\theta_r) = \frac{1}{\lambda_c}(\cos\theta_i - \cos\theta_c) \quad (11.71)$$

These equations should be familiar! They are just our previous non-paraxial Equations 11.47, 11.53 and 11.54. The PSM model can therefore be seen to lead to exactly the same results as our previous model, which sought to equate the average squared electric field distributions of two holograms recorded at different wavelengths and ray geometries.

Of course, we have not discussed the problem of Snell's law at the hologram-air boundary yet; as a result, all the wave vectors and angles discussed in the previous sections pertain to the holographic dielectric itself. However, converting the angles is simple, as we have seen in Section 11.4.1. As before, we see that Equations 11.69 and 11.70 are invariant under Snell's law. Equation 11.71 must be rewritten for exterior angles as

$$\frac{1}{\lambda_r} \left(\sqrt{1 - \frac{1}{n^2} \sin^2 \theta_o} \pm \sqrt{1 - \frac{1}{n^2} \sin^2 \theta_r} \right) = \frac{1}{\lambda_c} \left(\sqrt{1 - \frac{1}{n^2} \sin^2 \theta_i} \pm \sqrt{1 - \frac{1}{n^2} \sin^2 \theta_c} \right) \quad (11.72)$$

where a “+” sign describes the reflection grating and a “−” sign the transmission grating. It is sometimes quite useful to write Snell's law in terms of vectors

$$\frac{1}{n} \hat{\mathbf{k}}_{\text{ext}} \times \hat{\mathbf{N}} = \hat{\mathbf{k}}_{\text{int}} \times \hat{\mathbf{N}} \quad (11.73)$$

Here, $\hat{\mathbf{N}}$ is the unit normal vector of the hologram. This equation has solutions:

$$\begin{aligned} \hat{\mathbf{k}}_{\text{ext}} &= n \hat{\mathbf{k}}_{\text{int}} - \hat{\mathbf{N}} \left\{ n (\hat{\mathbf{k}}_{\text{int}} \cdot \hat{\mathbf{N}}) - \sqrt{(1 - n^2) + n^2 (\hat{\mathbf{k}}_{\text{int}} \cdot \hat{\mathbf{N}})^2} \right\} \\ \hat{\mathbf{k}}_{\text{int}} &= \frac{1}{n} \hat{\mathbf{k}}_{\text{ext}} - \hat{\mathbf{N}} \left\{ \frac{1}{n} (\hat{\mathbf{k}}_{\text{ext}} \cdot \hat{\mathbf{N}}) - \sqrt{\left(1 - \frac{1}{n^2}\right) + \frac{1}{n^2} (\hat{\mathbf{k}}_{\text{ext}} \cdot \hat{\mathbf{N}})^2} \right\} \end{aligned} \quad (11.74)$$

Equations 11.69 through 11.72 constitute a useful mathematical model of the thick hologram. They are valid for all angles, not just paraxial angles. In particular, they provide a simple and effective method when digital image data must be recalculated to compensate for changes in replay wavelength and reference beam geometry in the digital hologram.

11.7 Holograms of Finite Thickness

In the previous section, we described a basic variant of the PSM model of the thick hologram. In the limit that the number of grating planes sampled by an illuminating wave is large the model predicted

* Note that these angles are internal angles within the grating dielectric.

an essentially perfect wavelength discrimination as defined by Equation 11.66. This led to the relation (Equation 11.68) showing that an incident reference wave $\mathbf{E} = \mathbf{E}_c \exp(i\mathbf{k}_c \cdot \mathbf{r})$ provoked the hologram to produce a response wave $\mathbf{E} = \mathbf{E}_r \exp(i\mathbf{k}_c \cdot \mathbf{r} - i\mathbf{G} \cdot \mathbf{r})$.

The basic PSM model is however not fully consistent for real gratings. For example, in the case of the thick reflection hologram, an illuminating plane wave will not sample all the grating planes because, at each plane, a proportion of its amplitude will be reflected. At a certain depth, the illuminating wave will be completely depleted and effectively the hologram will “look” to this wave like a grating of finite thickness. The problem then is that we cannot assume that the Bragg condition must be exactly satisfied and so in general there will be a band of wavelengths and angles under which the hologram will replay. To analyse this situation properly, we must formulate the problem in terms of two waves—an illuminating wave, which gradually becomes depleted during its passage through the hologram, and a response wave, which is gradually created by the depletion of the first wave. By solving Maxwell’s equations within the grating and imposing proper boundary conditions at the hologram boundary, we can consistently* work out how this conversion process happens when the Bragg condition is satisfied only approximately. The first properly successful approach to this problem was published by Kogelnik [1] in 1969, and is known as the coupled wave theory. We shall describe this theory in detail in Section 11.10. A proper treatment of the PSM model may, however, also be used to solve this problem and has the important advantage of being able to treat full-colour gratings with great simplicity. We shall return to this more rigorous PSM theory in Chapter 12. For now, however, we shall restrict ourselves to making several naive (but nevertheless quite accurate) calculations of what a more basic version of the PSM model can say about a volume hologram of finite thickness.

Assuming a lossless phase hologram of finite thickness, we can use Equation 11.64 to estimate the strength of the response wave when we illuminate the hologram with a given illumination wave. Typical values of $\delta n/n$ for a modern panchromatic silver halide emulsion are in the region of $\delta n/n \sim 0.03$. The thickness of such emulsions is approximately 10 μm , giving approximately 20 grating planes for an unslanted reflection geometry. Coherent summation of the reflections then theoretically leads to a response of approximately 60% of the amplitude of the illuminating beam. Of course, this is a rather crude calculation, but it does serve to give some understanding that with reasonable modulation, volume phase holograms of the reflection variety do not have to be very thick to produce strong image signals on replay.

Another interesting calculation [4] is worthwhile. Suppose we do not exactly satisfy the Bragg condition within a hologram of N (unslanted) grating planes. For example, suppose our replay wavelength misses the Bragg wavelength by an amount $\delta\lambda$. In this case, the reflected signal from each grating plane will be dephased by $2\pi \delta\lambda/\lambda$. Coherently adding up the various reflections, we see that an estimate of the resultant image signal amplitude is given by

$$S \sim \frac{\delta n}{n} \left| \sum_{\beta=1}^N e^{i\beta(2\pi\delta\lambda/\lambda)} \right| = \frac{\delta n}{n} \left| \frac{\sin(N\pi\delta\lambda/\lambda)}{\sin(\pi\delta\lambda/\lambda)} \right| \quad (11.75)$$

For zero dephasing, we retrieve the relation we used above— $S = N\delta n/n$. At $N\delta\lambda/\lambda = 1$, however, S reaches zero. Now for an unslanted transmission grating, the number of grating planes crossed by a ray incident at angle θ_c (Figure 11.13) will be given by

$$N = \frac{d \tan \theta_c}{\Lambda} = \frac{|\mathbf{G}| d \tan \theta_c}{2\pi} \quad (11.76)$$

And so we can expect our hologram to replay over a range of wavelengths given by

$$\frac{\delta\lambda}{\lambda} \sim \frac{2\pi}{|\mathbf{G}| d \tan \theta_c} \sim \frac{\lambda}{2dn \sin \theta_c \tan \theta_c} \quad (11.77)$$

* One should note that for full consistency rigorous coupled wave theory must be used where higher diffraction orders are treated – see Appendix 8.

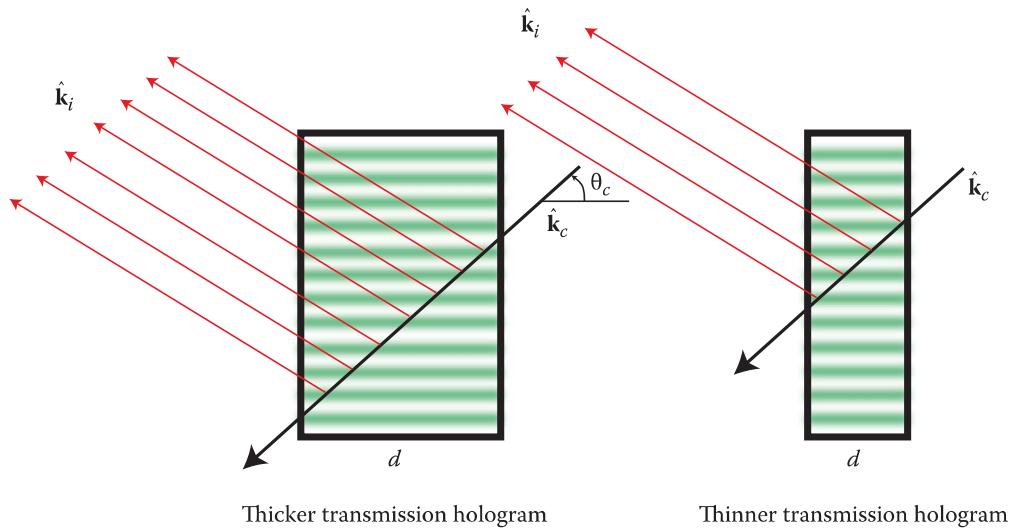


FIGURE 11.13 Replay from an unslanted transmission grating. The number of grating planes crossed by a ray incident at angle θ_c depends on the thickness of the grating.

Here, we have assumed that the hologram has been written at wavelength λ , which is also approximately the replay wavelength, and that the writing and replay reference angles are both equal to θ_c . For the corresponding case of the unslanted reflection grating, a ray transverses all the grating planes and so

$$N = \frac{d \tan \theta_c}{\Lambda} = \frac{|\mathbf{G}| d}{2\pi} \quad (11.78)$$

giving

$$\frac{\delta\lambda}{\lambda} \sim \frac{\lambda}{2dn \cos \theta_c} \quad (11.79)$$

Let us take the following as an example. We will use a wavelength of 532 nm and an emulsion of 12 μm thickness, recorded and replayed with an interior dielectric angle of $\theta_c = 15^\circ$. We will assume an average index of $n = 1.5$. Then, for the transmission hologram, we estimate a fractional bandwidth of 21%. For the corresponding reflection hologram, the bandwidth comes out as 1.5%. We shall see later that these crude estimations of the PSM model are actually rather accurate, differing only from those produced by the coupled wave theory by 10% to 15%.

These calculations serve to give us a very good understanding of what is going on in the hologram. Specifically, we can see that a ray in a reflection hologram undergoes much more wavelength discrimination than a ray in the corresponding transmission case. This is because of the angle of the grating planes but also because of the distance between the grating planes (this is much smaller in the reflection case). These two effects multiply and lead to a much stronger wavelength selectivity in the reflection hologram.

One might have naively thought that because a ray in a transmission grating generally samples far fewer Bragg planes than in the corresponding reflection case, then the transmission hologram might well be fundamentally less diffractively efficient than the reflection hologram of a similar thickness. However, the Fresnel formulae tell us that the reflection coefficient at a discontinuous interface in permittivity strongly depends on incidence angle, with smaller angles giving substantially higher reflections. This exact effect operates in the transmission hologram. So even though a ray in this type of hologram samples fewer Bragg planes than a ray in the corresponding reflection case, reflection from any one Bragg plane, being at a much higher angle of incidence, means that ultimately there is little difference in the diffractive powers of similar thickness transmission and reflection holograms.

We should note here that amplitude transmission holograms are fundamentally less diffractive than phase holograms. This was demonstrated very effectively by Kogelnik's work [1]. One can show, however, that phase holograms must in many cases be fundamentally more noisy than amplitude holograms [5], and originally, this created the opinion (circa 1970) that phase holograms would always produce inferior quality images. However, great advances in processing and materials have occurred since—for example the discovery of reversal bleaching—and as a consequence, modern day phase holograms are largely regarded as being superior to amplitude holograms in virtually all applications of display holography. As a consequence, we will have little to say about amplitude holograms in this book.

11.8 Emulsion Swelling and Change in Refractive Index on Processing

In general, any thick hologram, whether of the transmission or reflection type, will undergo a change in both the refractive index and the emulsion thickness on processing. Equation 11.67 must then be changed to the following form:

$$\frac{n_r}{\lambda_r} \underline{\alpha} \cdot (\hat{\mathbf{k}}_r - \hat{\mathbf{k}}_o) = \frac{n_c}{\lambda_c} (\hat{\mathbf{k}}_c - \hat{\mathbf{k}}_i) \quad (11.80)$$

Here, n_r represents the refractive index on recording, n_c is the refractive index after processing and $\underline{\alpha} = \text{diag}(\tau, 1, 1)$ is the emulsion swelling matrix with the scalar parameter τ representing the linear emulsion contraction in the x direction (normal to the hologram surface). The wave vectors in this formula are of course the wave vectors within the emulsion dielectric and Equations 11.73 and 11.74 must be used to convert them to a form valid outside the hologram-air boundary.

Emulsion swelling acts to change both the grating angle and the grating separation. Denoting unprimed quantities as pertaining to the situation before swelling and primed quantities for the situation after swelling, we can write

$$\begin{aligned} \hat{\mathbf{G}}' &= \frac{\underline{\alpha} \cdot \hat{\mathbf{G}}}{|\underline{\alpha} \cdot \hat{\mathbf{G}}|} \\ \Lambda' &= \Lambda |\underline{\alpha} \cdot \hat{\mathbf{G}}|^{-1} \end{aligned} \quad (11.81)$$

By substitution of Equation 11.59 into Equation 11.80, we see once again that the y and z components of Equation 11.80 are Snell invariant. The x component must, however, be modified. Accordingly, in terms of exterior (air) angles, we can write down the general non-paraxial equations relating the recorded object and the reconstructed image rays:

$$\frac{n_r}{\lambda_r} (\sin \phi_o \sin \theta_o - \sin \phi_r \sin \theta_r) = \frac{n_c}{\lambda_c} (\sin \phi_i \sin \theta_i - \sin \phi_c \sin \theta_c) \quad (11.82)$$

$$\frac{n_r}{\lambda_r} (\cos \phi_o \sin \theta_o - \cos \phi_r \sin \theta_r) = \frac{n_c}{\lambda_c} (\cos \phi_i \sin \theta_i - \cos \phi_c \sin \theta_c) \quad (11.83)$$

$$\frac{n_r \tau}{\lambda_r} \left(\sqrt{1 - \frac{1}{n_r^2} \sin^2 \theta_o} \pm \sqrt{1 - \frac{1}{n_r^2} \sin^2 \theta_r} \right) = \frac{n_c}{\lambda_c} \left(\sqrt{1 - \frac{1}{n_c^2} \sin^2 \theta_i} \pm \sqrt{1 - \frac{1}{n_c^2} \sin^2 \theta_c} \right) \quad (11.84)$$

From the form of these equations, we can expect an emulsion thickness change and a change of refractive index on processing to lead to slightly different behaviours. However, both effects will change the

replay characteristics of any transmission or reflection hologram. This will manifest itself as a different optimal replay wavelength being associated with a given replay geometry and a modification of the object point to image point mapping. It is perhaps worth pointing out that Equations 11.82 through 11.84 only have solutions for a given range of replay angles. Outside of these angles, there is no wavelength to satisfy the Bragg condition. This is particularly important for the reflection hologram.

11.9 Non-Paraxial Behaviour and Digital Image Predistortion

We have discussed at length the paraxial behaviour of both the thin and thick holograms. We have seen that as long as we replay our hologram with a point source, then a given object point will map into a well-defined image point. This is the case for both transmission and reflection holograms of both the thin and thick variety. For many holograms, however, we cannot make the paraxial approximation as the image and object points are characterised by large altitudinal angles. In this more general regime, Equations 11.82 through 11.84 show us that an object point will only map onto a well-defined image point if there is no swelling and no refractive index change on processing and, furthermore, if both the recording and replay geometries are the same and the replay wavelength is the same as the recording wavelength. If these rather stringent conditions are not satisfied, then we can expect different physical parts of the hologram to possess different object-image point mappings. The result is a chromatic and geometric aberration of the holographic image.

Non-paraxial image aberration may be very effectively corrected in digital holograms using the mathematical model we have developed. Often, digital reflection holograms are recorded at a slightly different wavelength than that desired to replay them. This may be due to intrinsic emulsion swelling with optimal chemistries or simply because available lasers have too high or too low a wavelength for optimal replay. Either way, without correction, the digital images suffer from noticeable chromatic and geometric aberration. Digital holograms are also very often written using a collimated reference beam—only to be illuminated using a spot lamp. We have seen in Chapter 7 that it is easier to use a collimated reference beam in a digital printer. In Appendix 4, we discuss in detail how the mathematical model of this chapter may be applied to the digital image data of a DWDH hologram to correct the induced geometric and chromatic aberrations, greatly enhancing the image realism.

11.10 Solving the Helmholtz Equation in Volume Gratings: Coupled Wave Theory

A simple but rather more rigorous approach to the volume hologram is to assume the existence of just two plane waves propagating in and outside a grating of finite thickness and to use the Helmholtz equation to calculate how a specific modulation in the dielectric permittivity intrinsically couples these waves. This approach originates from the field of acousto-optics but was first applied to holography by Kogelnik [1] in 1969. The first wave is assumed to be the illuminating reference wave and the second wave is the hologram's response or “signal” wave. The adoption of just two waves is made on the assumption that coupling to higher-order modes will be negligible. There is no rigorous mathematical proof for this *per se*. However numerical results from a fully accurate solution which is available from Moharam and Gaylord's rigorous coupled wave theory [12] (see Appendix 8) show good agreement with the two-wave theory for most practical gratings. The results of Kogelnik's coupled wave theory are simple and extremely useful for the purposes of display holography and holographic imaging.

11.10.1 One-Dimensional Coupled Wave Theory

Assuming a time dependence of $\sim \exp(i\omega t)$, we use Maxwell's equations and Ohm's law to write down the general wave equation in a dielectric in SI units:

$$\nabla \times (\nabla \times \mathbf{E}) - \gamma^2 \mathbf{E} = 0 \quad (11.85)$$

where

$$\gamma^2 = i\omega\mu\sigma - \omega^2\mu\epsilon \quad (11.86)$$

Here, μ is the permeability of the medium, ϵ its permittivity and σ represents its electrical conductivity. We shall now make two important assumptions. The first is that our grating is lossless, so $\sigma = 0$. The second is that the polarisation of our two waves is perpendicular to the grating vector or $\mathbf{E} \cdot \nabla\epsilon = 0$. This allows the simplification of Equation 11.85 to the Helmholtz equation:

$$\nabla^2\mathbf{E} - \gamma^2\mathbf{E} = 0 \quad (11.87)$$

The assumption of small conductivity means that we restrict our analysis to ideal phase holograms with no absorption. The assumption that $\mathbf{E} \cdot \nabla\epsilon = 0$ leads us to study gratings that have their grating vector in the propagation plane of two σ -polarised waves. These restrictions are not terribly constraining and the picture we get from the much-simplified analysis makes adopting them worthwhile. The interested reader is referred to Kogelnik [1] and Solymar and Cooke [4] for a discussion of more general models.

We will assume a one-dimensional grating extending from $x = 0$ to $x = d$. The relative permittivity is assumed to vary within the grating as

$$\epsilon_r = \epsilon_{r0} + \epsilon_{r1} \cos \mathbf{G} \cdot \mathbf{r} \quad (11.88)$$

We may therefore write the γ parameter as

$$\gamma^2 \sim -\beta^2 - 4\kappa\beta \cos \mathbf{G} \cdot \mathbf{r} \quad (11.89)$$

with

$$\beta = \omega(\mu\epsilon_0\epsilon_{r0})^{1/2} \quad (11.90)$$

and where we have introduced Kogelnik's coupling constant

$$\kappa = \frac{1}{4} \frac{\epsilon_{r1}}{\epsilon_{r0}} \beta \sim \frac{1}{2} \left(\frac{\delta n}{n} \right) \beta n. \quad (11.91)$$

11.10.2 Solution with Perfect Bragg Compliance

In Section 11.6, we saw how the Bragg condition related the image and response wave vectors to the grating vector. Specifically, Equation 11.68 told us that

$$\mathbf{k}_i = \mathbf{k}_c - \mathbf{G} \quad (11.92)$$

The magnitude of both \mathbf{k}_c and \mathbf{k}_i was also shown to be exactly $\beta = 2\pi n/\lambda_c$. We will first investigate the case of perfect Bragg compliance; due to symmetry arguments, we would strongly expect the highest diffractive response to be produced by such compliance.* Accordingly, we will assume that

$$\gamma^2 = -\beta^2 - 4\kappa\beta \cos(\mathbf{k}_c - \mathbf{k}_i) \cdot \mathbf{r} \quad (11.93)$$

and that

$$|k_c| = |k_i| = \beta = \frac{2\pi n}{\lambda_c} \quad (11.94)$$

We now choose a very particular trial solution of the form

$$E_z = R(x)e^{-i\mathbf{k}_c \cdot \mathbf{r}} + S(x)e^{-i\mathbf{k}_i \cdot \mathbf{r}} \quad (11.95)$$

* This is almost always the case.

The first term represents the input wave and the second term represents the response wave (Figure 11.14). Both are plane waves. Note that the complex functions R and S are functions of x only—even though the wave vectors \mathbf{k}_c and \mathbf{k}_i have both x and y components. The grating is assumed to be surrounded by a dielectric having the same permittivity and permeability as the average values within the grating so as not to unduly complicate the boundary conditions. Within the external dielectric, both R and S are constants. The choice of just using two waves in the calculation—the absolute minimum—with only a one-dimensional behaviour was inspired by the work of Bhatia and Noble [13] and Phariseau [14] in the field of acousto-optics. However, one really proceeds here with the philosophy that if the trial solution of Equation 11.95 cannot sensibly satisfy the Helmholtz equation given in Equation 11.87 with sensible

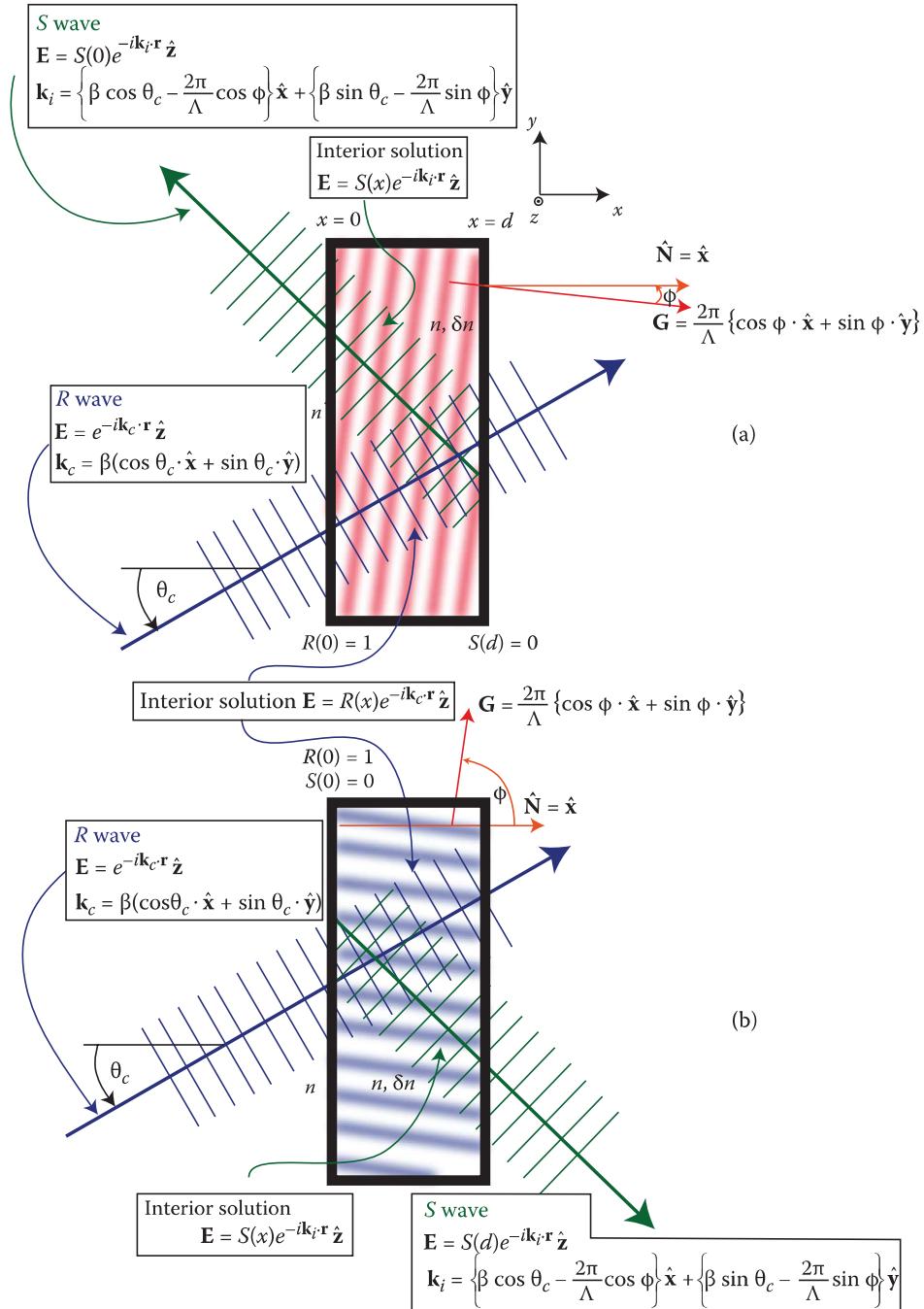


FIGURE 11.14 Diagram showing how the R and S waves of Kogelnik's coupled wave theory propagate within (a) a reflection grating and (b) a transmission grating.

boundary conditions, then the process will not provide a sensible answer. As we progress, we shall make more assumptions that appear at first sight to be at best plausible—but in the end, we shall see that the whole procedure provides a consistent method producing sensible results. With this in mind, we substitute Equations 11.93 through 11.95 into Equation 11.87 to obtain

$$e^{-ik_c \cdot r} \left\{ \frac{d^2 R}{dx^2} - 2ik_{cx} \frac{dR}{dx} + 2\beta\kappa S \right\} + e^{-ik_i \cdot r} \left\{ \frac{d^2 S}{dx^2} - 2ik_{ix} \frac{dS}{dx} + 2\beta\kappa R \right\} + 2\beta\kappa S e^{-i(2k_i - k_c) \cdot r} + 2\beta\kappa R e^{-i(2k_c - k_i) \cdot r} = 0 \quad (11.96)$$

Because we are assuming only two waves in the solution, we must now disregard the third and fourth terms of this expression assuming that they inherit only negligible energy from the primary modes. The next assumption is to neglect second-order derivatives on the premise that R and S are slowly varying functions. Given that the modulation of permittivity is small in all usual cases, this is quite plausible. With these assumptions in hand, Equation 11.96 implies the following two coupled first-order ordinary differential equations:

$$\frac{k_{cx}}{\beta} \frac{dR}{dx} + i\kappa S = 0 \quad (11.97)$$

$$\frac{k_{ix}}{\beta} \frac{dS}{dx} + i\kappa R = 0 \quad (11.98)$$

We can then use Equations 11.97 and 11.98 to write down identical uncoupled second-order differential equations for R and S :

$$\begin{aligned} \frac{d^2 R}{dx^2} + (\kappa^2 \sec \theta_c \sec \theta_i) R &= 0 \\ \frac{d^2 S}{dx^2} + (\kappa^2 \sec \theta_c \sec \theta_i) S &= 0 \end{aligned} \quad (11.99)$$

Here, the x component of the Bragg condition tells us that

$$\sec \theta_i = \left\{ \cos \theta_c - \frac{\lambda_c G_x}{2\pi n} \right\}^{-1} \quad (11.100)$$

Of course, if the grating has been written by a reference and object wave having angles of incidence of, respectively, θ_r and θ_o and at a wavelength of λ_r , then

$$G_x = \frac{2\pi n}{\lambda_r} (\cos \theta_r - \cos \theta_o) \quad (11.101)$$

11.10.3 Boundary Conditions

We have assumed that the R wave is the input wave and the S wave is the response. Normalising the input amplitude to unity, we can therefore write down different boundary conditions for transmission and reflection holograms. For transmission holograms, we must demand that $R(0) = 1$ (i.e., all the power is in the input wave at the point where the wave enters the grating) and $S(0) = 0$ (i.e., the power of the transmitted response wave must be zero at the input boundary as evidently no conversion has yet taken place). For reflection holograms, we demand that $R(0) = 1$ and $S(d) = 0$. Here, on the entrance boundary, we demand unit power in the input wave. Because the reflected response wave is travelling in the direction $x = d$ to $x = 0$, its amplitude must clearly be zero at the far boundary.

Armed with these boundary conditions, we can now solve Equation 11.99 (or equivalently Equations 11.97 and 11.98) for the transmission and reflection cases. For transmission holograms, we have

$$\begin{aligned} R &= \cos\left\{\kappa x(\sec\theta_c \sec\theta_i)^{1/2}\right\} \\ S &= -i\sqrt{\frac{\cos\theta_c}{\cos\theta_i}} \sin\left\{\kappa x(\sec\theta_c \sec\theta_i)^{1/2}\right\} \end{aligned} \quad (11.102)$$

And for reflection holograms

$$\begin{aligned} R &= \operatorname{sech}\left\{\kappa d\left(\sec\theta_c |\sec\theta_i|\right)^{1/2}\right\} \cosh\left\{\kappa(d-x)\left(\sec\theta_c |\sec\theta_i|\right)^{1/2}\right\} \\ S &= -i\sqrt{\frac{\cos\theta_c}{|\cos\theta_i|}} \operatorname{sech}\left\{\kappa d\left(\sec\theta_c |\sec\theta_i|\right)^{1/2}\right\} \sinh\left\{\kappa(d-x)\left(\sec\theta_c |\sec\theta_i|\right)^{1/2}\right\} \end{aligned} \quad (11.103)$$

These are remarkably simple solutions that paint a very logical picture. For the transmission case, we see that as the input wave enters the grating, it slowly donates power to the response wave which grows with increasing x . When the argument of the cosine function in Equation 11.102 reaches $\pi/2$, all power has been transferred to the S wave, which is now at a maximum. As x increases further, the waves change roles; the S wave now slowly donates power to a newly growing R wave. This process goes on until the waves exit the grating.

In the reflection case, the behaviour is different. Here, as one might well expect, there is simply a slow transfer of energy from the input wave to the reflected response wave. If the emulsion is thin, then the response wave is weak and most of the energy escapes as a transmitted R wave. If the emulsion is thick, then the amplitudes of both waves become exponentially small as x increases and all the energy is transferred from the R wave to the reflected S wave.

11.10.4 Power Conservation

Using Poynting's theorem, one can show that the power flowing along the x direction is given by

$$P = \cos\theta_c RR^* + \cos\theta_i SS^* \quad (11.104)$$

Multiplying Equations 11.97 and 11.98 by R^* and S^* , respectively, and then adding these equations and taking the real part, we see that $dP/dx = 0$. This is indeed indicative that the assumptions made were sensible.

11.10.5 Diffraction Efficiency

It is of particular interest to us to understand how efficient we can expect a holographic grating to be. With this in mind, we define the diffraction efficiency of a grating with reference wave of unit amplitude as

$$\eta = \frac{|k_{ix}|}{k_{ex}} SS^* \quad (11.105)$$

where S is evaluated on the exit boundary. It is now simple to use the forms for R and S given in Equations 11.102 and 11.103 to calculate the expected diffractive efficiencies for transmission and reflection holograms:

$$\eta_T = \sin^2\left\{\kappa d(\sec\theta_c \sec\theta_i)^{1/2}\right\} \quad (11.106)$$

$$\eta_R = \tanh^2\left\{\kappa d\left(\sec\theta_c |\sec\theta_i|\right)^{1/2}\right\} \quad (11.107)$$

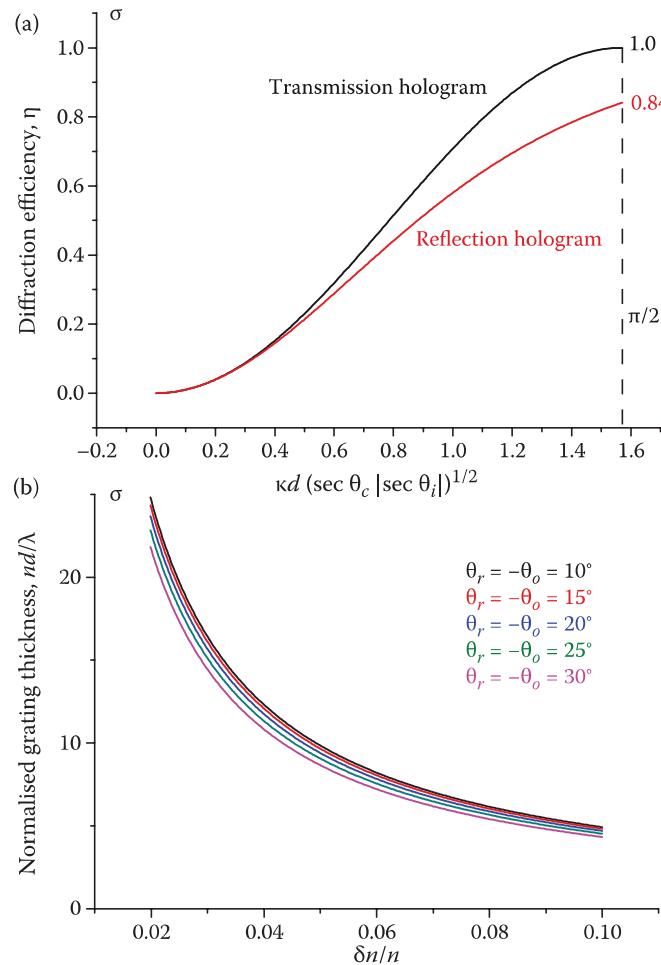


FIGURE 11.15 Perfect Bragg compliance: (a) diffractive replay efficiencies (σ -polarisation) of the transmission hologram (η_T) and the reflection hologram (η_R) versus the normalised grating thickness, $\kappa d(\sec \theta_c \sec \theta_i)^{1/2}$ according to Kogelnik's coupled wave theory. (b) Optimal value of the normalised grating thickness (providing $\eta_T = 1$) versus the modulation, $\delta n/n$ for a transmission hologram.

Figure 11.15a shows this graphically for $0 \leq \kappa d(\sec \theta_c \sec \theta_i)^{1/2} \leq \pi/2$. Clearly, for a small emulsion thickness or for a small permittivity modulation, the diffractive efficiencies of the reflection and transmission types of holograms are identical. As the parameter $\kappa d(\sec \theta_c \sec \theta_i)^{1/2}$ increases towards $\pi/2$, the transmission hologram becomes slightly more diffractive than its corresponding reflection counterpart. However, as we have remarked in the previous section, when $0 \leq \kappa d(\sec \theta_c \sec \theta_i)^{1/2} > \pi/2$, the transmission hologram decreases in diffractive response whereas the corresponding reflection hologram continues to produce an increasing response. Figure 11.15b shows the relationship between the optimum grating thickness at which the diffractive response of the transmission hologram peaks and the grating modulation.

A recent photopolymer emulsion from the commercial manufacturer Bayer is characterised by a value of $\delta n \sim 0.037$ and an emulsion thickness of approximately 13.7 μm . This corresponds to a value of $\kappa d = 2.52$ at 633 nm, giving an ideal diffractive efficiency of approximately 97% in reflection mode. Actual measured values indicated 94% [15].

11.10.6 Small Departure from the Bragg Condition

To study the case of a small departure from the Bragg condition, we continue to use Equation 11.92, but now, we relax the condition that $|\mathbf{k}_i| = \beta$. This has the effect that the phases of the contributions of the

signal wave from each Bragg plane do not add up coherently. We shall see that this choice leads naturally to the definition of an “off-Bragg” parameter, which allows us to quantify how much the Bragg condition is violated either in terms of wavelength or in terms of angle.

Proceeding in this fashion, Equation 11.97 remains the same, but Equation 11.98 generalises to

$$\frac{k_{ix}}{\beta} \frac{dS}{dx} + i \left(\frac{\beta^2 - |k_i|^2}{2\beta} \right) S + i\kappa R = 0 \quad (11.108)$$

We now define the “off-Bragg” or “dephasing” parameter

$$\vartheta = \frac{\beta^2 - |k_i|^2}{2\beta} = |\mathbf{G}| \cos(\phi - \theta_c) - \frac{|\mathbf{G}|^2}{2\beta} \quad (11.109)$$

where ϕ represents the slant angle between the grating normal and the grating vector (Figure 11.14). The value of ϑ is determined by the angle of incidence on reconstruction (θ_c) and by the wavelength of the illuminating light ($\lambda_c = 2\pi n/\beta$). Clearly, when $\vartheta = 0$, the Bragg condition is satisfied and $|k_i| = \beta$. We define the obliquity factors:^{*}

$$\begin{aligned} k_{ix}/\beta &= (|k_{ix}|/\beta) \cos \theta_i \equiv c_s \\ k_{cx}/\beta &= \cos \theta_c \equiv c_R \end{aligned} \quad (11.110)$$

Then, as previously, we can solve Equations 11.97 and 11.108 to find expressions for the diffractive efficiency. For the transmission hologram, the result is

$$\eta_T = \frac{\sin^2 \left(\frac{\kappa^2 d^2}{c_R c_S} + \frac{d^2 \vartheta^2}{4c_S^2} \right)^{1/2}}{1 + \frac{\vartheta^2 c_R}{4c_S \kappa^2}} \quad (11.111)$$

And for the reflection hologram,

$$\eta_R = \left\{ 1 + \frac{1 - \frac{\vartheta^2 c_R}{4|c_S| \kappa^2}}{\sinh^2 \left(\frac{\kappa^2 d^2}{c_R |c_S|} - \frac{d^2 \vartheta^2}{4c_S^2} \right)^{1/2}} \right\}^{-1} \quad (11.112)$$

Clearly, for $\vartheta = 0$, these equations revert, respectively, to Equations 11.106 and 11.107. We can better understand the parameter ϑ if we imagine having recorded the grating we are now seeking to play back with an object beam at angle of incidence θ_o and with a reference beam at angle θ_r . The recording wavelength is λ_r , and we take azimuthal angles $\phi_r = \phi_o = 0$. We assume no emulsion shrinkage and no change in average emulsion index. Then, we can write our various wave vectors as the following two vectors:[†]

^{*} See Equations 11.113 and 11.114.

[†] Note the simple change of sign from our previous discussions (see Equation 11.59) in which the reference beam came from the right instead of the left.

$$\mathbf{k}_r = \frac{2\pi n}{\lambda_r} \begin{pmatrix} \cos \theta_r \\ \sin \theta_r \end{pmatrix}; \quad \mathbf{k}_o = \frac{2\pi n}{\lambda_r} \begin{pmatrix} \cos \theta_o \\ \sin \theta_o \end{pmatrix} \quad (11.113)$$

$$\mathbf{k}_c = \frac{2\pi n}{\lambda_c} \begin{pmatrix} \cos \theta_c \\ \sin \theta_c \end{pmatrix}; \quad \mathbf{k}_i = |k_i| \begin{pmatrix} \cos \theta_i \\ \sin \theta_i \end{pmatrix} \quad (11.114)$$

Equation 11.109 can now be written as

$$\vartheta = \frac{2\pi n}{\lambda_r} \{ \cos(\theta_r - \theta_c) - \cos(\theta_o - \theta_c) \} - \frac{2\pi n \lambda_c}{\lambda_r^2} \{ 1 - \cos(\theta_r - \theta_o) \} \quad (11.115)$$

This tells us how the parameter ϑ behaves when $\lambda_c \neq \lambda_r$ and when $\theta_c \neq \theta_r$. Direct substitution of Equation 11.115 into Equations 11.111 and 11.112 leads trivially to general expressions for the diffractive response of a lossless holographic grating recorded with parameters $(\theta_r, \theta_o, \lambda_r)$ and replayed with (θ_c, λ_c) .* These expressions are of prime importance for estimating the diffractive response of a grating.

11.10.6.1 Behaviour of the Lossless Transmission Hologram

It is instructive to study separately the effect of hologram replay at a differing wavelength and at a differing angle. We shall start with differing angles. To simplify things, we shall take the case of an unslanted grating and set $\theta_r = -\theta_o$. To eliminate the contribution in ϑ due to wavelength, we set $\lambda_c = \lambda_r$, whereupon Equation 11.115 becomes

$$\vartheta = \frac{4\pi n}{\lambda_r} \sin \theta_r \{ \sin \theta_c - \sin \theta_r \} \quad (11.116)$$

The obliquity factors for the unslanted transmission grating are

$$c_R = \cos \theta_c \quad (11.117)$$

and

$$c_S = \frac{|k_i|}{\beta} \cos \theta_i = \cos \theta_c - \frac{\lambda_c}{\lambda_r} (\cos \theta_r - \cos \theta_o) = \cos \theta_c \quad (11.118)$$

Substitution of Equations 11.116 through 11.118 into 11.111 then yields

$$\eta_T = \frac{\sin^2 \left[\left(1 + \frac{4 \sin^2 \theta_r \{ \sin \theta_c - \sin \theta_r \}^2}{(\delta n/n)^2} \right)^{1/2} \left(\frac{\delta n}{n} \right) \frac{\pi}{\cos \theta_c} \frac{nd}{\lambda_r} \right]}{1 + \frac{4 \sin^2 \theta_r \{ \sin \theta_c - \sin \theta_r \}^2}{(\delta n/n)^2}} \quad (11.119)$$

This describes the diffractive response of the hologram in terms of the modulation $\delta n/n$, the record and illumination angles, θ_r and θ_c , and the normalised grating thickness, nd/λ_r . We can make one further useful simplification to this formula. The diffractive efficiency of a transmission hologram illuminated under perfect Bragg compliance is a maximum when

* One also needs Equation 11.110.

$$\left(\frac{\delta n}{n} \right) \frac{\pi}{\cos \theta_r} \frac{nd}{\lambda_r} = \frac{\pi}{2} \quad (11.120)$$

This was described in Figure 11.15. We can use this formula to define the thickness of the hologram. If we do this, we arrive at an expression for the diffractive response of the optimally thick transmission hologram:

$$\hat{\eta}_T = \frac{\sin^2 \left[\left(1 + \frac{4 \sin^2 \theta_r \{ \sin \theta_c - \sin \theta_r \}^2}{(\delta n/n)^2} \right)^{1/2} \frac{\pi \cos \theta_r}{2 \cos \theta_c} \right]}{1 + \frac{4 \sin^2 \theta_r \{ \sin \theta_c - \sin \theta_r \}^2}{(\delta n/n)^2}} \quad (11.121)$$

The function $\hat{\eta}_T(\theta_c)$ is plotted out in Figure 11.16a through 11.16c for various values of θ_c and for three cases of typical modulations. For a modulation characteristic of contemporary silver halide materials employed in display holography and at an internal recording angle of $\theta_r = 30^\circ$ (equivalent to an external angle of approximately 45° at typical index values), we see that the full-width half maximum (FWHM) of $\hat{\eta}_T(\theta_c)$ is approximately 3° . We should not, however, expect our results here to be accurate as θ_c becomes too small. In the unslanted transmission grating the Bragg condition will in general be satisfied for $\theta_c = \pm \theta_r$. However we have implicitly assumed that only the solution $\theta_c = \theta_r$ corresponds to perfect Bragg compliance. We are of course at liberty to choose the other root. In this case we would define $\mathbf{k}_c - \mathbf{k}_i = -\mathbf{G}$ instead of $\mathbf{k}_c - \mathbf{k}_i = +\mathbf{G}$. When θ_c is relatively large there is a large angular spacing between the two perfect Bragg angles. As such, for all but the thinnest gratings, only one root produces significant diffraction at a given angle. If we were to replot Figure 11.16b for the case of $\theta_r = 10^\circ$ assuming $\mathbf{k}_c - \mathbf{k}_i = -\mathbf{G}$ then we would see a peak value of η of approximately 0.1 which is about 10% of the value plotted. The corresponding graph at $\theta_r = 30^\circ$ however shows negligible diffraction for the alternate solution. As $\theta_c \rightarrow 0$, the assumption of a single diffracted mode breaks down. Indeed at $\theta_c = 0$ we encounter a fundamentally new type of diffraction known as Raman-Nath diffraction.

To investigate the effect of fringe slant on the diffractive replay, in Figure 11.17a, we plot $\eta_T(\theta_c)$ of Equation 11.111 for various values of the object beam recording angle θ_o for a very specific case ($d = 7 \mu\text{m}$, $n = 1.5$, $\delta n/n = 0.02$, $\lambda_r = \lambda_c = 532 \text{ nm}$, $\theta_r = 30^\circ$). Clearly, there is very little change in the diffractive response of the hologram as the slant of the grating changes over the normal range used in display holography. However, one does see a broadening of the curve, indicating a decrease in angle selectivity of the grating, as the recording and reference beams come closer together.

At this point, it is interesting to expand our discussion of the coupled wave theory to the case of the π -polarisation. You will remember that we have thus far restricted ourselves to the σ -polarisation. In fact, Kogelnik and others have shown [1,4] that the only difference to the coupled wave equations for the σ -polarisation is that, for the π -polarisation, the constant κ must be replaced by

$$\hat{\kappa} = -\kappa \cos 2(\theta_c - \phi) \quad (11.122)$$

where ϕ is the grating slant angle defined by

$$\phi = \tan^{-1} \frac{G_y}{G_x} = \tan^{-1} \left\{ \frac{\sin \theta_r - \sin \theta_o}{\cos \theta_r - \cos \theta_o} \right\} \quad (11.123)$$

Proving this result using the coupled wave theory is a little involved, and as such, the interested reader is referred to the references given.* However, we can easily use this to re-plot Figure 11.17a for the case of the π -polarisation. This is shown in Figure 11.17b. The graph is different from the σ -polarisation case because the diffractive efficiency drops somewhat as the object and reference beams become more

* We shall, however, use the PSM theory in Chapter 12 to prove this result. Whereas coupled wave theory is substantially easier and clearer to apply to the σ -polarisation, the PSM theory can be used to treat either polarisation with equal ease.

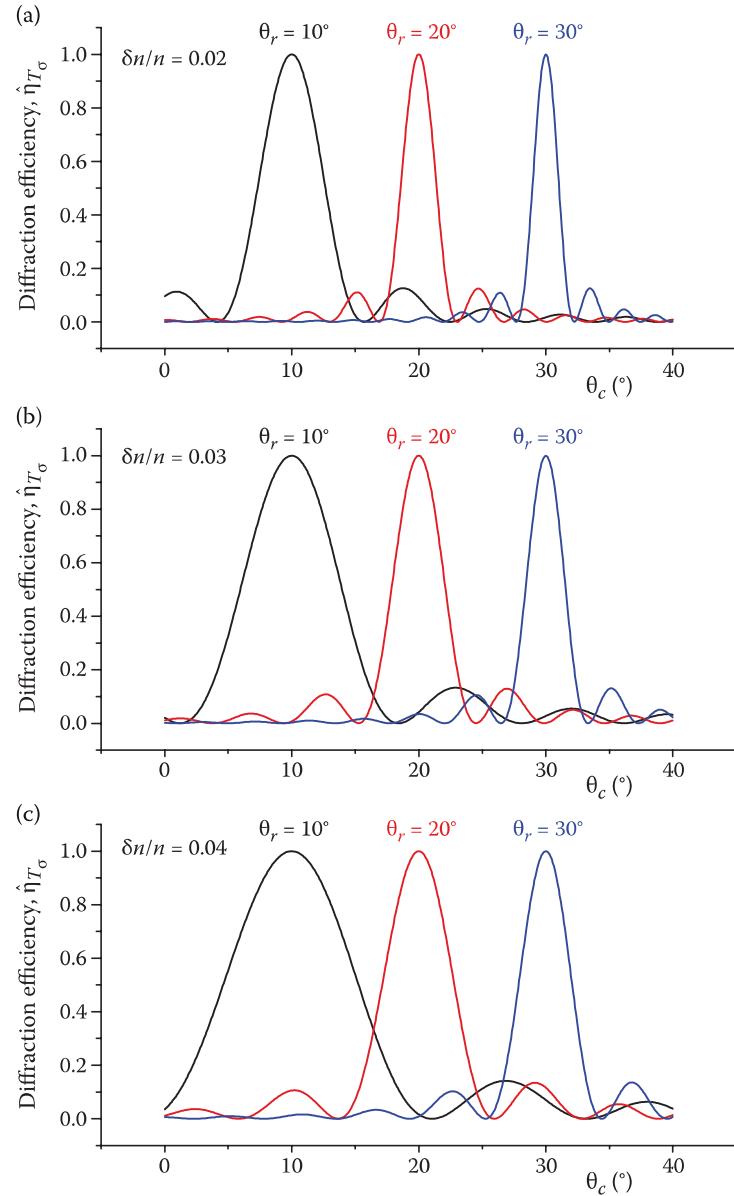


FIGURE 11.16 Optimal diffractive efficiency of the lossless unslanted transmission grating (σ -polarisation) as predicted by the coupled wave theory. The function $\hat{\eta}_T(\theta_c)$ is plotted for three values of $\theta_r = 10^\circ, 20^\circ, 30^\circ$ and for three cases of typical modulations: (a) $\delta n/n = 0.02$; (b) $\delta n/n = 0.03$; (c) $\delta n/n = 0.04$. Note that all angles are within the dielectric.

angularly spaced during recording. Clearly, from Equations 11.111 and 11.122, we can expect the diffractive efficiency to drop to zero as the difference in angle at recording tends to 90° . Intuitively, we can understand this by the fact that when $\theta_c = \theta_r$ and $|\theta_o - \theta_r| = \pi/2$, we are essentially asking the hologram to reconstruct a signal beam having an electric field vector in a direction in which the illuminating beam has zero electric field.

Returning to the σ -polarisation, we now investigate the behaviour of the lossless transmission hologram to variation of the replay wavelength by setting $\theta_c = \theta_r$. As before, we will first assume that the grating is unslanted so that $\theta_r = -\theta_o$. Equation 11.115 then becomes

$$\vartheta = \frac{2\pi n}{\lambda_r} \{1 - \cos(2\theta_r)\} \left(1 - \frac{\lambda_c}{\lambda_r}\right) \quad (11.124)$$

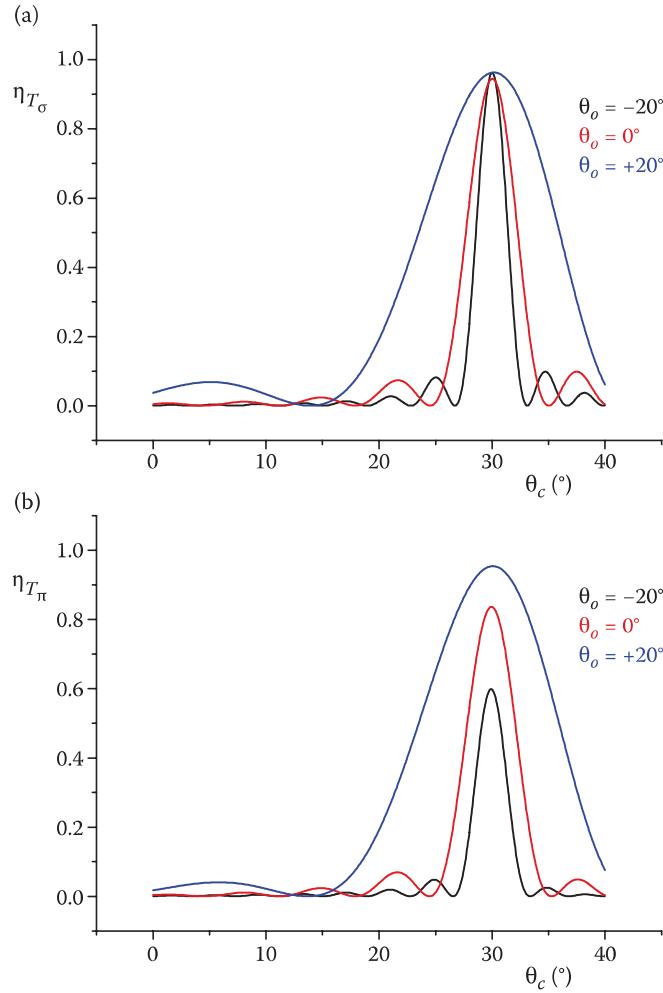


FIGURE 11.17 Diffractive efficiency versus replay angle, $\eta_T(\theta_c)$ as predicted by the coupled wave theory of the lossless transmission hologram for various values of the object beam recording angle $\theta_o = -20^\circ, 0^\circ, +20^\circ$ and for the grating parameters $d = 7 \mu\text{m}$, $n = 1.5$, $\delta n/n = 0.02$; $\lambda_r = \lambda_c = 532 \text{ nm}$, $\theta_r = 30^\circ$: (a) σ -polarisation and (b) π -polarisation. Angles quoted are within the dielectric.

Substituting this, Equation 11.117 and Equation 11.118 into Equation 11.111 yields

$$\eta_T = \frac{\sin^2 \left\{ \left(1 + \frac{\frac{\lambda_c^2}{\lambda_r^2} \{1 - \cos(2\theta_r)\}^2 \left(1 - \frac{\lambda_c}{\lambda_r} \right)^2}{(\delta n/n)^2} \right)^{1/2} \frac{\pi}{\cos \theta_r} \left(\frac{\delta n}{n} \right) n d \right\}}{\left(1 + \frac{\frac{\lambda_c^2}{\lambda_r^2} \{1 - \cos(2\theta_r)\}^2 \left(1 - \frac{\lambda_c}{\lambda_r} \right)^2}{(\delta n/n)^2} \right)} \quad (11.125)$$

Using Equation 11.120, we can then write down an expression for the diffractive response versus replay wavelength for the optimally thick transmission hologram:

$$\hat{\eta}_T = \frac{\sin^2 \left\{ \left[1 + \frac{\frac{\lambda_c^2}{\lambda_r^2} \{1 - \cos(2\theta_r)\}^2 \left(1 - \frac{\lambda_c}{\lambda_r}\right)^2}{(\delta n/n)^2} \right]^{1/2} \frac{\pi \lambda_r}{2 \lambda_c} \right\}}{\left\{ 1 + \frac{\frac{\lambda_c^2}{\lambda_r^2} \{1 - \cos(2\theta_r)\}^2 \left(1 - \frac{\lambda_c}{\lambda_r}\right)^2}{(\delta n/n)^2} \right\}} \quad (11.126)$$

The function $\hat{\eta}_T(\lambda_c)$ is plotted out in Figure 11.18a through 11.18c for the same values of θ_r and $\delta n/n$ as used in Figure 11.16 and for three popular recording wavelengths. Clearly, the replay characteristics of the unslanted transmission grating are rather insensitive to replay wavelength. At modulations characteristic of contemporary silver halide materials used in display holography and at an internal recording angle of $\theta_r = -\theta_o = 30^\circ$ (equivalent to an external angle of approximately 45° at typical index values), we see that the FWHM of $\hat{\eta}_T(\lambda_c)$ is well over 60 nm, and at values of $\theta_r = -\theta_o = 10^\circ$, the FWHM is greater than the entire visible spectrum.

To get an understanding of the effect of fringe slant on the diffractive replay, we will again choose a specific case. Figure 11.19a shows a plot of $\eta_T(\lambda_c)$ calculated directly from Equation 11.111 for three values of the object beam recording angle, $\theta_o = (-20^\circ, 0^\circ, +20^\circ)$ and for the following parameters: $d = 7 \mu\text{m}$, $n = 1.5$, $\delta n/n = 0.02$, $\theta_r = \theta_c = 30^\circ$. This is an interesting graph because it shows an effective blue shift on replay of the grating recorded at $\theta_r = 30^\circ$, $\theta_o = 20^\circ$, $\lambda_r = 532 \text{ nm}$. The Bragg condition is however exactly satisfied at $\theta_c = 30^\circ$, $\lambda_c = 532 \text{ nm}$ where $\vartheta = 0$, but this is not the peak of diffractive efficiency! This is because the reference and object beam are very close together during recording, leading to a large fringe spacing and a very broadband replay characteristic. As such, the diffractive efficiency depends more on $\kappa(\lambda_c)$ than on $\vartheta(\lambda_c)$; the dominant physical effect is therefore that a blue replay wavelength actually “feels” the grating to be thicker than the recording green wavelength and therefore exhibits a higher diffractive response. This is an effect of the thin grating only. As grating thickness is increased (with a proportionate decrease in modulation), the effect disappears. Figure 11.19b shows the corresponding case for the π -polarisation, which shows a lessening diffractive response when the angle between the electric field vectors of the diffracted and illuminating ray approaches 90° .

Such apparent violation, and it is of course only apparent, of the Bragg condition has a direct bearing on our prior discussion of the geometrical ray equations describing how a holographic image distorts when replayed. There, we implicitly assumed that the maximum diffractive ray was indeed to be associated with the ray that exactly satisfied the Bragg condition. However, we see here that for certain cases in moderately thin gratings, notably when the replay behaviour is extremely broadband, this assumption starts to break down.

11.10.6.2 Behaviour of the Lossless Reflection Hologram

We can usefully characterise the lossless reflection hologram in a very similar manner to the way we have proceeded in the previous section. We will start by restricting the discussion to the unslanted reflection grating for which $\theta_r = \pi - \theta_o$. Then, as previously, we will study the angle dependence by setting $\lambda_c = \lambda_r$. The obliquity factor c_S is still given by (11.117), but the form of c_S changes to

$$c_S = \cos \theta_c - 2 \cos \theta_r \quad (11.127)$$

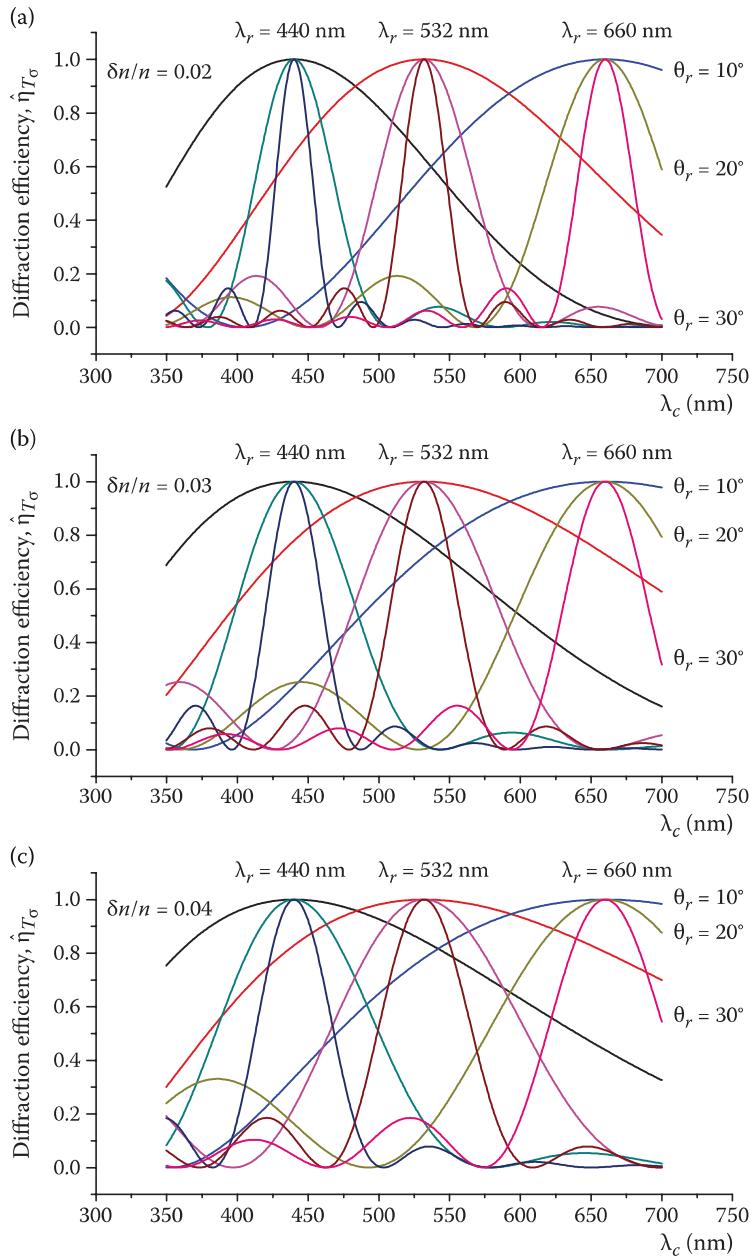


FIGURE 11.18 Optimal diffractive efficiency versus free-space replay wavelength, $\hat{\eta}_T(\lambda_c)$ of the lossless unslanted transmission grating (σ -polarisation) as predicted by the coupled wave theory for three values of $\theta_r = 10^\circ, 20^\circ, 30^\circ$; three recording wavelengths, $\lambda_r = 440 \text{ nm}$, $\lambda_r = 532 \text{ nm}$ and $\lambda_r = 660 \text{ nm}$; and for three cases of typical modulations, (a) $\delta n/n = 0.02$, (b) $\delta n/n = 0.03$, (c) $\delta n/n = 0.04$. Angles quoted are within the dielectric.

The parameter ϑ also changes:

$$\vartheta = \frac{4\pi n}{\lambda_r} \cos \theta_r (\cos \theta_c - \cos \theta_r) \quad (11.128)$$

With these formulae, we can now rewrite Equation 11.112 to give a formula for the diffraction efficiency of the lossless unslanted reflection grating in terms of replay angle:

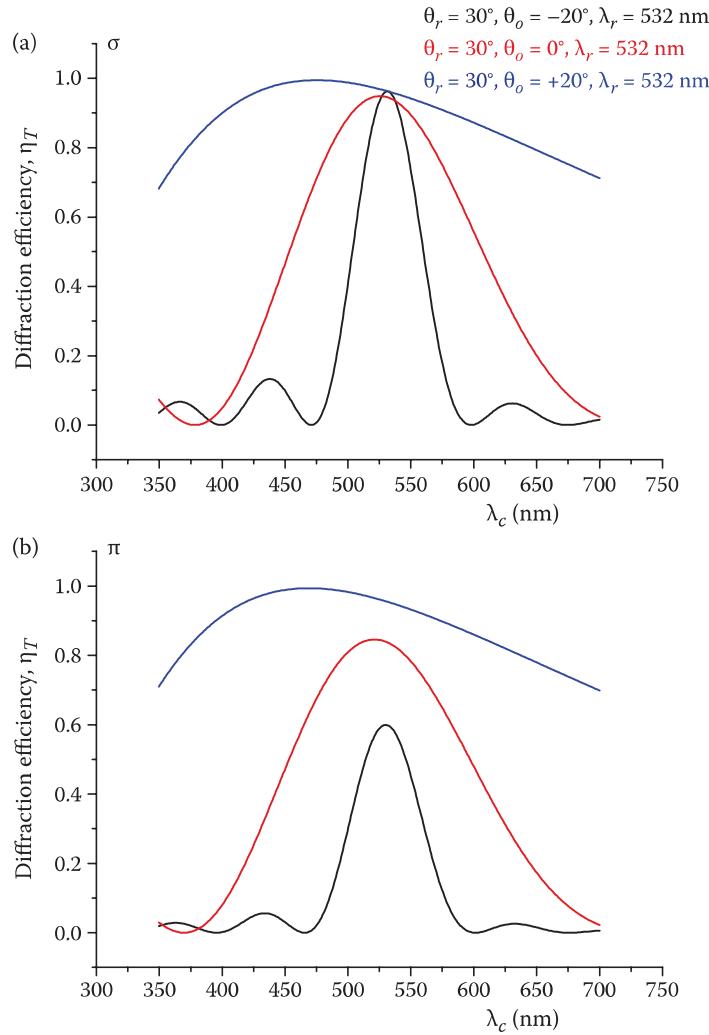


FIGURE 11.19 Diffractive efficiency versus free-space replay wavelength, $\eta_T(\lambda_c)$ as predicted by the coupled wave theory of the lossless transmission hologram for three values of the object beam recording angle, $\theta_o = (-20^\circ, 0^\circ, +20^\circ)$ and for the following parameters: $d = 7 \mu\text{m}$, $n = 1.5$, $\delta n/n = 0.02$, $\lambda_r = 532 \text{ nm}$, $\theta_r = \theta_c = 30^\circ$): (a) σ -polarisation and (b) π -polarisation. Angles quoted are within the dielectric.

$$\eta_R = \frac{\sinh^2 \left\{ \Upsilon \frac{\kappa d}{\sqrt{c_R |c_S|}} \right\}}{\sinh^2 \left\{ \Upsilon \frac{\kappa d}{\sqrt{c_R |c_S|}} \right\} + \Upsilon^2} \quad (11.129)$$

where

$$\Upsilon = \left(1 - \frac{4(\cos \theta_c - \cos \theta_r)^2 \cos \theta_c \cos^2 \theta_r}{|(\cos \theta_c - 2 \cos \theta_r)(\delta n/n)^2|} \right)^{1/2} \quad (11.130)$$

and

$$\frac{\kappa d}{\sqrt{c_R |c_S|}} = \frac{\pi}{\sqrt{|(\cos \theta_c - 2 \cos \theta_r)| \cos \theta_c}} \left(\frac{nd}{\lambda_r} \right) \left(\frac{\delta n}{n} \right) \quad (11.131)$$

Let us now define the thickness of the hologram to give a 90% diffractive efficiency at $\theta_c = \theta_r$:

$$\pi \left(\frac{nd}{\lambda_r} \right) \left(\frac{\delta n}{n} \right) = \cos \theta_r \tanh^{-1} \left\{ \sqrt{0.9} \right\} \quad (11.132)$$

We then arrive at the formula for the diffractive efficiency versus angle $\hat{\eta}_{R90}(\theta_c)$ for a 90% optimised lossless unslanted reflection grating:

$$\hat{\eta}_{R90} = \frac{\sinh^2 \left\{ \Upsilon \frac{\cos \theta_r \tanh^{-1} \left\{ \sqrt{0.9} \right\}}{\sqrt{|(\cos \theta_c - 2 \cos \theta_r)| \cos \theta_c}} \right\}}{\sinh^2 \left\{ \Upsilon \frac{\cos \theta_r \tanh^{-1} \left\{ \sqrt{0.9} \right\}}{\sqrt{|(\cos \theta_c - 2 \cos \theta_r)| \cos \theta_c}} \right\} + \Upsilon^2} \quad (11.133)$$

The function $\hat{\eta}_{R90}(\theta_c)$ is plotted out in Figure 11.20a through 11.20c for various values of θ_r and for three cases of typical modulations. It is interesting to note that, generally, the reflection hologram is characterised by slightly poorer replay angle discrimination than the transmission hologram.

To investigate the importance of fringe slant, in Figure 11.21a, we plot $\eta_R(\theta_c)$ of Equation 11.112 for three object beam recording angles ($\theta_o = 160^\circ, 180^\circ, 200^\circ$) and for the parameters ($d = 7 \mu\text{m}$, $n = 1.5$, $\delta n/n = 0.03$, $\lambda_r = \lambda_c = 532 \text{ nm}$, $\theta_r = 30^\circ$). For completeness, we plot the corresponding graph for the π -polarisation in Figure 11.21b. You will notice that the curves exhibit two peaks corresponding to a reflection of \mathbf{k}_c about the grating vector \mathbf{G} . For instance, when $\theta_o = 160^\circ$, the fringes are slanted at $\phi = 5^\circ$ and the angle between \mathbf{k}_r and \mathbf{G} is 25° . One would therefore expect a second peak at a replay angle of -20° because at this angle, the angle between \mathbf{k}_c and \mathbf{G} is again 25° . The two peaks for $\theta_o = 200^\circ$ are close enough together that they merge into a single wide peak.

Finally, we come to the wavelength dependence of the lossless reflection grating. To study this, we first assume an unslanted grating and put $\theta_o = \pi - \theta_r$ and $\theta_c = \theta_r \geq 0$ with $\lambda_c \neq \lambda_r$. Following our previous work, we then arrive at a formula for the diffractive efficiency versus wavelength, $\hat{\eta}_{R90}(\lambda_c)$ for a 90% optimised lossless unslanted reflection grating:

$$\hat{\eta}_{R90} = \sinh^2 \left\{ \Upsilon \frac{\frac{\lambda_r}{\lambda_c} \tanh^{-1} \left\{ \sqrt{0.9} \right\}}{\sqrt{1 - 2 \frac{\lambda_c}{\lambda_r}}} \right\} \left[\sinh^2 \left\{ \Upsilon \frac{\frac{\lambda_r}{\lambda_c} \tanh^{-1} \left\{ \sqrt{0.9} \right\}}{\sqrt{1 - 2 \frac{\lambda_c}{\lambda_r}}} \right\} + \Upsilon^2 \right]^{-1} \quad (11.134)$$

where

$$\Upsilon = \left\{ 1 - \frac{4 \left(\frac{\lambda_c}{\lambda_r} \right)^2 \cos^4 \theta_r \left[1 - \frac{\lambda_c}{\lambda_r} \right]^2}{\left(1 - 2 \frac{\lambda_c}{\lambda_r} \right) \left(\frac{\delta n}{n} \right)^2} \right\}^{1/2} \quad (11.135)$$

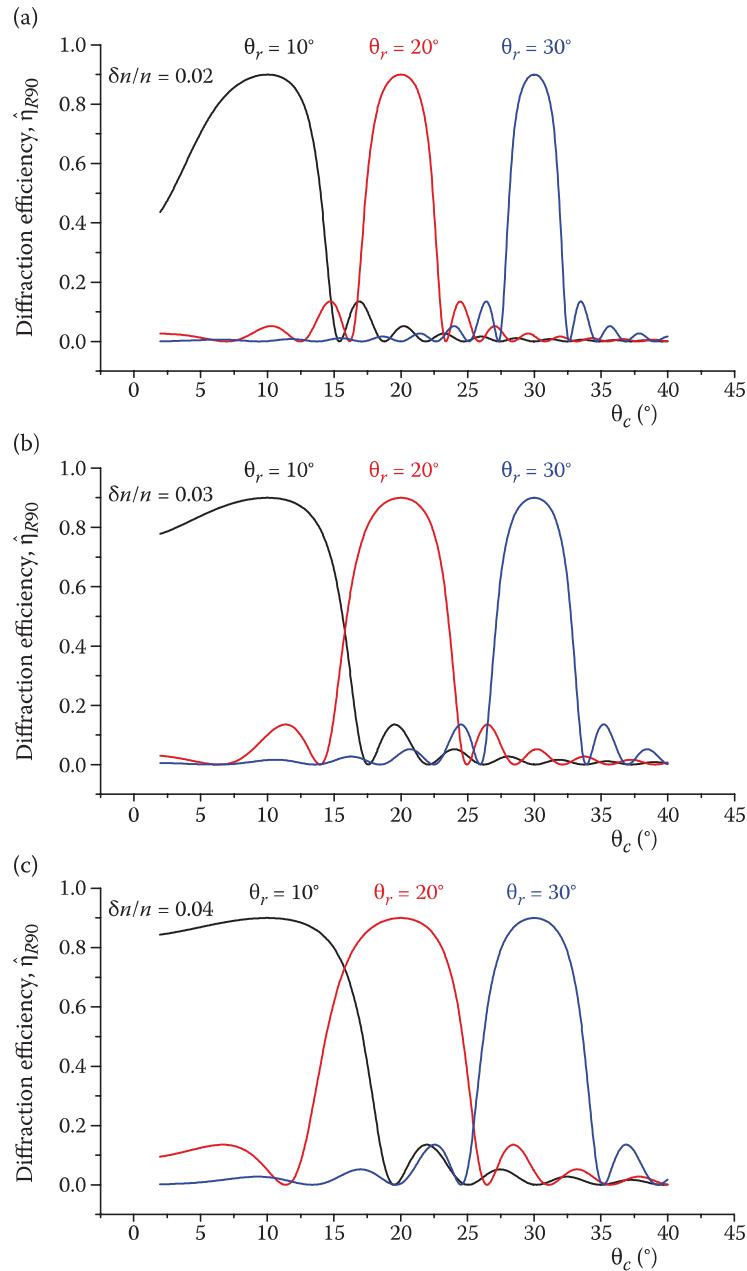


FIGURE 11.20 Optimal diffractive efficiency versus replay angle, $\hat{\eta}_{R90}(\theta_c)$ of the lossless unslanted reflection grating (σ -polarisation) as predicted by the coupled wave theory for three values of recording angle $\theta_r = 10^\circ, 20^\circ, 30^\circ$ and for three cases of typical modulations: (a) $\delta n/n = 0.02$ (corresponding to a grating thickness of $10.1 \mu\text{m}$ for $\theta_r = 10^\circ$, $9.65 \mu\text{m}$ for $\theta_r = 20^\circ$ and $8.89 \mu\text{m}$ for $\theta_r = 30^\circ$); (b) $\delta n/n = 0.03$ (corresponding to a grating thickness of $6.74 \mu\text{m}$ for $\theta_r = 10^\circ$, $6.43 \mu\text{m}$ for $\theta_r = 20^\circ$ and $5.93 \mu\text{m}$ for $\theta_r = 30^\circ$); (c) $\delta n/n = 0.04$ (corresponding to a grating thickness of $5.05 \mu\text{m}$ for $\theta_r = 10^\circ$, $4.82 \mu\text{m}$ for $\theta_r = 20^\circ$ and $4.44 \mu\text{m}$ for $\theta_r = 30^\circ$). The function $\hat{\eta}_{R90}(\theta_c)$ is chosen by defining the product of normalised grating thickness and index modulation such that $\hat{\eta}_{R90}(\theta_r) = 0.9$ for each value of θ_r plotted. All angles quoted are within the dielectric.

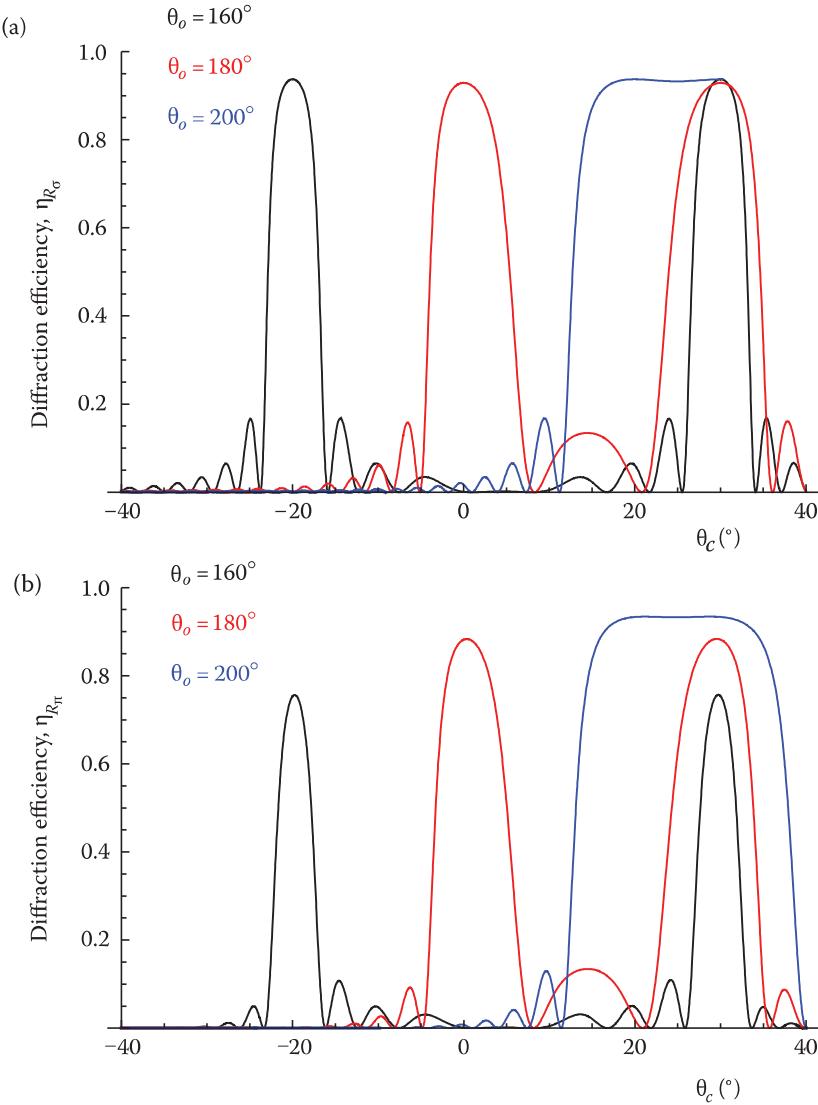


FIGURE 11.21 Diffractive efficiency versus replay angle, $\eta_R(\theta_c)$, as predicted by the coupled wave theory of the lossless reflection hologram for various values of the object beam recording angle $\theta_o = 160^\circ, 180^\circ, 200^\circ$ and for the grating parameters $d = 7 \mu\text{m}$, $n = 1.5$, $\delta n/n = 0.03$; $\lambda_r = \lambda_c = 532 \text{ nm}$, $\theta_r = 30^\circ$: (a) σ -polarisation and (b) π -polarisation. All angles quoted are within the dielectric.

The function $\hat{\eta}_{R90}(\lambda_c)$ is plotted out in Figure 11.22a through 11.22c for the same values of θ_r and $\delta n/n$ as used in Figure 11.18 for the transmission case. Clearly, and as expected from our previous discussions, the unslanted reflection grating is far more sensitive to replay wavelength than its transmission counterpart. In addition, the replay bandwidth is much less sensitive to recording angle.

Figure 11.23a and b show the effect of fringe slant on the replay wavelength dependence of the reflection hologram for the σ - and π -polarisations, respectively. The graphs are plotted for a typical silver halide emulsion grating thickness of $7 \mu\text{m}$ and an index modulation of $\delta n/n = 0.03$, giving a maximum diffraction efficiency of more than 90% and a typical FWHM replay bandwidth of 25 nm.

It is worthwhile spending a little time trying to understand the basic reasons, in terms of the coupled wave theory, behind the angle and wavelength behaviour of the transmission and reflection holograms. To do this, we will assume that the illumination wave on playback is of magnitude $|k_i| = 2\pi n/\lambda_r + \Delta\beta$ and that its angle of incidence is $\theta_c = \theta_r + \Delta\theta$. Then, Equations 11.92, 11.109, 11.113 and 11.114 permit us to write down the following simple expression, which relates θ to $\Delta\theta_c$ and $\Delta\beta$:

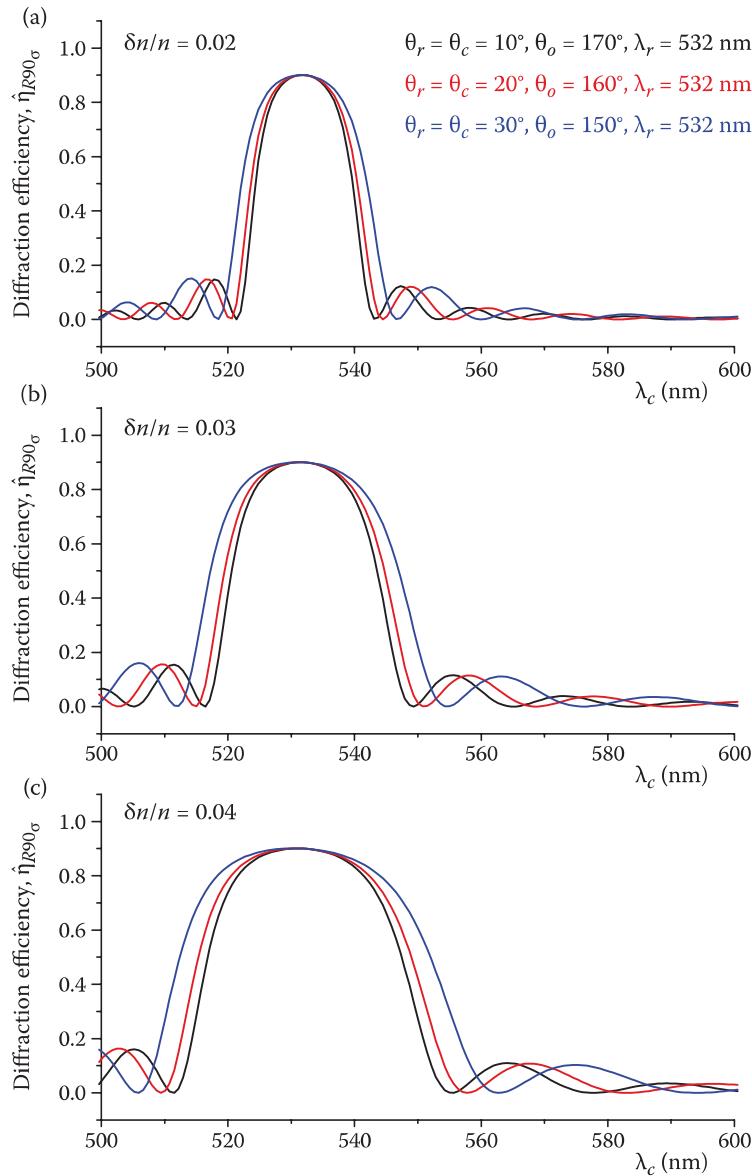


FIGURE 11.22 Optimal diffractive efficiency versus (freespace) replay wavelength, $\hat{\eta}_{R90}(\lambda_c)$, of the lossless unslanted reflection grating (σ -polarisation) as predicted by the coupled wave theory for three values of $\theta_r = 10^\circ, 20^\circ, 30^\circ$, a recording wavelength of $\lambda_r = 532 \text{ nm}$ and for three cases of typical modulations: (a) $\delta n/n = 0.02$, (b) $\delta n/n = 0.03$ and (c) $\delta n/n = 0.04$. The function $\hat{\eta}_{R90}(\lambda_c)$ is defined by choosing the product of the normalised grating thickness and index modulation, such that $\hat{\eta}_{R90}(\lambda_r, \theta_r) = 0.9$ for each case of (θ_r, λ_r) plotted. Assuming an average index of $n = 1.5$, case (a) corresponds to a grating thickness of $d = 10.1 \mu\text{m}$ for $\theta_r = 10^\circ$, $d = 9.65 \mu\text{m}$ for $\theta_r = 20^\circ$ and $d = 8.89 \mu\text{m}$ for $\theta_r = 30^\circ$; likewise, case (b) corresponds to a grating thickness of $d = 6.74 \mu\text{m}$ for $\theta_r = 10^\circ$, $d = 6.43 \mu\text{m}$ for $\theta_r = 20^\circ$ and $d = 5.93 \mu\text{m}$ for $\theta_r = 30^\circ$; and case (c) corresponds to a grating thickness of $d = 5.05 \mu\text{m}$ for $\theta_r = 10^\circ$, $d = 4.82 \mu\text{m}$ for $\theta_r = 20^\circ$ and $d = 4.44 \mu\text{m}$ for $\theta_r = 30^\circ$. All angles internal.

$$\vartheta = \frac{2\pi n}{\lambda_r} \Delta\theta \sin(\theta_r - \theta_o) + \Delta\beta \{1 - \cos(\theta_r - \theta_o)\} \quad (11.136)$$

Now let us adopt a value of $\kappa d / \sqrt{|c_R| c_S|} = \pi/2$. You will recall that this gives us perfect conversion from the R wave to the S wave in the transmission hologram when $\vartheta = 0$. It also corresponds to a diffractive efficiency for the reflection hologram of 0.84. We use Equations 11.111 and 11.112 to calculate the value of the dephasing parameter ϑ that is required to bring the diffraction to its first zero. This is given by

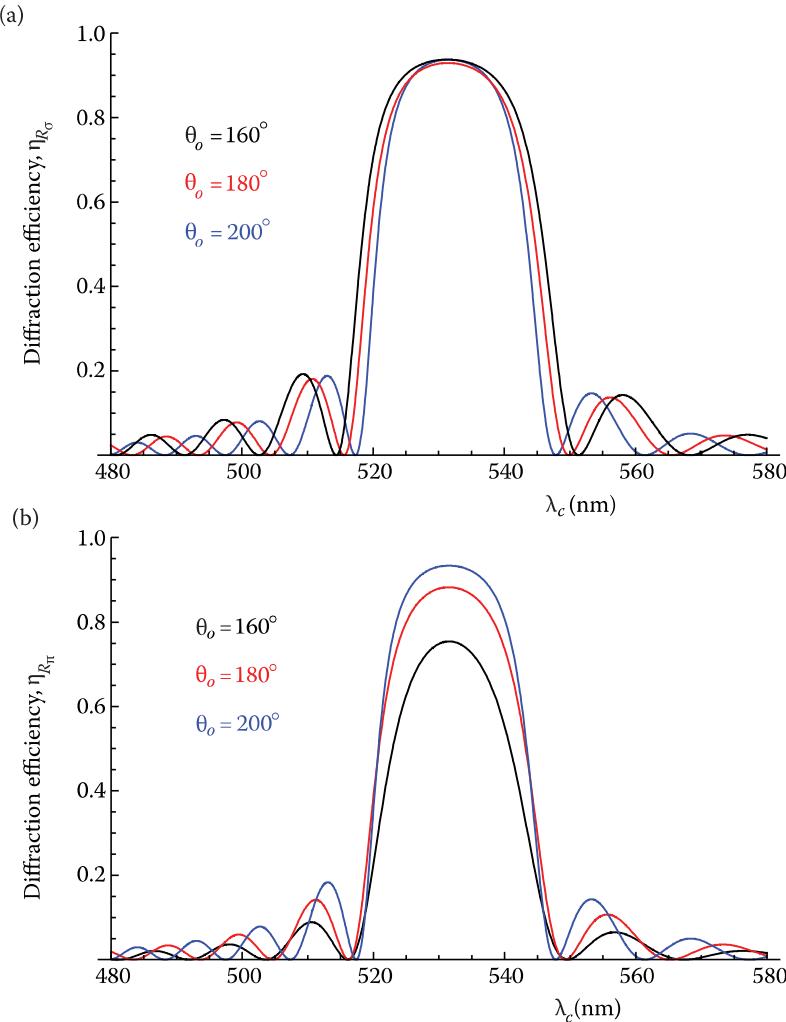


FIGURE 11.23 Diffractive efficiency versus (freespace) replay wavelength, $\eta_R(\lambda_c)$, as predicted by the coupled wave theory of the lossless reflection hologram for three values of the object beam recording angle, $\theta_o = (160^\circ, 180^\circ, 200^\circ)$ and for the following parameters: $d = 7 \mu\text{m}$, $n = 1.5$, $\delta n/n = 0.03$, $\lambda_r = 532$, $\theta_r = \theta_c = 30^\circ$: (a) σ -polarisation and (b) π -polarisation. Angles quoted are internal.

$$\vartheta_T = \sqrt{3}\pi \frac{c_s}{d} \quad (11.137)$$

$$\vartheta_R = \sqrt{5}\pi \frac{|c_s|}{d}$$

We may then use Equation 11.136 to show that for the unslanted transmission hologram,*

$$\Delta\theta_T \sim \frac{\sqrt{3}}{2} \frac{\Lambda}{d} = \frac{\sqrt{3}}{4} \frac{\lambda}{dn} \csc \theta_r \quad (11.138)$$

$$\left(\frac{\Delta\lambda}{\lambda} \right)_T \sim \frac{\sqrt{3}}{2} \frac{\Lambda}{d} \cot \theta_r = \frac{\sqrt{3}}{4} \frac{\lambda}{dn} \cos \theta_r \csc^2 \theta_r \quad (11.139)$$

* Note that Kogelnik [1] gives the following formulae for the FWHM: $\Delta\theta_{FWHM} = \Lambda/d$; $\Delta\lambda_{FWHM} = \cot \theta_r \cdot \Lambda/d$.

and for the corresponding reflection hologram,

$$\Delta\theta_R \sim \frac{\sqrt{5}}{2} \frac{\Lambda \cot \theta_r}{d} = \frac{\sqrt{5}}{4} \frac{\lambda}{dn} \csc \theta_r \quad (11.140)$$

$$\left(\frac{\Delta\lambda}{\lambda} \right)_R = \frac{\sqrt{5}}{2} \frac{\Lambda}{d} = \frac{\sqrt{5}}{4} \frac{\lambda \sec \theta_r}{dn} \quad (11.141)$$

We can now see why a transmission hologram is generally more selective in angle than a reflection hologram: $\Delta\theta_R/\Delta\theta_T = \sqrt{5}/3$, independent of wavelength and angle! Similarly, $\Delta\lambda_R/\Delta\lambda_T \sim \sqrt{5}/3 \tan^2 \theta_r$, which for small θ_r makes the reflection hologram much more selective than the corresponding transmission case. Note that the expressions in Equations 11.139 and 11.141 are remarkably close to the estimations of the PSM model that led us to Equations 11.77 and 11.79.

11.10.7 Effect of Loss in the Dielectric

Up until now, we have assumed that the electrical conductivity of the dielectric grating was zero. This led us to a form of the Helmholtz equation with real coefficients. It is, however, very straightforward to generalise the coupled wave theory to the case of a complex coefficient. To do this, we introduce the following companion equation to Equation 11.88:

$$\sigma = \sigma_0 + \sigma_1 \cos \mathbf{G} \cdot \mathbf{r} \quad (11.142)$$

Here, σ is the conductivity of the grating, which we divide into a constant and modulated part just as we did with the permittivity. If we then introduce the parameter

$$\alpha = \frac{\sigma_0}{2} \sqrt{\frac{\mu}{\epsilon_0 \epsilon_1}} \quad (11.143)$$

we can generalise Equation 11.89 to

$$\gamma^2 = -\beta^2 + 2i\alpha\beta - 4\kappa\beta \cos \mathbf{G} \cdot \mathbf{r} \quad (11.144)$$

where Kogelnik's coupling constant now becomes

$$\kappa = \frac{1}{4} \frac{\epsilon_{r1}}{\epsilon_{r0}} \beta + i \frac{\alpha}{2} \frac{\sigma_1}{\sigma_0} \quad (11.145)$$

It is then a simple matter to solve the Helmholtz equation exactly as before by setting up two coupled waves, R and S , which are then described by the differential equations

$$c_R \frac{dR}{dx} + \alpha R + i\kappa S = 0 \quad (11.146)$$

$$c_S \frac{dS}{dx} + (\alpha + i\vartheta) S + i\kappa R = 0 \quad (11.147)$$

Using these equations, we can see that the energy balance of the lossy coupled wave model is now described by

$$\frac{d}{dx} (c_R R R^* + c_S S S^*) + 2\alpha(RR^* + SS^*) + i(\kappa - \kappa^*)(RS^* + R^*S) = 0 \quad (11.148)$$

The presence of the obliquity factors in the first term indicates that power flows along the x axis. The second and third terms correspond to the expected ohmic heating, $\sigma E E^*$.

Equations 11.146 and 11.147 can be solved to find (rather complex) analytic expressions for S and R for both the transmission and reflection boundary conditions. We shall simply summarise the most

important results here that pertain to unslanted mixed-phase amplitude gratings. In the case of the transmission hologram, the angular selectivity is essentially independent of loss. Kogelnik suggests that this is because absorption does not really alter the phase relationship between R and S . The only practical effect of including absorption then is to reduce the diffractive efficiency of the hologram as one would naively expect.

The effect of absorption on the unslanted reflection hologram is perhaps a little more noticeable. Here, loss tends to broaden the $\eta(\theta d)$ curve in the wings and steepen it towards the centre. Again, the principal effect is, however, the reduction of diffractive efficiency. Even for losses that reduce the efficiency by 10 times, the broadening in $\Delta\lambda$ or $\Delta\theta$ for typical display holography type gratings is rarely more than 10%. One further effect is worth mentioning however. Loss does reduce somewhat the significant side lobes in the $\eta(\theta d)$ curve present at higher values of the parameter κd .

The formulae we have derived in the last few sections very much form part of the essential toolkit for emulsion design in display holography. By choosing values of $\delta n/n$ and the emulsion thickness, we can now calculate the expected diffractive efficiencies and the expected angular and wavelength selectivities for either transmission or reflection geometries. Experience has shown that Kogelnik's coupled wave theory works extremely well in most cases of interest. More complex theoretical work, numerical simulation and experimental comparison all show that the theory is remarkably good for the low modulations normally encountered.

11.10.8 Recording of Complex Wave Fronts and Multiple Gratings

Throughout this chapter, we have examined how a single grating behaves. We have seen how such gratings can be expected to faithfully replay "image points" in three-dimensional space when we employ a spherical-wave formulism. Using a simpler plane wave formulism, we have then answered questions about the diffractive response and the angular and wavelength selectivities of the hologram. Here, we have seen that, with the right design for the emulsion in terms of permittivity modulation and thickness, we can expect to produce high-brightness reproductions of simple recorded waves.

Two critical questions remain unanswered, however. How does the hologram cope with complex wave fronts of the kind required to synthesise the three-dimensional image of a real object? And how does the hologram cope with the polychromatic gratings, which will inevitably be required for full-colour reflection holography?

11.10.9 Multiple Gratings Generated by Many Object Points

The question of spatially multiplexed gratings may be addressed quite simply by an extension of the coupled wave theory. This is the two-dimensional N -coupled wave theory [2]. We will concentrate first on the problem of recording a complex object composed of many object points. In the limit, we can regard this as describing the holographic recording of a diffuse object. Here, we will follow Solymar and Cooke [4] and assume a reference wave (labelled by subscript 0) and $N - 1$ object waves of the form

$$E_j = A_{j0}a_j(x, y) \exp(-i\beta\rho_j) \quad (11.149)$$

Here, A_{j0} is a complex amplitude, a_j is the normalised amplitude distribution and ρ_j is the phase of the wave. We will assume an infinite two-dimensional slab of thickness d with z being an ignorable coordinate (Figure 11.24). As before, we take the σ -polarisation for simplicity.

We will now assume that the reference wave is large compared with the other waves. This is a vital assumption because, without it, radiation from individual points will interact together to form their own gratings. If we make this assumption, then the grating created will be free from cross-terms and we can write

$$\Delta\epsilon = g \left| \sum_{j=0}^{N-1} E_j \right|^2 = \frac{1}{2} \sum_{j=1}^{N-1} a_o a_j \left\{ \epsilon_j \exp(-i\beta[\rho_0 - \rho_j]) + \epsilon_j^* \exp(i\beta[\rho_0 - \rho_j]) \right\} \quad (11.150)$$

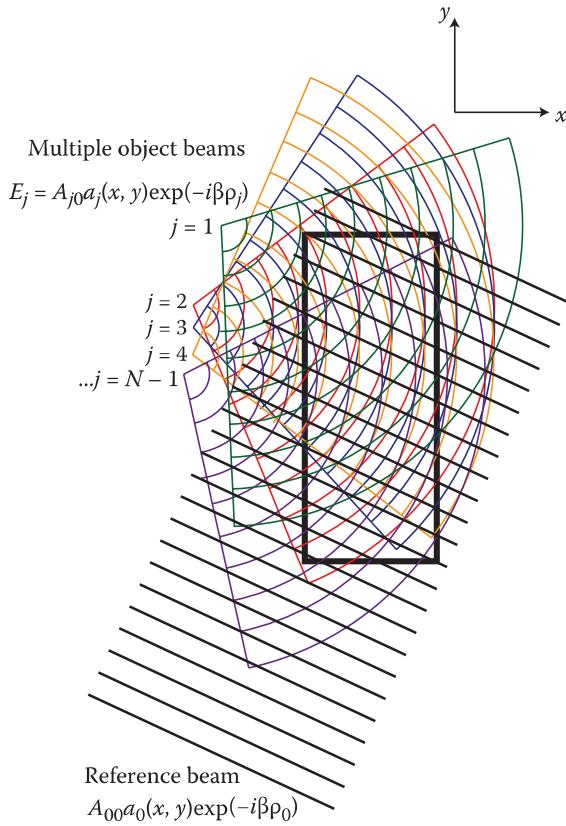


FIGURE 11.24 Two-dimensional N -coupled wave theory. An infinite two-dimensional slab grating is recorded with reference wave of complex amplitude A_{00} and j object waves of complex amplitude A_{j0} . All waves are assumed to be of σ -polarisation.

where

$$\varepsilon_j = 2gA_{00}A_{j0}^* \quad (11.151)$$

and where g is a constant. Of course, this expression would be exact if the individual object points were recorded sequentially rather than all at the same time. However, in the limit of a large-amplitude reference beam, it also provides a good approximation of the actual recorded grating. To see how this grating replays when illuminated by the original reference beam, we must solve the wave equation

$$\nabla^2 E + \beta^2(1 + \Delta\varepsilon/\varepsilon_{r0})E = 0 \quad (11.152)$$

Clearly, the Bragg condition will be satisfied for each of the $N - 1$ waves and so all will be reconstructed. Assuming no other waves are produced, we can then adopt a trial solution of the form

$$E = \sum_{j=0}^{N-1} A_j(x, y)a_j(x, y)\exp(-i\beta\rho_j) \quad (11.153)$$

One then proceeds exactly as with the one-dimensional coupled wave theory, equating coefficients of each exponential term to zero. Discarding second-order derivatives, assuming exact Bragg compliance for each signal wave, keeping only the signal waves corresponding to our trial function and using the geometric optics laws,

$$\begin{aligned} (\nabla\rho_j)^2 &= 1 \\ \nabla \cdot (a_j^2 \nabla\rho_j) &= 0 \end{aligned} \quad (11.154)$$

we are led to the following set of coupled differential equations:

$$\nabla A_0 \cdot \nabla p_0 + i \sum_{j=1}^{N-1} \kappa_j a_j^2 A_j = 0 \quad (11.155)$$

$$\nabla A_j \cdot \nabla p_j + i \kappa_j^* a_0^2 A_0 = 0 \quad \forall j \in \{1, 2, 3, \dots, (N-1)\} \quad (11.156)$$

where

$$\kappa_j = \frac{\epsilon_j \beta}{4\epsilon_{r0}} \quad (11.157)$$

These equations can now be solved analytically for the case that all the waves are plane waves. In fact, this turns out to be a simple extension of the one-dimensional coupled wave theory. Taking the angle of incidence of the j th wave as θ_j , we may rewrite Equations 11.155 and 11.156 as

$$\cos \theta_0 \cdot \frac{dA_0}{dx} + i \sum_{j=1}^{N-1} \kappa_j A_j = 0 \quad (11.158)$$

and

$$\frac{dA_j}{dx} \cos \theta_j + i \kappa_j^* A_0 = 0 \quad \forall j \in \{1, 2, 3, \dots, (N-1)\} \quad (11.159)$$

Note that $a_j = 1 \forall j$. These equations may then be recast into a simple second-order ordinary differential equation:

$$\frac{d^2 A_0}{dx^2} + A_0 \left\{ \sec \theta_0 \sum_{j=1}^{N-1} \sec \theta_j |\kappa_j|^2 \right\} = 0 \quad (11.160)$$

Now the boundary conditions for a transmission hologram are $A_0 = 1$ and $A_j = 0 \forall j \neq 0$ at $x = 0$. This then yields the solution

$$\begin{aligned} A_0 &= \cos(Lx) \\ A_j &= -i \frac{\kappa_j^*}{L} \sin(Lx) \sec \theta_j \quad \forall j \in \{1, 2, 3, \dots, (N-1)\} \end{aligned} \quad (11.161)$$

where

$$L^2 = \sec \theta_0 \sum_{j=1}^{N-1} |\sec \theta_j| |\kappa_j|^2 \quad (11.162)$$

If we take $j = 1$, then Equation 11.161 reduces to Equation 11.102—the standard result from the one-dimensional coupled wave theory. As we discussed earlier, in this case, when $Lx = \pi/2$, complete power transfer occurs between the provoking wave and the signal wave—that is, when $Lx = \pi/2$, $A_0 = 0$ and A_1 is a maximum. If we now, however, consider the case of $N > 1$, then $Lx = \pi/2$ and A_0 is still zero, but this time, the total power in this wave has been donated to the $N-1$ signal waves. We see very clearly then that the N -coupled wave theory predicts that complete power transfer should be possible from a single reference wave to many signal waves under the assumption of a strong reference beam at recording. From this, we can deduce that a hologram made up of many image points may theoretically at least be expected to play back with perfect diffractive efficiency under the correct conditions. From the above

solution, we see that each of the plane waves reconstructed faithfully reproduces the directional information recorded. The relative amplitudes are, however, subject to some modification for using Equation 11.161, we see that

$$\frac{A_i}{A_j} = \frac{A_{i0} \cos \theta_j}{A_{j0} \cos \theta_i} \quad (11.163)$$

This means that the image is only reconstructed perfectly in amplitude in the central field (i.e., paraxially)—at higher angles, the amplitudes are modified according to an angle law. This is not, however, a great problem, as the angular field is usually restricted, and even at large angles of view, Snell's law acts to strongly confine the angular field within the emulsion layer of most display holograms. Finally, any residual amplitude modification can of course be compensated for in the case of digital holograms.

Exactly the same argument applies to reflection holograms. Here, the N -wave analytical solution for plane waves is

$$\begin{aligned} A_0 &= \frac{\cosh(L\{d-x\})}{\cosh(Ld)} \\ A_j &= -\frac{i\kappa_j^*}{L|\cos \theta_j|} \frac{\sinh(L\{d-x\})}{\cosh(Ld)} \end{aligned} \quad (11.164)$$

If we define the diffractive efficiency of the j th wave as

$$\eta_j = \frac{|\cos \theta_j|}{|\cos \theta_0|} A_j(0) A_j^*(0) \quad (11.165)$$

then each individual wave will have a diffractive power given by

$$\eta_j = \frac{|k_j|^2}{L^2} |\sec \theta_j| |\sec \theta_0| \tanh^2(Ld) \quad (11.166)$$

or

$$\eta_j = \frac{|k_j|^2}{|\cos \theta_j| \sum_{j=1}^{N-1} |\sec \theta_j| |k_j|^2} \tanh^2 \left[d \sqrt{\sec \theta_0 \sum_{j=1}^{N-1} |\sec \theta_j| |k_j|^2} \right] \quad (11.167)$$

The total diffractive efficiency is likewise given by

$$\eta = \sum_{j=1}^{N-1} \eta_j = \tanh^2 \left[d \sqrt{\sec \theta_0 \sum_{j=1}^{N-1} |\sec \theta_j| |k_j|^2} \right] \quad (11.168)$$

showing clearly that the $N-1$ signal waves can inherit 100% of the power from the driving wave.

We should perhaps put in a slight word of caution here. N -wave theory as applied to plane waves does seem to offer an extremely simple model of holographic replay. Nevertheless, the coupling between the waves which we discard in N -wave theory is in some ways equivalent to a deviation from the Bragg condition in the remaining gratings—and this leads, in the real world, to imperfect power coupling between the reference reading wave and the signal waves.

A treatment of N -wave theory that deals with the non-compliance of the Bragg condition has been given by Peri and Friesem [16]. Cooke and Solymar [3] have further shown that analytic solutions exist for a version of the theory in which no couplings are ignored—in fact, in this work, too much coupling is included. The result is therefore to be regarded as a pessimistic estimate of any final diffractive efficiency. For small angles of incidence, Cooke and Solymar find that for a reference to object intensity ratio of 5, a maximum diffractive efficiency of 95.4% is found for a transmission grating. For equal object and reference intensities, this decreases to 80%.

11.10.10 Recording Multiple Colour Gratings

The problem of co-recording two or more gratings using different wavelengths of light is very similar to the problem of multiple gratings, which we discussed in the last section. The only difference is that now there will exist gratings within the volume hologram which do not satisfy Bragg compliance at all. A hologram made with a red reference and red object beam and a green reference and green object beam will lead to a permittivity distribution formally described by an equation similar to Equation 11.150. The main difference is that half of the signal waves will not reconstruct under stimulation by the red reference reading beam as the “green” gratings will not be in Bragg compliance for the red reference beam. If we assume naively that this non-compliance does not matter, then again, we are led to the conclusion that perfect diffractive conversion should be theoretically possible in a multiple-colour hologram. We shall see in the next chapter that the PSM model brings us to this same conclusion but rather more clearly.

The incoherent storage of many holograms is a subject of primary interest to the field of holographic memory. Here, a given material is characterised by its M number [17]. This is a number that characterises the dynamic range of the material—that is, how many optimally diffractive holograms may be incoherently multiplexed for a material of given thickness. The M number is defined as

$$M/\# = \sum_{i=1}^M \eta_i^{1/2} \quad (11.169)$$

where η_i is the maximum diffractive efficiency of the i th co-written grating.

Ulibarrena [18,19] has recently investigated the panchromatic silver halide emulsion, BB640, produced by the UK company Colour Holographics Ltd., and compared experimental measurements of the diffractive efficiency of three-colour multiplexed reflection gratings (442, 532 and 632.8 nm) made on this material with the predictions of the coupled wave theory. A best-fitting procedure of the measured and theoretically expected diffractive efficiency versus wavelength produced an estimated emulsion thickness of 7.3 μm and an index modulation of each component grating of $\delta n/n = 0.027/1.579$. The measured diffractive efficiency of unslanted gratings at each of their three peak replay wavelengths was more than 52% (for $\theta_r = \theta_c = 0^\circ$). When only a one-colour grating was recorded in the emulsion, typical diffraction efficiencies exceeded 72%. Ulibarrena also recorded diffuse images and measured approximate diffractive efficiencies for three-colour BB640 reflection holograms of 40% at 632.8 nm, and 20% at 442 nm/532 nm.

11.10.11 Dispersion Equation Theory

In the previous sections, we have given a simple exposition of the coupled wave theory. As we have stated, the coupled wave theory is a simple but highly effective method of analysing a volume grating. In its two-wave form, it is not mathematically rigorous, but for all practical applications of display holography, it is extremely useful. The interested reader is referred to the excellent book by Solymar and Cooke [4] for a discussion of how this theory may be generalised to two and three dimensions. There are, however, some alternative approaches to the coupled wave theory.* In particular, the analysis techniques used in the field of x-ray and electron diffraction yielded an alternative method for analysing the volume

* One such variant is the PSM model, which we shall discuss in detail in Chapter 12.

hologram known as dispersion equation theory. This theory was first suggested in 1967 by Saccoccio [20], two years before Kogelnik's theory was published. It was not until 1971, however, that Aristov and Shekhtman [21] outlined the first proper treatment of a two-wave and N -wave dispersion treatment of the volume hologram assuming perfect Bragg compliance. In 1976, Sheppard [22] produced a version in which Bragg compliance was not assumed, essentially reproducing an equivalent but alternative theory to Kogelnik's. Dispersion equation theory is, in many ways, very similar to the coupled wave theory. In the simplest variant, a trial solution of the wave equation assumes just two waves, R and S , just as in Kogelnik's theory. However, these two waves are assumed constant throughout the grating, leading to a set of algebraic equations rather than differential equations. Each of the two waves is then split into a further two component waves and these waves beat to produce the one-dimensional variation implicit to Kogelnik's theory. The results of dispersion equation theory are usually very close to the coupled wave theory, although there are differences. The theory is, however, fundamentally based on plane waves and as such cannot be generalised in the way the coupled wave theory can be. However, as a model of what is going on in the grating, dispersion equation theory offers an alternative way of looking at things, which is often useful.

11.11 Blurring Revisited

In Section 11.2.3, we worked out that the presence of a broadband of wavelengths at reconstruction would act to blur the image of a hologram in the vertical and depth directions. We saw that in the vertical direction, an image point in the thin transmission or reflection hologram would be blurred by an amount

$$\delta y_i \sim x_o \lambda_r \sin \theta_c \frac{\delta \lambda}{\langle \lambda_c \rangle^2} \quad (11.170)$$

We can now use Equation 11.141 to calculate, for the case of an unslanted reflection grating with $\kappa d / \sqrt{c_R |c_S|} = \pi/2$ *, that at the point of diffraction, the reconstructed image ray will produce a cone of light having an angle of approximately†

$$\delta \theta \sim \frac{\sqrt{5}}{4} \frac{\lambda}{dn} \tan \theta_c \quad (11.171)$$

Note that this formula is different from Equation 11.140, which gave an estimate for the expected tolerance in reconstruction angle at a fixed wavelength. Equation 11.171 predicts that for an average fine-grain silver halide emulsion, such as the Slavich VRP-M, which has a thickness of approximately 7 µm, one can expect an angular blurring angle of around 1° at 532 nm or approximately 17 times that of the human eye's resolution.‡ For the latest Bayer photopolymer, which is approximately twice as thick, the blurring halves, but it is still more than eight times worse than the practical human eye resolution. The effect of this chromatic blurring is to produce a finite depth and a finite front projection in which a clear image can be displayed by a reflection hologram illuminated by a broadband source. We saw in Section 11.2.3 that the critical defocussing distance behind the hologram was given by

$$x_{\text{CRIT}} = \frac{h \delta \theta_{\text{Eye}}}{\delta \theta - \delta \theta_{\text{Eye}}} \quad (11.172)$$

* This represents a pretty good design choice, giving a diffractive efficiency of 0.84 under the lossless coupled wave theory.

† Note that we could have derived a similar formula by considering $\theta_i(\lambda_c) = \tan^{-1}[k_{iy}/k_{ix}]$ under a variation of $\lambda_c \rightarrow \lambda_c + \delta \lambda_c$.

This leads to $\delta \theta \sim \frac{\sqrt{5}}{2} \frac{\lambda}{dn} \sin \theta_r$.

‡ This is assuming an angle θ_c within the emulsion of 30°.

For the VRP-M material, at an observation distance of 1 m and at 532 nm, this gives a value of approximately 7 cm. The front projection is also around this figure leading to a total clear image depth of field of approximately 14 cm. For the Bayer photopolymer and for the ultrafine grain silver halide emulsion PFG-03CN from Sfera-S, this roughly doubles to a total clear field of 28 cm.

Of course, these calculations pertain to analogue reflection holograms and full-parallax digital reflection holograms. For HPO digital holograms illuminated by a distant point source, we have already seen that there is no chromatic blurring in the paraxial approximation. We come to an important conclusion then. For full-parallax reflection holographic images to be unaffected by the chromatic blurring inherent in broadband illumination, we must increase the emulsion thickness to more than 100 μm . The permittivity modulation must then decrease significantly so that $\kappa d/\sqrt{c_R |c_S|}$ continues to be of the order of $\pi/2$. There is, however, a sizeable problem here. In materials such as silver halide, chemical processing needs to effect a change throughout the emulsion thickness, and for emulsions much thicker than approximately 15 μm , one starts to see problems of penetration of the processing chemicals [23]. As a result, non-uniform development occurs and it becomes difficult to guarantee the uniformity of the grating. Silver halides can be made sensitive to pulsed radiation, and as such, these materials have a great advantage for the generation of digital holograms. Realistically, they work best at lower thicknesses. Photopolymers, on the other hand, are usually not easily sensitised to pulsed laser radiation, but they are easier to produce in large thicknesses.* Bayer, for example, has recently tested a 50 μm material sensitive in the red and green spectral regions with a value of $\delta n \sim 0.012$ [24]. Such material may be expected to show a clear depth behind an analogue or full-parallax reflection hologram of more than 1 m when viewed from a distance of 1 m.

There is, however, another problem with thick emulsions. The very fact that they have such great wavelength selectivity means that they only use a fraction of any broadband light illuminating them! As we shall discuss in Chapter 13, modern light-emitting diode sources can currently produce bandwidths down to approximately 15 nm at the time of writing. A 50 μm emulsion has a value of $\delta\lambda$ at 532 nm of approximately 2.5 nm. So one could expect that maybe only 20% of the light will actually be used depending on the form of the distribution. As such, until smaller bandwidth light-emitting diodes or other cheap illumination sources become available, most full-parallax holograms are always going to be either dimmer or shallower than HPO holograms. The obvious exception is of course when lasers are used to illuminate the hologram.

We have used the coupled wave theory to derive expressions for the angular replay sensitivity of both the transmission and reflection hologram (Equations 11.138 and 11.140) in the absence of fringe slant. For the reflection case, we can use Equation 11.140 to see that for a typical silver halide emulsion of 7 μm thickness, we can expect an angular selectivity of the order of 5° depending on replay angle and wavelength. In fact, Equation 11.140 gives us an estimate of the half-width at zero diffractive efficiency. However, this figure is also very close to the FWHM as can be verified from the plots in Figures 11.20 and 11.21. This value is very much larger than the human eye value of $\delta\theta_{\text{Eye}} \sim 0.06^\circ$ meaning that source-size blurring will be essentially unaffected by the Bragg effect of the volume grating. Unfortunately, the only way with these types of emulsion to avoid source-size blurring is to use a very compact light source!

11.11.1 Blurring in the Digital Hologram

We have discussed at length the method of DWDH printing in Chapters 7, 8 and 9. Here, the hologram is split into a matrix of abutting square microholograms or hogels. This technique has been extremely successful in generating both small and large full-colour digital holograms of excellent quality. In a certain obvious sense, just like standard two-dimensional dot matrix printing, the smaller the hogel, the better the image quality. Digital holograms have been produced with hogels down to almost 250 μm [25]. However, there are two principal disadvantages to reducing the hogel size. The first is the time required to write a hologram—clearly, if one reduces the linear dimension of the hogel by two times, then it is going to take four times as long to print the hologram. For large holograms, which may measure several metres in linear dimension, this is an important constraint. Usually, such holograms are viewed at large distances; as such, the observer will simply not notice the difference between a hogel size of 1 mm and 0.5 mm.

* One should note, however, that thicker photopolymers can suffer from absorption of the recording light by the photosensitiser.

The second disadvantage of reducing the hogel size is diffractive blurring. For a hogel diameter of d , each hogel will radiate a cone of light, due to diffraction, having an angle of $\delta\theta \sim \lambda/(\gamma d)$, in which the form factor γ is 1.0 for a square hogel and 1.22 for a circular hogel. This will lead to a blurring of the holographic image if $\delta\theta$ is greater than the smallest angle resolvable by the human eye. If $\delta\theta > \delta\theta_{\text{Eye}}$, then the hologram will exhibit a maximum “in focus” depth after which it will appear blurred in accordance with our discussion in Section 11.2.2. Taking the most constraining wavelength as 440 nm and estimating $\delta\theta_{\text{Eye}} \sim 0.06^\circ$, we see that the critical square hogel diameter for the onset of diffractive blurring is approximately 0.5 mm. Of course, this figure depends strongly on $\delta\theta_{\text{Eye}}$ which varies somewhat depending on the person. As a result, it is not surprising that some observers report that large-depth and large-projection holograms made with 0.8 mm-diameter hogels appear more blurred than when written with 1.6 mm hogels. Of course this is under good illumination conditions which ensure that source-size and chromatic blurring are both less than $\delta\theta_{\text{Eye}}$.

Hogel sizes less than 0.5 mm can demonstrate sizeable blurring. For example, a square hogel of 0.1 mm diameter will show a value of $\delta\theta$ equal to around 0.26° . Using Equation 11.172, this then leads to a blur-free depth of 30 cm when the hologram is observed from a distance of 1 m.

11.12 Computational Methods of Calculating Diffractive Efficiency of Planar Gratings

11.12.1 Rigorous Coupled Wave Theory and Rigorous Modal Theory

In 1981, Moharam and Gaylord [12] described a rigorous coupled wave theory, which provided the first fully accurate and convenient computational solution to the problem of diffraction from slanted planar gratings. It was shown that Kogelnik’s analytic theory provided a rather good description of the unslanted reflection grating, but at high index modulations ($\delta n/n \sim 0.16$), it somewhat overestimated the diffraction efficiency of the slanted grating. Differences in the transmission grating were shown to appear at a rather lower modulation ($\delta n/n \sim 0.06$). The main reason for the difference between Kogelnik’s theory and the rigorous coupled wave theory in the slanted reflection grating is in the treatment of boundary diffraction. This is related to the neglect of the second-order derivatives in Kogelnik’s theory. Differences in the transmission hologram arise because other diffractive orders become important.

An earlier approach to the computational solution of diffraction from planar gratings is the modal approach [26–36]. Both the rigorous coupled wave approach and the (rigorous) modal approach analyse the planar grating diffraction problem by solving the wave equation in three regions—the grating region and two exterior regions—and then matching the tangential electric and the magnetic fields at the two boundaries to determine the unknowns. The main difference between the two approaches is in the technique used to find solutions of the wave equation in the grating region. In the rigorous coupled wave approach, the system of coupled wave equations is usually formulated as a simple matrix relation for which the solution is obtained by calculating the eigenvalues and the eigenvectors. The modal approach requires a more complicated transcendental relationship in the form of a continued fraction expansion to be solved to find the wave numbers and their related coefficients, which are needed to satisfy the wave equation in the grating region. The primary difficulty when applying the modal approach to the analysis of slanted gratings is the difficulty in formulating a systematic technique capable of solving the general transcendental continued fraction relation. Rigorous coupled wave theory and its application to simple polychromatic and more complex spatially multiplexed gratings are described in more detail in Appendix 8.

11.12.2 Rigorous Chain Matrix Method

The rigorous coupled wave theory provides a useful way to calculate accurate diffraction efficiencies for any planar grating. Another method, the rigorous chain matrix method [37–41] can, however, be programmed extremely quickly on any PC for the case of the unslanted planar grating. The method is based on dividing up the grating into thin layers each of a given index. A general index profile can then be constructed by using many layers. For the case of the reflection grating, a reference wave, R , and a

counterpropagating signal wave, S , are assumed. The tangential components of the electric and magnetic fields are then matched at each layer boundary.

At the m th layer, this gives

$$\begin{aligned} R_m &= t_m^{-1} \left\{ R_{m-1} e^{i\delta_{m-1}} + r_m S_{m-1} e^{-i\delta_{m-1}} \right\} \\ S_m &= t_m^{-1} \left\{ r_m R_{m-1} e^{i\delta_{m-1}} + S_{m-1} e^{-i\delta_{m-1}} \right\} \end{aligned} \quad (11.173)$$

where for the π -polarisation, the reflection and transmission coefficients are

$$\begin{aligned} r_m &= \frac{n_m \cos \phi_{m-1} - n_{m-1} \cos \phi_m}{n_m \cos \phi_{m-1} + n_{m-1} \cos \phi_m} \\ t_m &= \frac{2n_m \cos \phi_m}{n_m \cos \phi_{m-1} + n_{m-1} \cos \phi_m} \end{aligned} \quad (11.174)$$

and for the σ -polarisation

$$\begin{aligned} r_m &= \frac{n_m \cos \phi_m - n_{m-1} \cos \phi_{m-1}}{n_m \cos \phi_m + n_{m-1} \cos \phi_{m-1}} \\ t_m &= \frac{2n_m \cos \phi_m}{n_m \cos \phi_m + n_{m-1} \cos \phi_{m-1}} \end{aligned} \quad (11.175)$$

Here, n_m and ϕ_m are, respectively, the index and ray angle of incidence at the m th layer. The phase parameters are likewise given by

$$\delta_m = \frac{2\pi n_m}{\lambda} \delta \cos \phi_m \quad (11.176)$$

with δ being the distance between successive layers. Finally, the angles of incidence at each layer are given by Snell's law:

$$n_m \sin \phi_m = n_{m-1} \sin \phi_{m-1} \quad (11.177)$$

Equation 11.173 may be written in the matrix form

$$\begin{pmatrix} R_m \\ S_m \end{pmatrix} = \frac{1}{t_m} M_{m-1} \begin{pmatrix} R_{m-1} \\ S_{m-1} \end{pmatrix} \quad (11.178)$$

The solution for a stack of k layers on a substrate may then be written using the chain matrix form

$$\begin{pmatrix} R_{k+1} \\ S_{k+1} \end{pmatrix} = \frac{M_k M_{k-1} \dots M_2 M_1}{t_k t_{k-1} \dots t_2 t_1} \begin{pmatrix} R_0 \\ S_0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} R_0 \\ S_0 \end{pmatrix} \quad (11.179)$$

Using the boundary conditions, $R_0 = 1$ and $S_{k+1} = 0$, the diffraction efficiency of the reflection grating may then be calculated as

$$\eta = S_0 S_0^* = \left| \frac{c}{d} \right|^2 \quad (11.180)$$

The chain matrix method provides an extremely practical method for calculating the diffraction efficiency of any unslanted planar grating—and this includes polychromatic gratings! The method can also be extended to a complex index to model lossy polychromatic gratings. In the next chapter, we shall see that a differential formulation of this method provides a powerful analytic alternative to the coupled wave theory.

REFERENCES

1. H. Kogelnik, "Coupled wave theory for thick hologram gratings," *Bell Syst. Tech. J.* **48**, 2909–2947 (1969).
2. L. Solymar, "Two-dimensional N-coupled-wave theory for volume holograms," *Opt. Commun.* **23**, 199–202 (1977).
3. D. J. Cooke and L. Solymar, "Comparison of two-wave geometrical optics and N-wave theories for volume phase holograms," *J. Opt. Soc. Am. B* **70**, 1631A (1980).
4. L. Solymar and D. J. Cooke, *Volume Holography and Volume Gratings*, Academic Press Inc., New York (1981).
5. H. M. Smith, *Principles of Holography*, 2nd Edn, John Wiley and Sons Inc. (1975).
6. S. A. Benton, "Hologram reconstructions with extended incoherent sources," *J. Opt. Soc. Am.* **59**, 1545–1546A (1969).
7. J. Wolff, N. Phillips and A. Furst, *Light Fantastic*, Bergström + Boyle Books Ltd, London (1977) and J. Wolff, N. Phillips and A. Furst. *Light Fantastic 2*, Bergström + Boyle Books Ltd, London (1978).
8. P. Dawson and P. A. Wilksch, "Laser transmission holograms maximum permissible exposure," in *Laser Florence 2009*, AIP Conf. Proc. **1226**, 147–154 (2009).
9. P. Dawson, *Virtual Encounters; Paula Dawson Holograms*, Macquarie Univ. and Newcastle Region Art Gallery, Australia (2010).
10. D. Brotherton-Ratcliffe, "A treatment of the general volume holographic grating as an array of parallel stacked mirrors," *J. Mod. Optic* **59**, 1113–1132 (2011).
11. J. D. Jackson, *Classical Electrodynamics*, 3rd Edn, John Wiley and Sons Inc. (1974).
12. M. G. Moharam and T. K. Gaylord, "Rigorous coupled wave analysis of planar grating diffraction," *J. Opt. Soc. Am.* **71**, 811–818.
13. A. B. Bhatia and W. J. Noble, "Diffraction of light by ultrasonic waves," *Proc. Roy. Soc. A* **220**, 356–385 (1953).
14. P. Phariseau, "On the diffraction of light by supersonic waves," *Proc. Ind. Acad. Sci.* **44A**, 165–170 (1956).
15. F.-K. Bruder, F. Deuber, T. Fäcke, R. Hagen, D. Hönel, D. Jurbergs, M. Kogure, T. Rölle and M.-S. Weiser, "Full-color self-processing holographic photopolymers with high sensitivity in red – the first class of instant holographic photopolymers," *J. Photopolym. Sci. Technol.* **22**(2) 257–260 (2009).
16. D. Peri and A. A. Friesem, "Volume hologram for image restoration," *J. Opt. Soc. Am.* **70**, 515–522 (1980).
17. H. J. Coufal, D. Psaltis and G. T. Sincerbox, *Holographic Data Storage*, Springer, New York (2000).
18. M. Ulibarrena, *Estudio y caracterización de la emulsión fotográfica de grano ultrafino BB640 como material de registro holográfico*, [Transl: Study and characterization of the ultra-fine-grain photographic emulsion BB640 as a holographic recording material] PhD Thesis, Univ. Miguel Hernandez, Elche, Spain (2003).
19. M. Ulibarrena, L. Carretero, R. F. Madrigal, S. Blaya and A. Fimia, "Multiple band holographic reflection gratings recorded in new ultra-fine-grain emulsion BBVPan," *Opt. Expr.* **11**, 3385–3392 (2003).
20. E. J. Saccoccio, "Application of the dynamical theory of X-ray diffraction to holography," *J. Appl. Phys.* **38**, 3994–3998 (1967).
21. V. V. Aristov and V. Shekhtman, "Properties of three-dimensional holograms," *Sov. Phys. Uspekhi* **14**, 263–277 (1971).
22. C. J. R. Sheppard, "The application of the dynamical theory of X-ray diffraction to thick hologram gratings," *Int. J. Electron.* **41**, 365–373 (1976).
23. D. Dainton, A. R. Gattiker and W. O. Lock, "The processing of thick photographic emulsions," *Phil. Mag.*, (Series 7) **42**, Issue 327 (1951).
24. T. Fäcke, F.-K. Bruder, M.-S. Weiser, T. Rölle and D. Hönel, *Novel holographic media and photopolymers*, US patent application US 2011/0065827 A1 (2011).
25. S. Zacharovas, Geola Digital UAB, Private Communication (2012).
26. T. Tamir, H. C. Wang and A. A. Oliner, "Wave propagation in sinusoidally stratified dielectric media," *IEEE Trans. Microwave Theory Tech.* **MTT-12**, 323–335 (1964).
27. T. Tamir and H. C. Wang, "Scattering of electromagnetic waves by a sinusoidally stratified half-space: I. Formal solution and analysis approximations," *Can. J. Phys.* **44**, 2073–2094 (1966).

28. T. Tamir, "Scattering of electromagnetic waves by a sinusoidally stratified half-space: II. Diffraction aspects of the Rayleigh and Bragg wavelengths," *Can. J. Phys.* **44**, 2461–2494 (1966).
29. C. B. Burckhardt, "Diffraction of a plane wave at a sinusoidally stratified dielectric grating," *J. Opt. Soc. Am.* **56**, 1502–1509 (1966).
30. L. Bergstein and D. Kermisch, "Image storage and reconstruction in volume holography," *Proc. Symp. Modern Opt.* **17**, 655–680 (1967).
31. R. S. Chu and T. Tamir, "Guided-wave theory of light diffraction by acoustic microwaves," *IEEE Trans. Microwave Theory Tech.* **MTT-18**, 486–504 (1970).
32. R. S. Chu and T. Tamir, "Wave propagation and dispersion in space-time periodic media," *Proc. IEE* **119**, 797–806 (1972).
33. F. G. Kaspar, "Diffraction by thick periodically stratified gratings with complex dielectric constant," *J. Opt. Soc. Am.* **63**, 37–45 (1973).
34. S. T. Peng, T. Tamir, and H. L. Bertoni, "Theory of periodic dielectric waveguides," *IEEE Trans. Microwave Theory Tech.* **MTT-23**, 123–133 (1975).
35. R. S. Chu and J. A. Kong, "Modal theory of spatially periodic media," *IEEE Trans. Microwave Theory Tech.* **MTT-25**, 18–24 (1977).
36. R. Magnusson and T. K. Gaylord, "Equivalence of multiwave coupled-wave theory and modal theory of periodic-media diffraction," *J. Opt. Soc. Am.* **68**, 1777–1779 (1978).
37. O. S. Heavens, "Optical Properties of Thin Films," *Reports on Progress in Physics*, Vol **XXIII**, p. 1, (1960).
38. F. Abeles, "Recherches sur la propagation des ondes électromagnétiques sinusoïdales dans les milieux stratifiés. Application aux couches minces," *Ann. Phys. (Paris)* **5**, 596–640 (1950).
39. M. G. Moharam and T. K. Gaylord, "Chain-matrix analysis of arbitrary-thickness dielectric reflection gratings," *J. Opt. Soc. Am.* **72**, 187–190 (1982).
40. X. Ning, "Analysis of multiplexed-reflection holographic gratings," *J. Opt. Soc. Am. A* **7**, 1436–1440 (1990).
41. D.W. Diehl and N. George, "Analysis of multitone holographic interference filters by use of a sparse Hill matrix method," *Appl. Opt.* **43**, 88–96 (2004).

12

Diffraction Efficiency: An Alternative Approach Using the PSM Model

12.1 Introduction

An alternative model to Kogelnik's coupled wave theory [1] of the volume holographic grating can be developed through a more refined version of the parallel stacked mirror (PSM) model [2,3], which we presented in the last chapter. Indeed, in 1982, Ludman [4] published a simple analysis of the volume transmission phase grating, which used the Fresnel reflection formulae to provide an estimation of the diffraction efficiency at Bragg resonance. In 2009, Heifetz et al. [5] extended this idea somewhat. However, we shall see that the PSM model can be taken *much* further and that it in fact provides a truly alternative picture to the coupled wave theory. It also sheds some useful light on the assumptions made in Kogelnik's work. This more rigorous treatment of the PSM model allows a proper consideration of multiple-colour reflection gratings. We will see that it is also capable of analytically describing spatially multiplexed polychromatic reflection gratings [3] and in providing useful formulae for the scaling of the diffractive efficiency of full-colour reflection holograms.

The PSM model is based on a particular mathematical description of the permittivity distribution of the unslanted volume holographic grating. The distribution is broken up into an infinite number of infinitesimal discontinuities or step functions. The Fresnel reflections of an incident plane wave from each of these discontinuities are then summed up in a consistent way. The resulting first-order coupled partial differential equations are solved in a rotated frame of reference to arrive at analytical expressions for the diffraction efficiency of the general slanted grating. The PSM model can be viewed as a differential representation of the chain matrix model [6,7]. Both the chain matrix model and the basic differential equations of the PSM model rigorously describe the phenomenon of diffraction from unslanted planar reflection gratings.

12.2 Formulation of the Simplest Model—The Unslanted Reflection Grating at Normal Incidence

We will consider a reflection holographic grating with the following index profile

$$n = n_0 + n_1 \cos\left(\frac{4\pi n_0}{\lambda_r} y\right) = n_0 + \frac{n_1}{2} \left\{ e^{\frac{4i\pi n_0}{\lambda_r} y} + e^{-\frac{4i\pi n_0}{\lambda_r} y} \right\} \quad (12.1)$$

Here, n_0 is the average index and n_1 is generally a small number representing the index modulation. We can imagine that this grating has been created by the interference of two counterpropagating normal-incidence plane waves within a photosensitive material, each with a wavelength of λ_r .*

We wish to understand the response of the grating to a plane reference wave of the form

$$R^{\text{ext}} = e^{i\beta y} \quad (12.2)$$

* Both λ_r and λ_c refer to wavelengths in vacuum.

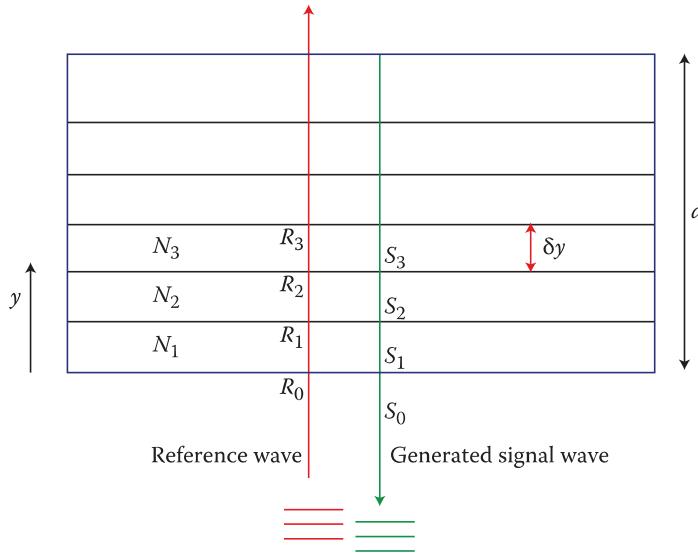


FIGURE 12.1 A simple model of the unslanted normal-incidence reflection grating. The grating is assumed to be made up from discrete layers of different refractive indices. The reference wave illuminates the grating and is reflected and transmitted according to the standard Fresnel rules, generating the signal wave.

where

$$\beta = \frac{2\pi n_0}{\lambda_c} \quad (12.3)$$

We shall assume that the grating is surrounded by a zone of constant index, \$n_0\$, to circumvent the complication of refraction and reflection at the grating interface. We start by modelling the grating of Equation 12.1 by a series of many thin constant-index layers, \$N_0, N_1, N_2, \dots, N_M\$, between each of which there exists an index discontinuity (Figure 12.1). Across each such discontinuity, we may derive the well-known Fresnel formulae [e.g., 8] for the amplitude reflection and transmission coefficients from Maxwell's equations by demanding that the tangential components of the electric and magnetic fields be continuous. An illuminating plane wave will, in general, generate many mutually interfering reflections from each discontinuity. We therefore imagine two plane waves within the grating—the driving reference wave, \$R(y)\$, and a created signal wave, \$S(y)\$. Using the Fresnel formulae, we may then write the following relationship for either the \$\sigma\$- or \$\pi\$-polarisations*

$$\begin{aligned} R_J &= 2e^{i\beta n \delta y / n_0} \left\{ \frac{N_{J-1}}{N_J + N_{J-1}} \right\} R_{J-1} + e^{i\beta n \delta y / n_0} \left\{ \frac{N_{J-1} - N_J}{N_J + N_{J-1}} \right\} S_J \\ S_J &= 2e^{i\beta n \delta y / n_0} \left\{ \frac{N_{J+1}}{N_{J+1} + N_J} \right\} S_{J+1} + e^{i\beta n \delta y / n_0} \left\{ \frac{N_{J+1} - N_J}{N_{J+1} + N_J} \right\} R_J \end{aligned} \quad (12.4)$$

Here, the terms in curly brackets are just the Fresnel amplitude reflection and transmission coefficients and the exponential is a phase propagator which advances the phase of the \$R\$ and \$S\$ waves as they travel the distance \$\delta y\$ between discontinuities. We now let

$$X_{J-1} = X_J - \frac{dX}{dy} \delta y - \dots \quad (12.5)$$

* Note we have used the shorthand here that \$n = N_j\$.

and consider the limit $\delta y \rightarrow 0$. Further expanding the exponential terms as Taylor series and ignoring quadratic terms in δy , we arrive at the differential counterpart to Equation 12.4

$$\begin{aligned}\frac{dR}{dy} &= \frac{R}{2} \left(2i\beta \frac{n}{n_0} - \frac{1}{n} \frac{dn}{dy} \right) - \frac{1}{2n} \frac{dn}{dy} S \\ \frac{dS}{dy} &= -\frac{S}{2} \left(\frac{1}{n} \frac{dn}{dy} + 2i\beta \frac{n}{n_0} \right) - \frac{1}{2n} \frac{dn}{dy} R\end{aligned}\quad (12.6)$$

These equations are an *exact* representation of Maxwell's equations for an arbitrary index profile, $n(y)$ —as letting $u(y) = R(y) - S(y)$, we see that they simply reduce to the Helmholtz equation

$$\frac{d^2u}{dy^2} + \frac{\beta^2 n^2}{n_0^2} u = 0 \quad (12.7)$$

and the conservation of energy

$$\frac{d}{dy} (nR^* R - nS^* S) = 0 \quad (12.8)$$

When $dn/dy = 0$, Equations 12.6 describe two counterpropagating and non-interacting plane waves. A finite index gradient couples these waves.

We now make the transformation

$$\begin{aligned}R &\rightarrow R'(y) e^{i\beta y} \\ S &\rightarrow S'(y) e^{-i\beta y}\end{aligned}\quad (12.9)$$

where the primed quantities are slowly varying compared with $e^{i\beta y}$. Because they are slowly varying, we can write

$$\begin{aligned}\langle R' \rangle &\sim R \\ \langle S' \rangle &\sim S\end{aligned}\quad (12.10)$$

where the operator $\langle \rangle$ takes an average over several cycles of $e^{i\beta y}$. Substituting Equation 12.9 in Equation 12.6 and using Equation 12.10, we then arrive at the following differential equations

$$\begin{aligned}\frac{dR}{dy} &= -i\alpha\kappa S e^{2i\beta y(\alpha-1)} \\ \frac{dS}{dy} &= i\alpha\kappa R e^{-2i\beta y(\alpha-1)}\end{aligned}\quad (12.11)$$

where we have defined

$$\alpha = \frac{\lambda_c}{\lambda_r} \quad (12.12)$$

which is just the ratio of the replay wavelength to the recording wavelength. Introducing the pseudo-field,

$$\hat{S} = S e^{2\beta y(\alpha-1)} \quad (12.13)$$

and defining Kogelnik's constant,

$$\kappa = \frac{\pi n_1}{\lambda_c} \quad (12.14)$$

these equations may now be written in the form of Kogelnik's equations for the normal-incidence unslanted sinusoidal grating:

$$\begin{aligned} c_R \frac{dR}{dy} &= -i\kappa \hat{S} \\ c_S \frac{d\hat{S}}{dy} &= -i\vartheta \hat{S} - i\kappa R \end{aligned} \quad (12.15)$$

where

$$c_R = \frac{1}{\alpha}; c_S = -\frac{1}{\alpha}; \vartheta = 2\frac{\beta}{\alpha}(1-\alpha) \quad (12.16)$$

For comparison, Kogelnik's coefficients are

$$c_R = 1; c_S = (2\alpha - 1); \vartheta = 2\alpha\beta(1-\alpha) \quad (12.17)$$

By imposing boundary conditions appropriate for the reflection hologram,

$$\begin{aligned} R(y=0) &= 1 \\ \hat{S}(y=d) &= 0 \end{aligned} \quad (12.18)$$

where d is the grating thickness, Equations 12.15 may be solved analytically. We can then define the diffraction efficiency for both the PSM and Kogelnik models as

$$\begin{aligned} \eta &= \left| \frac{c_S}{c_R} \right| \hat{S}(0) \hat{S}^*(0) \\ &= \left\{ 1 - \frac{c_R c_S}{\kappa^2} \Upsilon^2 \operatorname{csh}^2(d\Upsilon) \right\}^{-1} \end{aligned} \quad (12.19)$$

where

$$\Upsilon^2 = -\frac{\vartheta^2}{4c_S^2} - \frac{\kappa^2}{c_R c_S} \quad (12.20)$$

Note that we should ensure that

$$d = m \left(\frac{\pi}{2\alpha\beta} \right) \quad (12.21)$$

where m is a non-zero integer to prevent a discontinuity in index at $y = d$ (see Moharam and Gaylord [9] for a detailed discussion of the starting and ending conditions of a grating).

For cases of practical interest for display and optical element holography, substitution of Equation 12.16 (the PSM coefficients) or Equation 12.17 (Kogelnik's coefficients) into Equations 12.19 and 12.20 yields very similar results. However, one should note that the only approximation made in deriving the PSM equations (Equations 12.15, 12.16, 12.19 and 12.20) has been that of Equation 12.10. This is an assumption that one would reasonably expect to hold in most gratings of interest. At Bragg resonance, when $\alpha = 1$, both the Kogelnik and PSM models reduce to

$$\eta = \tanh^2(\kappa d) \quad (12.22)$$

The PSM model provides a useful insight into what is happening within the grating: multiple reflections of the reference wave simply synthesise the signal wave by classical Fresnel reflection and transmission at

each infinitesimal discontinuity. This is a rigorous picture for the normal-incidence unslanted reflection grating because Equation 12.6 is an exact representation of Maxwell's equations. The fact that we explicitly need to introduce a "pseudo-field," \hat{S} , to get the PSM equations into the same form as Kogelnik's equations reminds us that indeed Kogelnik's signal wave is not the physical electric field of the signal wave for $\alpha \neq 1$. Kogelnik's theory models the dephasing away from Bragg resonance by letting the non-physical wave propagate differently from the physical signal wave. In the PSM analytical theory (Equations 12.15 and 12.16), the pseudo-field is also not the real electric field—but here, the transformations in Equations 12.10 and 12.13 make the relationship between the real and the pseudo-field perfectly clear.

12.2.1 Comparison with a Numerical Solution of the Helmholtz Equation

The Helmholtz equation (see Equation 12.7) for the normal-incidence reflection grating of Equation 12.1 can be written as

$$\frac{d^2u}{dy^2} - \gamma^2 u = 0 \quad (12.23)$$

where

$$\gamma^2 = -\beta^2 \left\{ 1 + \frac{n_1}{n_0} \cos \left(\frac{4\pi n_0}{\lambda_r} y \right) \right\}^2 \quad (12.24)$$

and where u is the tangential electric field. The corresponding boundary conditions to Equation 12.18 can be shown to be

$$\begin{aligned} u'(0) + i\beta u(0) &= 2i\beta \\ \beta u(d) &= -iu'(d) \end{aligned} \quad (12.25)$$

These equations, which constitute a rigorous solution to the diffraction problem at hand, can be conveniently solved on a personal computer (PC) using standard Runge–Kutta integration. Comparison of such numerical results offers a simple way to rate the analytical PSM and coupled-wave predictions of diffractive efficiency for gratings of interest.

In Figures 12.2 and 12.3, the diffraction efficiency is plotted versus replay wavelength for two cases with values typical to display holography. Figure 12.2 pertains to a grating recorded at 532 nm and Figure 12.3 pertains to a grating recorded at 660 nm. These graphs show that the PSM analytical model agrees closely with the numerical Helmholtz calculation for all replay wavelengths. On the other hand,

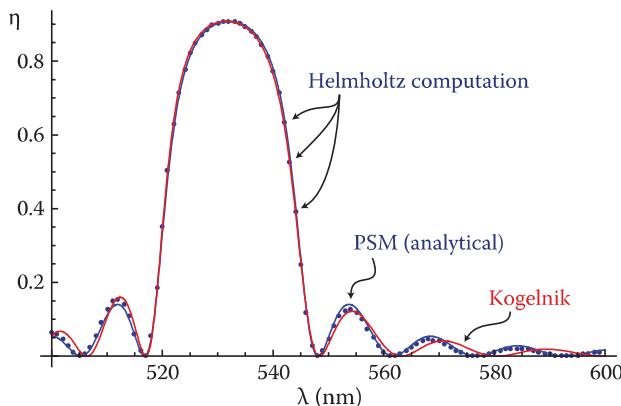


FIGURE 12.2 Comparison of the predictions of the PSM analytical model, Kogelnik's coupled wave theory and a Runge–Kutta numerical solution of the Helmholtz equation for a typical normal-incidence unslanted reflection grating with a thickness of 7 μm and with $n_0 = 1.5$, $n_1 = 0.045$. The grating was recorded at 532 nm.

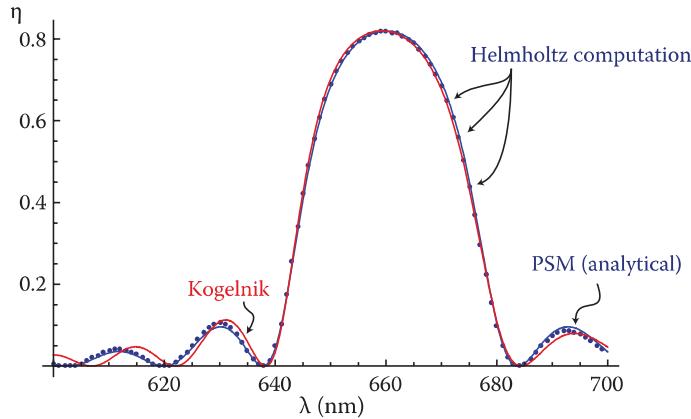


FIGURE 12.3 Comparison of the predictions of the PSM analytical model, Kogelnik's coupled wave theory and a numerical Runge–Kutta solution of the Helmholtz equation for identical parameters to Figure 12.2 with the exception that the grating was recorded here at 660 nm.

Kogelnik's coupled wave model appears to agree closely with the numerical calculation around Bragg resonance but does not predict an accurate behaviour of the sideband structure. In fact, this is broadly the behaviour seen over a wide range of parameters. As a general rule, Kogelnik's off-resonance solutions seem to become more accurate with lower modulations as one might expect. As would also be expected, a direct Runge–Kutta integration of the Helmholtz equation agrees within computational rounding errors to a similar numerical integration of Equation 12.6.

12.3 Unslanted Multiple-Colour Gratings at Normal Incidence

One of the advantages of the PSM model is that it does not limit the grating to a sinusoidal form. This is a great advantage over the simplest variants of standard coupled wave theories, including Kogelnik's.

We start by assuming a general (polychromatic unslanted) index profile

$$\begin{aligned} n &= n_0 + n_1 \cos(2\alpha_1 \beta y) + n_2 \cos(2\alpha_2 \beta y) + \dots \\ &= n_0 + \frac{n_1}{2} \left\{ e^{2i\beta\alpha_1 y} + e^{-2i\beta\alpha_1 y} \right\} + \frac{n_2}{2} \left\{ e^{2i\beta\alpha_2 y} + e^{-2i\beta\alpha_2 y} \right\} + \dots \end{aligned} \quad (12.26)$$

Equations 12.6 then reduce to the following form

$$\begin{aligned} \frac{dR}{dy} &= -S \sum_{j=1}^N i\kappa_j \alpha_j e^{2i\beta y(\alpha_j - 1)} \\ \frac{dS}{dy} &= R \sum_{j=1}^N i\kappa_j \alpha_j e^{-2i\beta y(\alpha_j - 1)} \end{aligned} \quad (12.27)$$

Assuming that the individual gratings have very different spatial frequencies, these equations lead to a simple expression for the diffractive efficiency when the reference wave is in Bragg resonance with one or another of the multiplexed gratings:

$$\eta_j = \tanh^2(\kappa_j d) \quad (12.28)$$

where

$$\kappa_j = \frac{n_j \pi}{\lambda_c} \quad (12.29)$$

In addition, in the region of the j th Bragg resonance, Equation 12.27 leads to the approximate analytical form

$$\eta_j = \frac{\alpha_j^2 \kappa_j^2}{\beta^2(1-\alpha_j)^2 + (\alpha_j^2 \kappa_j^2 - \beta^2(1-\alpha_j)^2) \coth^2 \left\{ d \sqrt{\alpha_j^2 \kappa_j^2 - \beta^2(1-\alpha_j)^2} \right\}} \quad (12.30)$$

When the spatial frequencies of the different gratings are too close to one another, these relations break down. For many cases of interest, however, Equations 12.28 through 12.30 provide a rather accurate picture of the normal-incidence polychromatic reflection phase grating. Indeed the following form can often be used to accurately describe an N -chromatic grating:

$$\eta = \sum_{j=1}^N \frac{\alpha_j^2 \kappa_j^2}{\beta^2(1-\alpha_j)^2 + (\alpha_j^2 \kappa_j^2 - \beta^2(1-\alpha_j)^2) \coth^2 \left\{ d \sqrt{\alpha_j^2 \kappa_j^2 - \beta^2(1-\alpha_j)^2} \right\}} \quad (12.31)$$

For example, Diehl and George [10] have used a sparse Hill's matrix technique to computationally calculate the diffraction efficiency of a lossless trichromatic phase reflection grating at normal incidence. They used free-space recording wavelengths of 400, 500 and 700 nm. The grating thickness was 25 μm and the index parameters were taken as $n_0 = 1.5$, $n_1 = n_2 = n_3 = 0.040533$. Comparison of Equation 12.31 with Diehl and George's published graphical results shows very good agreement.

In cases where the gratings are too close to one another in wavelength, Equation 12.6 or Equation 12.27 must be solved numerically.

12.3.1 Numerical Solution for Two-Colour Normal-Incidence Reflection Gratings

Equation 12.6 can be solved numerically using a standard Runge–Kutta method and the results compared with a similar numerical integration of the Helmholtz equation (Equation 12.23). If we compare these computational integrations with the PSM analytical expressions for diffractive efficiency, we see extremely good agreement.

This is illustrated in Figure 12.4, where diffraction efficiency versus replay wavelength is plotted for a two-colour grating recorded at 532 and 660 nm in a 7 μm emulsion. A peak index modulation of 0.045 was chosen as in Figures 12.2 and 12.3, and equal modulation was used for the two wavelengths. Comparison of Figures 12.2 and 12.3 with Figure 12.4 shows that the two-colour grating exhibits a somewhat lower diffractive efficiency than separate monochromatic gratings of an equivalent index modulation depth.

Figure 12.4 shows excellent agreement between the numerical PSM model and the direct Helmholtz integration as expected. In addition, the analytical expressions (Equations 12.28–12.31) are seen to

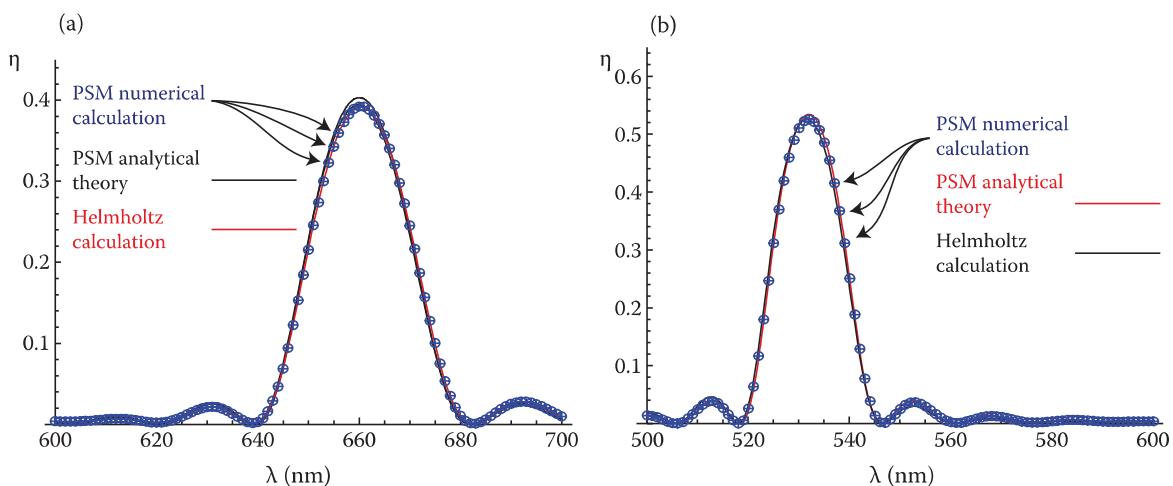


FIGURE 12.4 A normal-incidence unslanted reflection grating recorded using two wavelengths: 532 and 660 nm. A value of $n_1 = n_2 = 0.045/2$ was used giving a total index modulation of 0.045 ($n_0 = 1.5$). The grating thickness was 7 μm . (a) Replay near 660 nm and (b) replay near 532 nm.

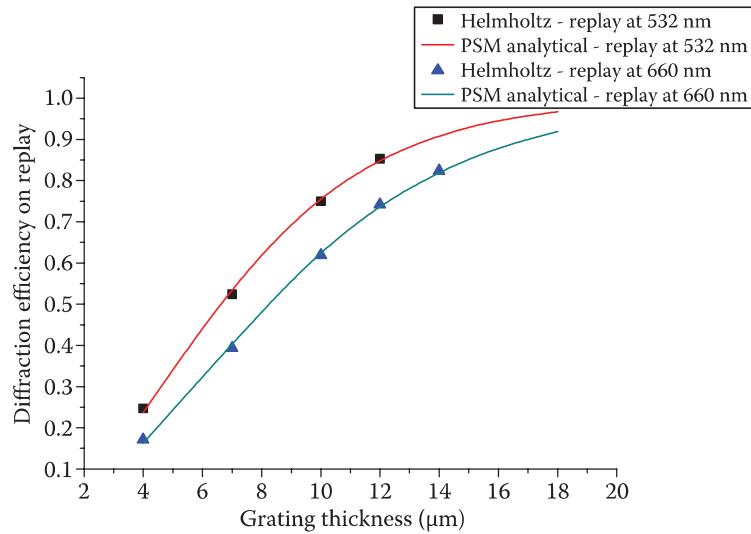


FIGURE 12.5 Diffraction efficiency versus grating thickness for a normal-incidence two-colour unslanted reflection grating. The grating was recorded at 660 and at 532 nm with equal modulations: $n_1 = n_2 = 0.045/2$ ($n_0 = 1.5$). The red and green curves show the analytical PSM prediction of Equation 12.28 for replay at 532 and 660 nm. The markers show numerical simulation results.

give very similar results. Calculations in which the higher 660 nm wavelength is lowered progressively towards the lower 532 nm wavelength show that the analytical expressions seem usefully accurate until the two diffractive peaks start to coalesce at approximately 555 nm.

To investigate whether the diffractive response of a two-colour reflective grating is always less than that of the corresponding one-colour grating with equivalent index modulation, we plot in Figure 12.5 a graph of the diffraction efficiencies at 532 and 660 nm against grating thickness. Clearly, at high thicknesses, two-colour gratings suffer rather less diffractive impairment. Of course, this does assume that the grating is lossless.

Finally, we should make the observation that the choice of the second recording wavelength affects the diffractive efficiency of a two-colour holographic grating quite significantly. Figure 12.6 shows a numerical calculation of the peak diffraction efficiency of a two-colour grating versus the second recording wavelength. This seems to imply that a proper choice of wavelengths may possibly optimise the diffractive response of colour holograms.

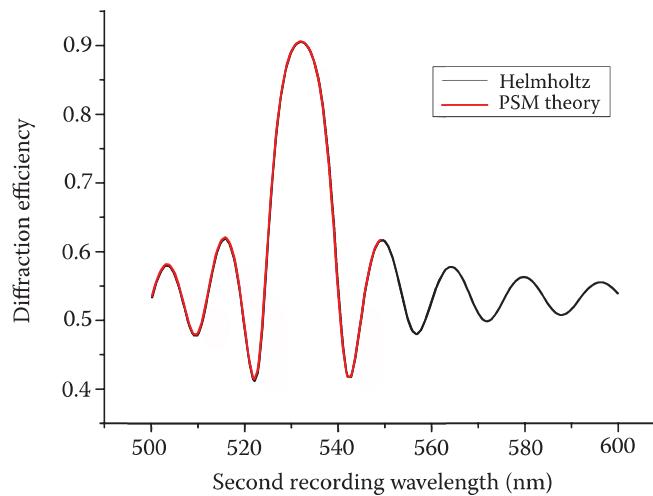


FIGURE 12.6 Two-colour normal-incidence unslanted reflection gratings have different diffractive efficiencies according to the exact choice of the wavelength pair. Here, a grating is recorded at 532 nm, and the peak 532 nm diffractive efficiency is plotted for different choices of the second recording wavelength. Equal modulation is used for the two recording wavelengths with $n_1 = n_2 = 0.045/2$ ($n_0 = 1.5$). Grating thickness is 7 μm .

12.3.2 Numerical Solution for Three-Colour Normal-Incidence Reflection Gratings

Figure 12.7 shows the case of an equal-modulation, three-colour normal-incidence unslanted reflection grating for the same total index modulation used in the previous graphs and for a grating thickness of 7 μm . Recording wavelengths of 440, 532 and 660 nm have been used. Clearly, the agreement between the numerical PSM model and the direct Helmholtz integration is excellent. Once again, the results are also described very closely by Equations 12.30 and 12.31.

With three colours, the diffractive efficiency falls substantially—particularly in the red spectral region—as compared with the monochromatic cases. Figure 12.8 shows a graph of the peak diffractive efficiency versus

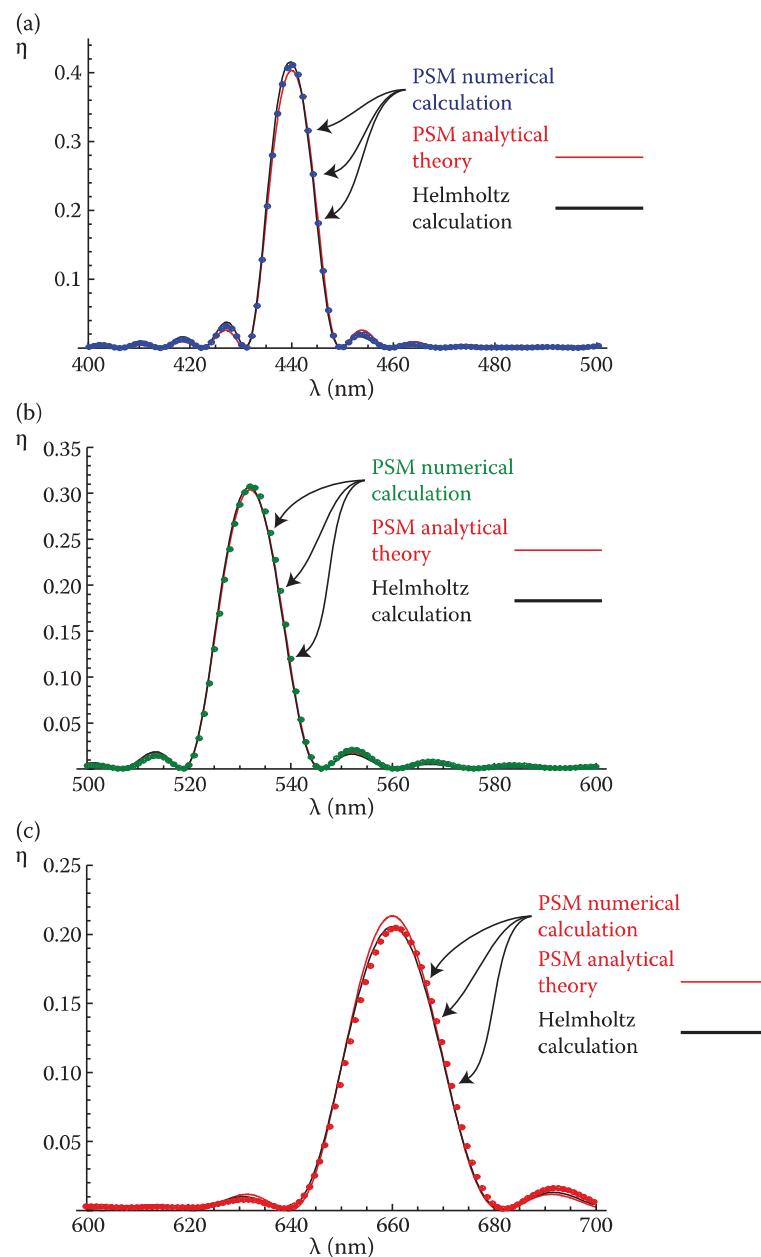


FIGURE 12.7 A normal-incidence unslanted reflection grating recorded using three wavelengths: 440, 532 and 660 nm. Values of $n_1 = n_2 = n_3 = 0.045/3$ were used, giving a total index modulation of 0.045 ($n_0 = 1.5$). The grating thickness is 7 μm . (a) Replay near 440 nm, (b) replay near 532 nm and (c) replay near 660 nm. The markers indicate a solution of the numerical PSM equations. The black lines represent a Runge–Kutta numerical integration of the Helmholtz equation. The red lines indicate the predictions of the PSM analytical theory.

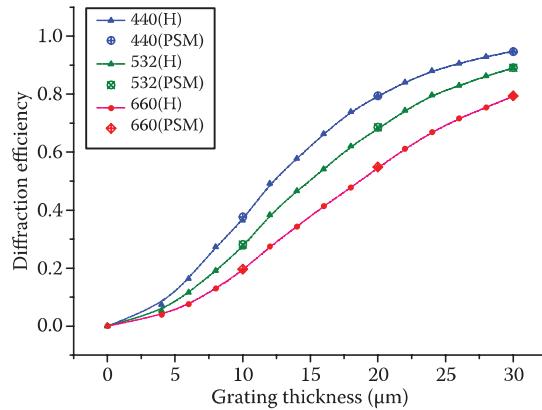


FIGURE 12.8 Peak diffraction efficiency for a three-colour normal-incidence unslanted reflection grating. The grating was recorded at 440, 532 and 660 nm with equal modulations: $n_1 = n_2 = n_3 = 0.03/3$ ($n_0 = 1.5$). The red, green and blue curves show the analytical PSM prediction of Equation 12.28 for replay at 440, 532 and 660 nm. The markers show numerical Runge–Kutta simulation results of the Helmholtz equation.

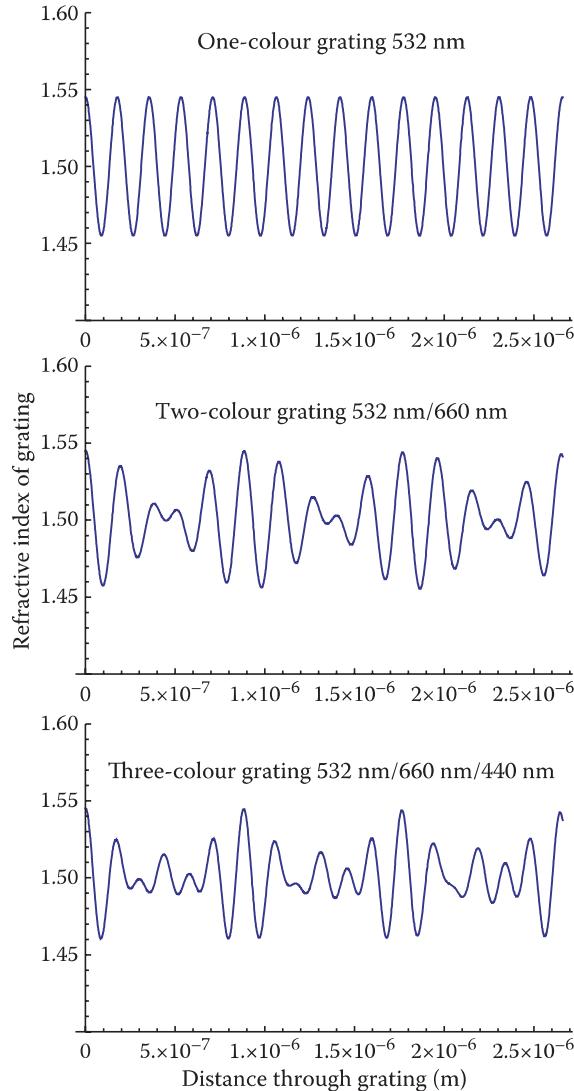


FIGURE 12.9 Index profiles used for the studies of one-, two- and three-colour holographic gratings with total modulation of 0.045.

grating thickness for a total index modulation of 0.03. This graph underlines the fact that thicker emulsions are required for bright, multicolour gratings. Figure 12.9 shows the index profiles used for the one-, two- and three-colour cases.

12.4 Unslanted Reflection Grating at Oblique Incidence

To treat the case of reference wave incidence at finite angle to the grating planes, we must redraw Figure 12.1 using two-dimensional fields, R and S , which we now endow with two indices instead of the previous single index (Figure 12.10). We shall make the approximation that the index modulation is small enough such that the rays of both the R and S waves are not deviated in angle. We shall, however, retain the proper Fresnel amplitude coefficients.

12.4.1 σ -Polarisation

The Fresnel amplitude coefficients for this polarisation may be written as

$$r_{k,k+1} = \frac{N_{k+1} \sqrt{1 - \frac{n_0^2}{N_{k+1}^2} \sin^2 \theta_c} - N_k \sqrt{1 - \frac{n_0^2}{N_k^2} \sin^2 \theta_c}}{N_{k+1} \sqrt{1 - \frac{n_0^2}{N_{k+1}^2} \sin^2 \theta_c} + N_k \sqrt{1 - \frac{n_0^2}{N_k^2} \sin^2 \theta_c}}$$

$$t_{k,k+1} = \frac{2N_k \sqrt{1 - \frac{n_0^2}{N_k^2} \sin^2 \theta_c}}{N_{k+1} \sqrt{1 - \frac{n_0^2}{N_{k+1}^2} \sin^2 \theta_c} + N_k \sqrt{1 - \frac{n_0^2}{N_k^2} \sin^2 \theta_c}} \quad (12.32)$$

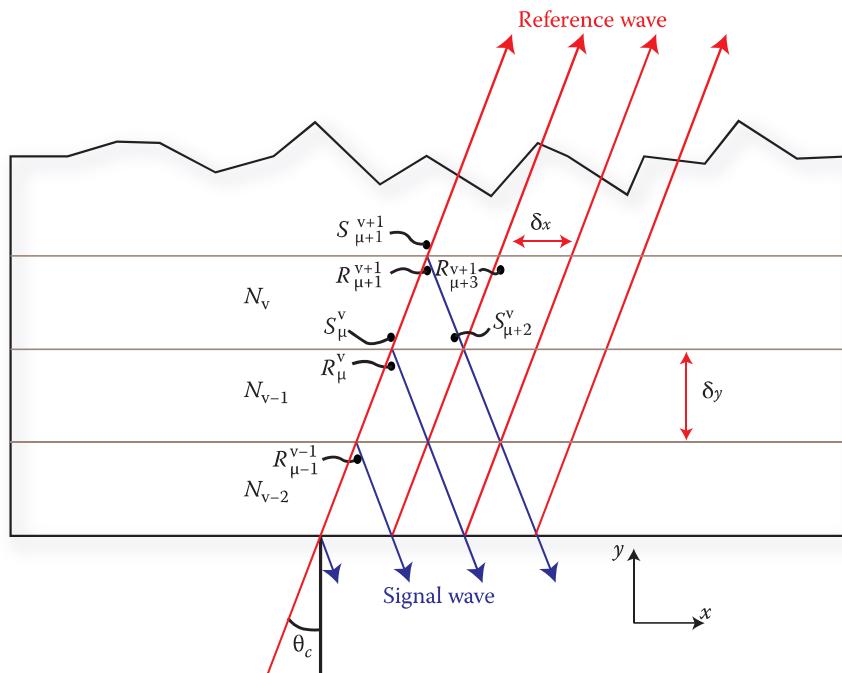


FIGURE 12.10 PSM model of the two-dimensional slanted reflection grating composed of layers of different refractive index. A plane reference wave is incident to the bottom boundary at angle θ_c . Fresnel transmission and reflection at the layer boundaries synthesise the signal or image wave.

where r and t pertain, respectively, to reflection and transmission occurring at the index discontinuity between layers k and $k + 1$ for a wave travelling from k to $k + 1$. The R and S waves in the exterior medium of index n_0 are assumed to be plane waves of the form

$$\begin{aligned} R &= e^{i(k_{cx}x+k_{cy}y)} \\ S &= S_0 e^{i(k_{ix}x+k_{iy}y)} \end{aligned} \quad (12.33)$$

where S_0 is a constant. Within the grating, we shall assume that R and S are functions of x and y . Using the normal rules of Fresnel reflection, the wave vectors can be written explicitly as

$$\begin{aligned} \mathbf{k}_c &= \beta \begin{pmatrix} \sin \theta_c \\ \cos \theta_c \end{pmatrix} \\ \mathbf{k}_i &= \beta \begin{pmatrix} \sin \theta_c \\ -\cos \theta_c \end{pmatrix} \end{aligned} \quad (12.34)$$

where the angle θ_c is the angle of incidence of the R wave.

We can now use Figure 12.10 to write down an expression relating the discrete values of R and S within the grating:

$$\begin{aligned} R_{\mu+1}^{v+1} &= e^{i\beta n(\sin \theta_c \delta x + \cos \theta_c \delta y)/n_0} R_\mu^v \left\{ \frac{2N_{v-1} \sqrt{1 - \frac{n_0^2}{N_{v-1}^2} \sin^2 \theta_c}}{N_{v-1} \sqrt{1 - \frac{n_0^2}{N_{v-1}^2} \sin^2 \theta_c} + N_v \sqrt{1 - \frac{n_0^2}{N_v^2} \sin^2 \theta_c}} \right\} \\ &\quad + e^{i\beta n(\sin \theta_c \delta x + \cos \theta_c \delta y)/n_0} S_\mu^v \left\{ \frac{N_{v-1} \sqrt{1 - \frac{n_0^2}{N_{v-1}^2} \sin^2 \theta_c} - N_v \sqrt{1 - \frac{n_0^2}{N_v^2} \sin^2 \theta_c}}{N_{v-1} \sqrt{1 - \frac{n_0^2}{N_{v-1}^2} \sin^2 \theta_c} + N_v \sqrt{1 - \frac{n_0^2}{N_v^2} \sin^2 \theta_c}} \right\} \end{aligned} \quad (12.35)$$

$$\begin{aligned} S_{\mu+1}^{v-1} &= e^{i\beta n(\sin \theta_c \delta x + \cos \theta_c \delta y)/n_0} S_\mu^v \left\{ \frac{2N_v \sqrt{1 - \frac{n_0^2}{N_v^2} \sin^2 \theta_c}}{N_{v-1} \sqrt{1 - \frac{n_0^2}{N_{v-1}^2} \sin^2 \theta_c} + N_v \sqrt{1 - \frac{n_0^2}{N_v^2} \sin^2 \theta_c}} \right\} \\ &\quad + e^{i\beta n(\sin \theta_c \delta x + \cos \theta_c \delta y)/n_0} R_\mu^v \left\{ \frac{N_v \sqrt{1 - \frac{n_0^2}{N_v^2} \sin^2 \theta_c} - N_{v-1} \sqrt{1 - \frac{n_0^2}{N_{v-1}^2} \sin^2 \theta_c}}{N_{v-1} \sqrt{1 - \frac{n_0^2}{N_{v-1}^2} \sin^2 \theta_c} + N_v \sqrt{1 - \frac{n_0^2}{N_v^2} \sin^2 \theta_c}} \right\} \end{aligned} \quad (12.36)$$

Because we are assuming that δx and δy are small, we can use Taylor expansions for the fields and index profile

$$\begin{aligned} R_{\mu+1}^{v+1} &= R_\mu^v + \frac{\partial R_\mu^v}{\partial x} \delta x + \frac{\partial R_\mu^v}{\partial y} \delta y + \dots \\ S_{\mu+1}^{v-1} &= S_\mu^v + \frac{\partial S_\mu^v}{\partial x} \delta x - \frac{\partial S_\mu^v}{\partial y} \delta y + \dots \\ N_{v-1} &= N_v - \frac{\partial N_v}{\partial y} \delta y + \dots \end{aligned} \quad (12.37)$$

The exponentials are also written using a Taylor expansion. Before substituting these expressions in Equations 12.35 and 12.36, we make the following additional approximations

$$\sqrt{1 - \frac{n_0^2}{N_v^2} \sin^2 \theta_c} = \sqrt{1 - \frac{a}{N_v^2}} \sim \cos \theta_c = b \quad (12.38)$$

$$\sqrt{1 - \frac{n_0^2}{N_{v-1}^2} \sin^2 \theta_c} \sim b - \frac{\partial N_v}{\partial y} \frac{a \delta y}{b N_v^3} + O(\delta y^2) \quad (12.39)$$

Then, letting

$$R_\mu^v \rightarrow R; S_\mu^v \rightarrow S; N_v \rightarrow n \quad (12.40)$$

Equation 12.35 becomes

$$\begin{aligned} &\left\{ R + \frac{\partial R}{\partial x} \delta x + \frac{\partial R}{\partial y} \delta y \right\} \left\{ \left(n - \frac{\partial n}{\partial y} \delta y \right) \left(b - \frac{\partial n}{\partial y} \frac{a \delta y}{b n^3} \right) + nb \right\} \\ &= (1 + i\beta \sin \theta_c \delta x + i\beta \cos \theta_c \delta y) R \left\{ 2 \left(n - \frac{\partial n}{\partial y} \delta y \right) \left(b - \frac{\partial n}{\partial y} \frac{a \delta y}{b n^3} \right) \right\} \\ &\quad + (1 + i\beta \sin \theta_c \delta x + i\beta \cos \theta_c \delta y) S \left\{ \left(n - \frac{\partial n}{\partial y} \delta y \right) \left(b - \frac{\partial n}{\partial y} \frac{a \delta y}{b n^3} \right) - nb \right\} \end{aligned} \quad (12.41)$$

Ignoring quadratic terms such as δx^2 , δy^2 and $\delta x \delta y$ and letting $n\beta/n_o \rightarrow \beta(n) = \beta$, this expression simplifies to

$$\begin{aligned} &\frac{\partial R}{\partial x} \delta x + \frac{\partial R}{\partial y} \delta y = \\ &-R \frac{1}{2n} \frac{\partial n}{\partial y} \delta y \left(1 + \frac{a}{b^2 n^2} \right) + i\beta R (\sin \theta_c \delta x + \cos \theta_c \delta y) - S \delta y \frac{1}{2n} \frac{\partial n}{\partial y} \left(1 + \frac{a}{b^2 n^2} \right) \end{aligned} \quad (12.42)$$

Letting $\delta x, \delta y \rightarrow 0$, we arrive at a partial differential equation for the R field:

$$\frac{\mathbf{k}_c \cdot \nabla R}{\beta} = \sin \theta_c \frac{\partial R}{\partial x} + \cos \theta_c \frac{\partial R}{\partial y} = \frac{R}{2} \left\{ 2i\beta - \frac{1}{n \cos \theta_c} \frac{\partial n}{\partial y} \right\} - \frac{S}{2n \cos \theta_c} \frac{\partial n}{\partial y} \quad (12.43)$$

Equation 12.36 may be treated in exactly the same manner, yielding a corresponding partial differential equation for the S field:

$$\frac{\mathbf{k}_i \cdot \nabla S}{\beta} = \sin \theta_c \frac{\partial S}{\partial x} - \cos \theta_c \frac{\partial S}{\partial y} = \frac{S}{2} \left\{ 2i\beta + \frac{1}{n \cos \theta_c} \frac{\partial n}{\partial y} \right\} + \frac{R}{2n \cos \theta_c} \frac{\partial n}{\partial y} \quad (12.44)$$

Note the similarity of Equations 12.43 and 12.44 to Equation 12.6. Note also that if we set $\theta_c = 0$, then we retrieve Equation 12.6 exactly. Equations 12.43 and 12.44 are the PSM equations for an unslanted reflection grating at oblique incidence with the σ -polarisation. In the following sections, we shall use these equations to provide useful analytical expressions for both unslanted and slanted gratings. However, before proceeding, we shall briefly derive the PSM equations appropriate for the π -polarisation.

12.4.2 π -Polarisation

The Fresnel amplitude coefficients for this polarisation may be written as

$$r_{k,k+1} = \frac{N_{k+1} \sqrt{1 - \frac{n_0^2}{N_{k+1}^2} \sin^2 \theta_c} - N_k \sqrt{1 - \frac{n_0^2}{N_k^2} \sin^2 \theta_c}}{N_{k+1} \sqrt{1 - \frac{n_0^2}{N_{k+1}^2} \sin^2 \theta_c} + N_k \sqrt{1 - \frac{n_0^2}{N_k^2} \sin^2 \theta_c}} \quad (12.45)$$

$$t_{k,k+1} = \frac{2N_k \sqrt{1 - \frac{n_0^2}{N_k^2} \sin^2 \theta_c}}{N_{k+1} \sqrt{1 - \frac{n_0^2}{N_{k+1}^2} \sin^2 \theta_c} + N_k \sqrt{1 - \frac{n_0^2}{N_k^2} \sin^2 \theta_c}}$$

The analogues of Equations 12.35 and 12.36 then become

$$\left\{ N_{v-1} + N_v \left(1 - \frac{\partial N_v}{\partial y} \frac{a\delta y}{b^2 N_v^3} \right) \right\} R_{\mu+1}^{v+1} = 2N_{v-1} e^{i\beta \sin \theta_c \delta x + i\beta \cos \theta_c \delta y} R_\mu^v + e^{i\beta \sin \theta_c \delta x + i\beta \cos \theta_c \delta y} S_\mu^v \left\{ N_{v-1} - N_v \left(1 - \frac{\partial N_v}{\partial y} \frac{a\delta y}{b^2 N_v^3} \right) \right\} \quad (12.46)$$

$$\left\{ N_{v-1} + N_v \left(1 - \frac{\partial N_v}{\partial y} \frac{a\delta y}{b^2 N_v^3} \right) \right\} S_{\mu+1}^{v-1} = e^{i\beta \sin \theta_c \delta x + i\beta \cos \theta_c \delta y} S_\mu^v \left\{ 2N_v \left(1 - \frac{\partial N_v}{\partial y} \frac{a\delta y}{b^2 N_v^3} \right) \right\} + e^{i\beta \sin \theta_c \delta x + i\beta \cos \theta_c \delta y} R_\mu^v \left\{ N_v \left(1 - \frac{\partial N_v}{\partial y} \frac{a\delta y}{b^2 N_v^3} \right) - N_{v-1} \right\} \quad (12.47)$$

Following our previous analysis, we can then derive the corresponding partial differential equations for the π -polarisation:

$$\frac{\mathbf{k}_e \cdot \nabla R}{\beta} = \sin \theta_c \frac{\partial R}{\partial x} + \cos \theta_c \frac{\partial R}{\partial y} = \frac{R}{2} \left\{ 2i\beta - \frac{1}{n} \frac{\cos 2\theta_c}{\cos \theta_c} \frac{\partial n}{\partial y} \right\} - \frac{S}{2n} \frac{\cos 2\theta_c}{\cos \theta_c} \frac{\partial n}{\partial y} \quad (12.48)$$

$$\frac{\mathbf{k}_i \cdot \nabla S}{\beta} = \sin \theta_c \frac{\partial S}{\partial x} - \cos \theta_c \frac{\partial S}{\partial y} = \frac{S}{2} \left\{ 2i\beta + \frac{1}{n} \frac{\cos 2\theta_c}{\cos \theta_c} \frac{\partial n}{\partial y} \right\} + \frac{R}{2n} \frac{\cos 2\theta_c}{\cos \theta_c} \frac{\partial n}{\partial y} \quad (12.49)$$

12.4.3 Simplification of PSM Equations to Ordinary Differential Equations

The PSM equations may be simplified under boundary conditions corresponding to monochromatic illumination of the grating.

Let

$$\begin{aligned} R &\rightarrow R(y) e^{i\beta \sin \theta_c x} \\ S &\rightarrow S(y) e^{i\beta \sin \theta_c x} \end{aligned} \quad (12.50)$$

Under this transformation, Equations 12.43 and 12.44 yield the following pair of ordinary differential equations

$$\begin{aligned} \cos \theta_c \frac{dR}{dy} &= \frac{R}{2} \left\{ 2i\beta \cos^2 \theta_c - \frac{1}{n \cos \theta_c} \frac{dn}{dy} \right\} - \frac{S}{2} \left\{ \frac{1}{n \cos \theta_c} \frac{dn}{dy} \right\} \\ -\cos \theta_c \frac{dS}{dy} &= \frac{S}{2} \left\{ 2i\beta \cos^2 \theta_c + \frac{1}{n \cos \theta_c} \frac{dn}{dy} \right\} + \frac{R}{2} \left\{ \frac{1}{n \cos \theta_c} \frac{dn}{dy} \right\} \end{aligned} \quad (12.51)$$

Similarly, the π -polarisation equations yield

$$\begin{aligned} \cos \theta_c \frac{dR}{dy} &= \frac{R}{2} \left\{ 2i\beta \cos^2 \theta_c - \frac{\cos 2\theta_c}{n \cos \theta_c} \frac{dn}{dy} \right\} - \frac{S}{2} \left\{ \frac{\cos 2\theta_c}{n \cos \theta_c} \frac{dn}{dy} \right\} \\ -\cos \theta_c \frac{dS}{dy} &= \frac{S}{2} \left\{ 2i\beta \cos^2 \theta_c + \frac{\cos 2\theta_c}{n \cos \theta_c} \frac{dn}{dy} \right\} + \frac{R}{2} \left\{ \frac{\cos 2\theta_c}{n \cos \theta_c} \frac{dn}{dy} \right\} \end{aligned} \quad (12.52)$$

Equations 12.51 and 12.52 are approximate only because we have assumed an approximate form for the direction vector of the waves within the grating. We may, however, approach the problem differently and derive exact equations directly from Equation 12.6. For example, in the case of the σ -polarisation, we use the optical invariant

$$\tilde{\beta}(y) \rightarrow \tilde{\beta}(y) \cos \theta(y) \quad (12.53)$$

where

$$\tilde{\beta}(y) = \frac{\beta n(y)}{n_0} \quad (12.54)$$

Then using Snell's law

$$\frac{d\tilde{\beta}(y)}{dy} \sin \theta(y) + \tilde{\beta}(y) \frac{d\theta(y)}{dy} \cos \theta(y) = 0 \quad (12.55)$$

it is simple to see that Equation 12.6 reduces to

$$\begin{aligned}\cos\theta \frac{dR}{dy} &= \frac{R}{2} \left\{ 2i\tilde{\beta} \cos^2 \theta - \frac{1}{\tilde{\beta} \cos\theta} \frac{d\tilde{\beta}}{dy} \right\} - \frac{S}{2} \left\{ \frac{1}{\tilde{\beta} \cos\theta} \frac{d\tilde{\beta}}{dy} \right\} \\ -\cos\theta \frac{dS}{dy} &= \frac{S}{2} \left\{ 2i\tilde{\beta} \cos^2 \theta + \frac{1}{\tilde{\beta} \cos\theta} \frac{d\tilde{\beta}}{dy} \right\} + \frac{R}{2} \left\{ \frac{1}{\tilde{\beta} \cos\theta} \frac{d\tilde{\beta}}{dy} \right\}\end{aligned}\quad (12.56)$$

where θ is now a function of y throughout the grating. If we now replace Equation 12.50 with the more general behaviour

$$\begin{aligned}R &\rightarrow R(y) e^{i\tilde{\beta}(y)\sin\theta(y)x} \\ S &\rightarrow S(y) e^{i\tilde{\beta}(y)\sin\theta(y)x}\end{aligned}\quad (12.57)$$

then Equation 12.56 can be seen to be an exact solution of the Helmholtz equation. It follows therefore that solution of Equations 12.55 and 12.56, subject to the boundary conditions of Equation 12.18 and $\theta(0) = \theta_c$, constitutes a rigorous solution of the Helmholtz equation. Note that this is independent of periodicity required by a Floquet solution. Because these equations are none other than a differential representation of the chain matrix method of thin films [6,7], it is simple to show that this implies that the chain matrix method is itself rigorous.

Equations 12.55 and 12.56, or alternatively the approximate Equations 12.51 and 12.52, are simple to solve on a PC for arbitrary index profiles. However, Equations 12.51 and 12.52 also possess analytic solutions of interest.

12.4.4 Analytic Solutions for Sinusoidal Gratings

We start by defining an unslanted grating with the following index profile

$$n = n_0 + n_1 \cos(2\alpha\beta \cos\theta_r y) = n_0 + \frac{n_1}{2} \left\{ e^{2i\alpha\beta \cos\theta_r y} + e^{-2i\alpha\beta \cos\theta_r y} \right\} \quad (12.58)$$

where we imagine θ_r to be the recording angle of this grating. Then letting

$$\begin{aligned}R &\rightarrow R(y) e^{i\beta \cos\theta_c y} \\ S &\rightarrow S(y) e^{-i\beta \cos\theta_c y}\end{aligned}\quad (12.59)$$

and using Equation 12.10, Equation 12.51 reduces to

$$\begin{aligned}\cos\theta_c \frac{dR}{dy} &= -\frac{1}{2n_0} n_1 i\beta \alpha \frac{\cos\theta_r}{\cos\theta_c} \left\langle \left\{ e^{2i\beta \cos\theta_c y} + \dots \right\} S e^{-2i\beta \cos\theta_c y} \right\rangle \\ &= -\frac{n_1 i\beta (\alpha \cos\theta_r)}{2n_0 \cos\theta_c} S e^{2i\beta y(\alpha \cos\theta_r - \cos\theta_c)} \\ \cos\theta_c \frac{dS}{dy} &= \frac{1}{2n_0} n_1 i\beta \alpha \frac{\cos\theta_r}{\cos\theta_c} \left\langle \left\{ e^{-2i\beta \cos\theta_c y} + \dots \right\} R e^{2i\beta \cos\theta_c y} \right\rangle \\ &= \frac{n_1 i\beta (\alpha \cos\theta_r)}{2n_0 \cos\theta_c} R e^{-2i\beta y(\alpha \cos\theta_r - \cos\theta_c)}\end{aligned}\quad (12.60)$$

As before, we now define the pseudo-field

$$\hat{S} = S e^{2j\beta_y(\alpha \cos \theta_r - \cos \theta_c)} \quad (12.61)$$

at which point Equations 12.60 reduce to the standard form of Kogelnik's equations

$$\begin{aligned} c_R \frac{dR}{dy} &= -i\kappa \hat{S} \\ c_S \frac{d\hat{S}}{dy} &= -i\vartheta \hat{S} - i\kappa R \end{aligned} \quad (12.62)$$

The coefficients for the PSM model and for Kogelnik's model are as follows:

$$\begin{aligned} c_{R(PSM)} &= \frac{\cos^2 \theta_c}{\alpha \cos \theta_r} & c_{R(KOG)} &= \cos \theta_c \\ c_{S(PSM)} &= -\frac{\cos^2 \theta_c}{\alpha \cos \theta_r} & c_{S(KOG)} &= \cos \theta_c - 2\alpha \cos \theta_r \\ \vartheta_{PSM} &= 2\beta \left(1 - \frac{\cos \theta_c}{\alpha \cos \theta_r} \right) \cos^2 \theta_c & \vartheta_{KOG} &= 2\alpha \beta \cos \theta_r (\cos \theta_c - \alpha \cos \theta_r) \end{aligned} \quad (12.63)$$

Equation 12.62, in conjunction with the boundary conditions in Equation 12.18, then leads to the general analytic expression for the diffractive efficiency of the unslanted reflection grating:

$$\eta_\sigma = \frac{|c_S|}{c_R} \hat{S}(0) \hat{S}^*(0) = \frac{\kappa^2 \sinh^2(d\Upsilon)}{\kappa^2 \sinh^2(d\Upsilon) - c_R c_S \Upsilon^2} \quad (12.64)$$

where

$$\Upsilon^2 = -\frac{\vartheta^2}{4c_S^2} - \frac{\kappa^2}{c_R c_S} \quad (12.65)$$

Note that at Bragg resonance, both the PSM theory and Kogelnik's theory reduce to the well-known formula

$$\eta_\sigma = \tanh^2(\kappa d \sec \theta_c) \quad (12.66)$$

The π -polarisation may be treated in an exactly analogous way, leading to the following pair of ordinary differential equations for R and \hat{S} :

$$\begin{aligned} c_R \frac{dR}{dy} &= -i\kappa \cos 2\theta_c \hat{S} \\ c_S \frac{d\hat{S}}{dy} &= -i\vartheta \hat{S} - i\kappa \cos 2\theta_c R \end{aligned} \quad (12.67)$$

These are just Kogelnik's equations with a modified κ parameter. The PSM model therefore distinguishes the π - and σ -polarisations in exactly the same manner as Kogelnik's theory does! In both theories, in the case of the unslanted grating, Kogelnik's constant is simply transformed according to the rule

$$|\kappa| \rightarrow |\kappa \cos 2\theta_c| \quad (12.68)$$

We shall see shortly that the practical predictions of Kogelnik's model and the PSM model are very close for gratings of interest to display and optical element holography. This is partly due to the effect

of Snell's law, which acts to steepen the angle of incidence in most situations. At very high angles of incidence within the grating, larger differences appear.

12.4.5 Multiple Colour Gratings

Let

$$\begin{aligned} n &= n_0 + n_1 \cos(2\alpha_1 \beta \cos \theta_{r1} y) + n_2 \cos(2\alpha_2 \beta \cos \theta_{r2} y) + \dots \\ &= n_0 + \frac{n_1}{2} \left\{ e^{2i\alpha_1 \beta \cos \theta_{r1} y} + e^{-2i\alpha_1 \beta \cos \theta_{r1} y} \right\} + \frac{n_2}{2} \left\{ e^{2i\alpha_2 \beta \cos \theta_{r2} y} + e^{-2i\alpha_2 \beta \cos \theta_{r2} y} \right\} + \dots \\ &= n_0 + \frac{1}{2} \sum_{j=1}^N n_j \left\{ e^{2i\alpha_j \beta \cos \theta_{rj} y} + e^{-2i\alpha_j \beta \cos \theta_{rj} y} \right\} \end{aligned} \quad (12.69)$$

In this case, the PSM σ -polarisation equations yield

$$\begin{aligned} \cos \theta_c \frac{dR}{dy} &= -S \sum_{j=1}^N i \kappa_j \alpha_j \frac{\cos \theta_{rj}}{\cos \theta_c} e^{2i\beta y (\alpha_j \cos \theta_{rj} - \cos \theta_c)} \\ \cos \theta_c \frac{dS}{dy} &= R \sum_{j=1}^N i \kappa_j \alpha_j \frac{\cos \theta_{rj}}{\cos \theta_c} e^{-2i\beta y (\alpha_j \cos \theta_{rj} - \cos \theta_c)} \end{aligned} \quad (12.70)$$

where

$$\kappa_j = \frac{n_j \pi}{\lambda_c} \quad (12.71)$$

Once again, as in Section 12.3, if we assume that the individual gratings have very different spatial frequencies, then these equations lead to a simple expression for the diffractive efficiency when the reference wave is in Bragg resonance with one or another of the multiplexed gratings:

$$\eta_{PSM/\sigma_j} = \tanh^2(\kappa_j d \sec \theta_c) \quad (12.72)$$

The corresponding result for the π -polarisation is

$$\eta_{PSM/\pi_j} = \tanh^2(\kappa_j d \sec \theta_c \cos 2\theta_c) \quad (12.73)$$

If we recalculate Figures 12.5 and 12.8 using Equation 12.72, for typical values of θ_c , we see very little difference in the plots. The effect for the σ -polarisation is usually to increase the diffractive response of the grating by a small amount. The π -polarisation generally produces an inverse behaviour, at least at small angles of incidence.

As in Section 12.3 in the region of the j th Bragg resonance, Equations 12.70 lead to the approximate analytical form

$$\left. \begin{aligned} \eta_{\sigma j} &= \frac{\kappa_j^2 \sinh^2(dY_{\sigma j})}{\kappa_j^2 \sinh^2(dY_{\sigma j}) - c_R c_S Y_{\sigma j}^2} \\ Y_{\sigma j}^2 &= -\frac{\vartheta_{\sigma j}^2}{4c_S^2} - \frac{\kappa_j^2}{c_R c_S} \\ \vartheta_{\sigma j} &= 2\beta \left(1 - \frac{\cos \theta_c}{\alpha_j \cos \theta_r} \right) \cos^2 \theta_c \end{aligned} \right\} \quad (12.74)$$

where c_R and c_S are given by $c_{R(PSM)}$ and $c_{S(PSM)}$ of Equation 12.63.

Again, as long as there is a sufficient difference in the spatial frequencies of each grating, we can add each response to give a convenient analytical expression for the total diffraction efficiency:

$$\eta_\sigma = \sum_{j=1}^N \frac{\kappa_j^2 \sinh^2(dY_{\sigma j})}{\kappa_j^2 \sinh^2(dY_{\sigma j}) - c_R c_S Y_{\sigma j}^2} \quad (12.75)$$

In cases in which the individual gratings are too close to one another in wavelength or in which small amplitude interaction effects between gratings are to be described, Equation 12.70 must be solved numerically.

12.4.6 Comparison of Kogelnik's Theory with the PSM Theory for Unslanted Gratings at Oblique Incidence

In Figures 12.11 and 12.12, we plot the diffractive efficiency (σ -polarisation) against replay wavelength and replay angle as predicted by both Kogelnik's theory and by the PSM analytical theory for typical sinusoidal gratings of interest to display and optical element holography. It is immediately clear that the two theories produce very similar predictions.

In Figure 12.13, results are presented from a rigorous chain matrix computation (see Section 11.12.2) for the π -polarisation case at finite incidence angle. These are compared with predictions of the analytical Kogelnik and PSM models. The chain matrix algorithm used is based on the technique described by Heavens [7], and models the grating as a stack of 100,000 layers. As can be seen, in this case, the chain matrix computation is somewhat closer to the PSM model than Kogelnik's, although the differences

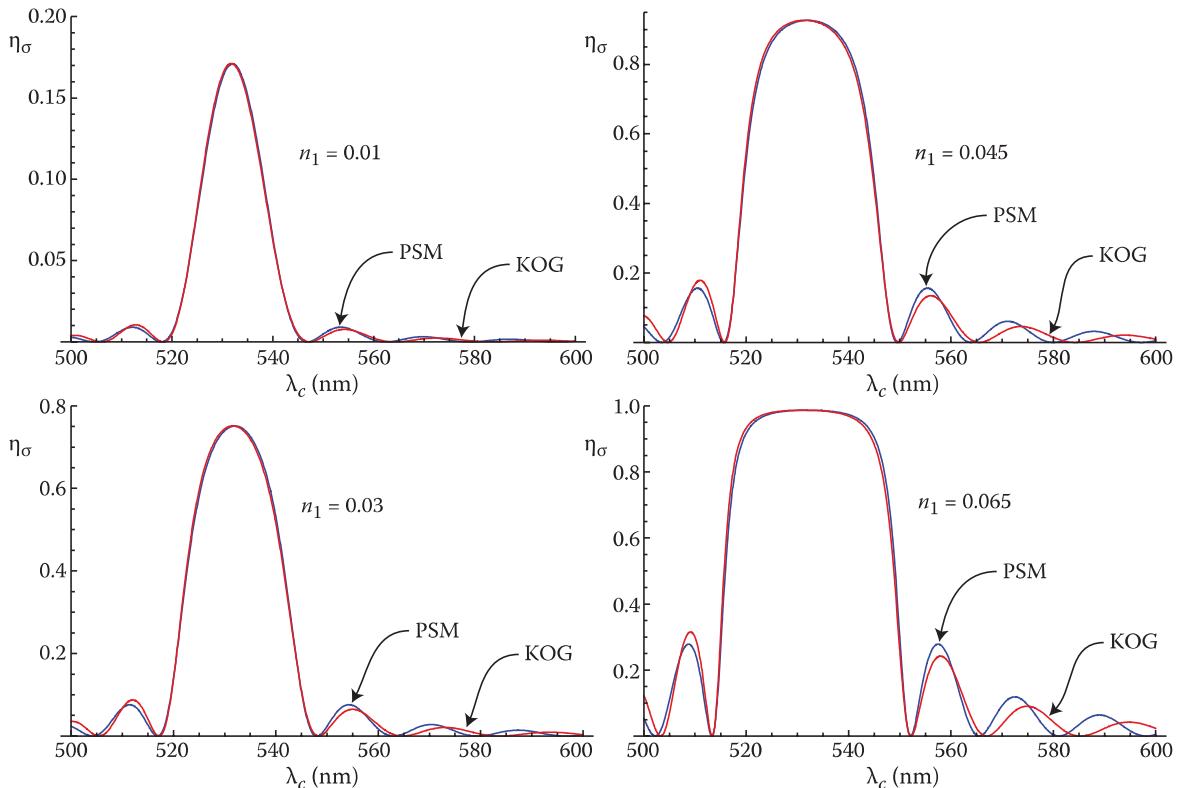


FIGURE 12.11 Four graphs showing the predicted diffractive response (σ -polarisation) versus replay wavelength of typical unslanted single-colour reflection gratings using the PSM analytical model (blue lines) and Kogelnik's theory (red lines). Recording wavelength, $\lambda_r = 532$ nm; recording angle, $\theta_r = 20^\circ$; replay angle, $\theta_c = 20^\circ$; grating thickness, $d = 7 \mu\text{m}$; $n_0 = 1.5$; index modulations shown on graphs. Note that all angles are interior angles.

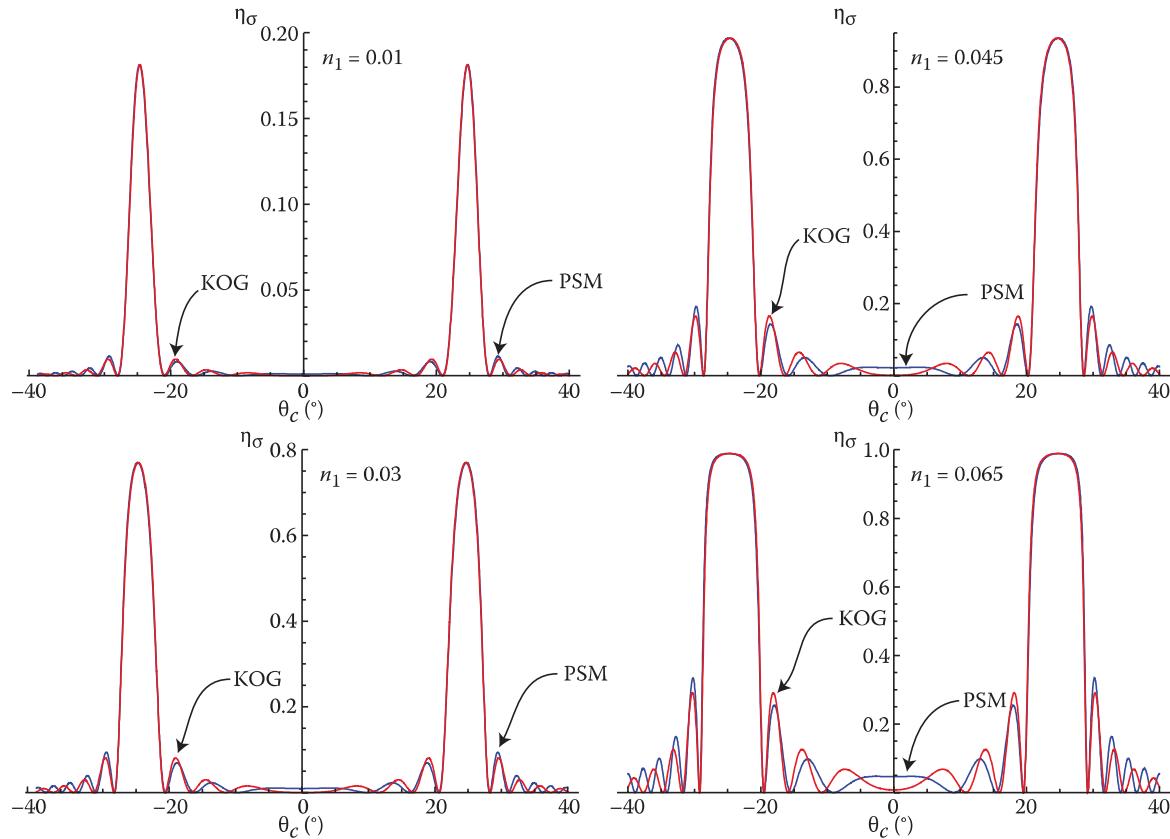


FIGURE 12.12 Four graphs showing the predicted diffractive response (σ -polarisation) versus reconstruction angle of typical unslanted single-colour reflection gratings using the PSM analytical model (blue lines) and Kogelnik's theory (red lines). Recording wavelength, $\lambda_r = 550$ nm; reconstruction wavelength, $\lambda_c = 532$ nm; recording angle, $\theta_r = 20^\circ$; grating thickness, $d = 7 \mu\text{m}$; $n_0 = 1.5$; index modulations shown on graphs. Note that all angles are interior angles.

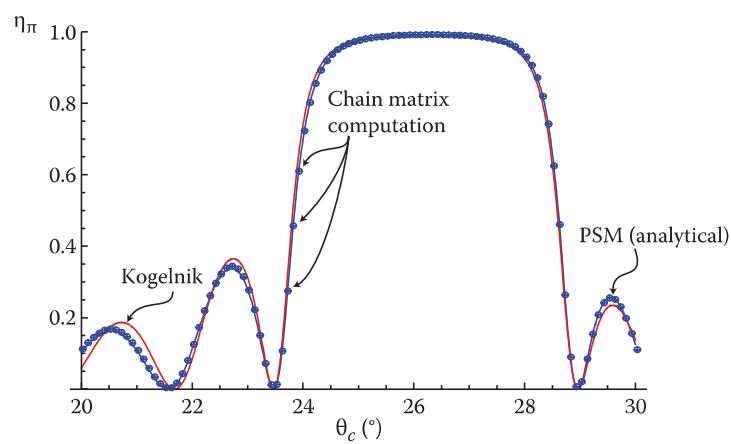


FIGURE 12.13 Comparison of the PSM (analytic) and Kogelnik models with a rigorous chain matrix computation for the π -polarisation diffractive efficiency versus replay angle, θ_c . For the purpose of the numerical calculation, the grating index profile has been modelled using 100,000 discrete thin films. The 12 μm unslanted reflection grating ($n_0 = 1.5$; $n_1 = 0.065$) has been recorded at $\lambda_r = 532$ nm with a recording angle of $\theta_r = 30^\circ$. The replay wavelength is $\lambda_r = 550$ nm. Note that all angles are interior angles.

are small. However, this is not always the case. In particular, as the angle of incidence (within the grating) becomes very high, the balance can shift to favouring Kogelnik's model. Indeed, the PSM model has been derived as an approximate differential representation of the rigorous chain matrix model. In the limit of zero incidence angle, this representation is exact, but as the incidence angle increases, the model should be expected to become somewhat less accurate as can be appreciated from the discussion in Section 12.4.3.

12.5 Slanted Reflection Gratings

We may use the PSM equations for the unslanted grating to derive corresponding equations for the general slanted grating. To do this, we define rotated Cartesian coordinates (x', y') that are related to the unprimed Cartesian system by

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (12.76)$$

This is illustrated in Figure 12.14. In the unprimed frame, we have

$$\mathbf{k}_c = \beta \begin{pmatrix} \sin \theta_c \\ \cos \theta_c \end{pmatrix}; \quad \mathbf{k}_i = \beta \begin{pmatrix} \sin \theta_c \\ -\cos \theta_c \end{pmatrix} \quad (12.77)$$

whereas in the primed frame, we have

$$\mathbf{k}'_c = \beta \begin{pmatrix} \sin(\theta_c - \psi) \\ \cos(\theta_c - \psi) \end{pmatrix}; \quad \mathbf{k}'_i = \beta \begin{pmatrix} \sin(\theta_c + \psi) \\ -\cos(\theta_c + \psi) \end{pmatrix} \quad (12.78)$$

Derivatives in the primed system are related to those in the unprimed system by Leibnitz's chain rule

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial x'}{\partial x} \frac{\partial}{\partial x'} + \frac{\partial y'}{\partial x} \frac{\partial}{\partial y'} = \cos \psi \frac{\partial}{\partial x'} + \sin \psi \frac{\partial}{\partial y'} \\ \frac{\partial}{\partial y} &= \frac{\partial y'}{\partial y} \frac{\partial}{\partial y'} + \frac{\partial x'}{\partial y} \frac{\partial}{\partial x'} = -\sin \psi \frac{\partial}{\partial x'} + \cos \psi \frac{\partial}{\partial y'} \end{aligned} \quad (12.79)$$

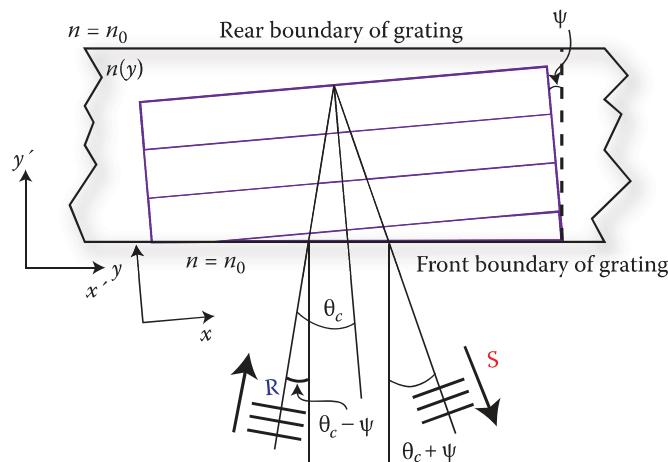


FIGURE 12.14 Geometry for rotating the unslanted reflection grating in the (x, y) system to a slanted grating in the primed system.

The PSM equations for the σ -polarisation may therefore be written as

$$\begin{aligned}\frac{\mathbf{k}'_c}{\beta} \cdot \nabla' R &= \frac{\partial R}{\partial x'} \sin(\theta_c - \psi) + \frac{\partial R}{\partial y'} \cos(\theta_c - \psi) \\ &= \frac{R}{2} \left\{ 2i\beta - \frac{1}{n \cos \theta_c} \frac{\partial n}{\partial y} \right\} - \frac{S}{2n \cos \theta_c} \frac{\partial n}{\partial y}\end{aligned}\quad (12.80)$$

and

$$\begin{aligned}\frac{\mathbf{k}'_i}{\beta} \cdot \nabla' S &= \frac{\partial S}{\partial x'} \sin(\theta_c + \psi) - \frac{\partial S}{\partial y'} \cos(\theta_c + \psi) \\ &= \frac{S}{2} \left\{ 2i\beta + \frac{1}{n \cos \theta_c} \frac{\partial n}{\partial y} \right\} + \frac{R}{2n \cos \theta_c} \frac{\partial n}{\partial y}\end{aligned}\quad (12.81)$$

Note that we have kept the unprimed frame on the right hand side (RHS) on purpose because, in this system, the index profile is one-dimensional and as such, much easier to evaluate.

12.5.1 Analytical Solutions for Single-Colour Gratings

To study the single-colour grating, we shall use the unslanted index profile (Equation 12.58) in the unprimed frame leading to the following profile in the primed frame

$$n = n_0 + n_1 \cos(2\alpha\beta \cos \theta_r \{ \sin \psi x' - \cos \psi y' \}) = n_0 + \frac{n_1}{2} \left\{ e^{2i\beta \alpha \cos \theta_r \hat{\mathbf{K}} \cdot \mathbf{r}'} + e^{-2i\beta \alpha \cos \theta_r \hat{\mathbf{K}} \cdot \mathbf{r}'} \right\} \quad (12.82)$$

where $\hat{\mathbf{K}}$ is the unit grating vector in the primed frame.

We now let

$$\begin{aligned}R &\rightarrow Re^{i\beta \{ \sin(\theta_c - \psi)x' + \cos(\theta_c - \psi)y' \}} \\ S &\rightarrow Se^{i\beta \{ \sin(\theta_c + \psi)x' - \cos(\theta_c + \psi)y' \}}\end{aligned}\quad (12.83)$$

Equations 12.80 and 12.81 then become

$$\begin{aligned}\sin(\theta_c - \psi) \frac{\partial R}{\partial x'} + \cos(\theta_c - \psi) \frac{\partial R}{\partial y'} \\ &= -\frac{S}{2} \left\{ \frac{1}{n \cos \theta_c} \frac{\partial n}{\partial y} \right\} e^{i\beta \{ [\sin(\theta_c + \psi) - \sin(\theta_c - \psi)]x' - [\cos(\theta_c - \psi) + \cos(\theta_c + \psi)]y' \}} \\ &= -\frac{S}{2} \left\langle \left\{ \frac{1}{n_0 \cos \theta_c} \frac{\partial}{\partial y} \frac{n_1}{2} \left\{ e^{2i\beta \alpha (\cos \theta_r)y} + e^{-2i\beta \alpha (\cos \theta_r)y} \right\} \right\} e^{2i\beta \cos \theta_c \{ [\sin \psi]x' - [\cos \psi]y' \}} \right\rangle \\ &= -\frac{S}{2} \left\langle \left\{ \frac{n_1 i \beta \alpha (\cos \theta_r)}{n_0 \cos \theta_c} \left\{ e^{2i\beta \alpha (\cos \theta_r)y} + ... \right\} \right\} e^{-2i\beta \cos \theta_c y} \right\rangle \\ &= -\frac{i \beta n_1}{2n_0} \alpha \frac{\cos \theta_r}{\cos \theta_c} S e^{2i\beta (\alpha \cos \theta_r - \cos \theta_c)(y' \cos \psi - x' \sin \psi)}\end{aligned}\quad (12.84)$$

and

$$\sin(\theta_c + \psi) \frac{\partial S}{\partial x'} - \cos(\theta_c + \psi) \frac{\partial S}{\partial y'} = -\frac{i \beta n_1}{2n_0} \alpha \frac{\cos \theta_r}{\cos \theta_c} R e^{-2i\beta (\alpha \cos \theta_r - \cos \theta_c)(y' \cos \psi - x' \sin \psi)} \quad (12.85)$$

Next, we make the transformation

$$\begin{aligned}\hat{S} &= S(y') e^{2i\beta(\alpha \cos \theta_r - \cos \theta_c)(y' \cos \psi - x' \sin \psi)} \\ \hat{R} &= R(y')\end{aligned}\quad (12.86)$$

at which point the PSM equations once again reduce to a simple pair of ordinary differential equations in the form of Kogelnik's equations:

$$\begin{aligned}c_R \frac{dR}{dy'} &= -i\kappa \hat{S} \\ c_S \frac{d\hat{S}}{dy'} &= -i\vartheta \hat{S} - i\kappa R\end{aligned}\quad (12.87)$$

with coefficients

$$\begin{aligned}c_{R(\text{PSM})} &= \frac{\cos \theta_c \cos(\theta_c - \psi)}{\alpha \cos \theta_r} \\ c_{S(\text{PSM})} &= -\frac{\cos \theta_c \cos(\theta_c + \psi)}{\alpha \cos \theta_r} \\ \vartheta_{\text{PSM}} &= 2\beta \left(1 - \frac{\cos \theta_c}{\alpha \cos \theta_r} \right) \cos^2 \theta_c\end{aligned}\quad (12.88)$$

For comparison, Kogelnik's coefficients are

$$\begin{aligned}c_{R(\text{KOG})} &= \cos(\theta_c - \psi) \\ c_{S(\text{KOG})} &= \cos(\theta_c - \psi) - 2\alpha \cos \theta_r \cos \psi \\ \vartheta_{\text{KOG}} &= 2\alpha \beta \cos \theta_r (\cos \theta_c - \alpha \cos \theta_r)\end{aligned}\quad (12.89)$$

With the usual reflective boundary conditions, $\hat{R}(0) = 1$ and $\hat{S}(d) = 0$, we can then use the standard formula to describe the diffraction efficiency of the slanted reflection grating:

$$\eta_\sigma = \frac{|c_S|}{c_R} \hat{S}(0) \hat{S}^*(0) = \frac{\kappa^2 \sinh^2(d\Upsilon)}{\kappa^2 \sinh^2(d\Upsilon) - c_R c_S \Upsilon^2} \quad (12.90)$$

where

$$\Upsilon^2 = -\frac{\vartheta^2}{4c_S^2} - \frac{\kappa^2}{c_R c_S} \quad (12.91)$$

Substitution of either Equation 12.88 or Equation 12.89 into Equation 12.90 gives the required expression for the diffractive efficiency in either the Kogelnik or PSM model.

When $\psi = 0$, $\eta_{\text{PSM}/\sigma}$ reduces to the unslanted formula that was derived in Section 12.4.4. In the case of finite slant and Bragg resonance (where $\cos \theta_c = \alpha \cos \theta_r$), we have

$$\eta_{PSM/\sigma} = \tanh^2(d\kappa \sqrt{\sec(\theta_c - \psi) \sec(\theta_c + \psi)}) \quad (12.92)$$

which is identical to Kogelnik's solution. Note that the behaviour of the π -polarisation is simply described by making the transformation (Equation 12.68) in all formulae of interest.

The PSM model for the slanted grating under either the σ - or π -polarisations gives expressions very similar to Kogelnik's theory. We shall see shortly that for most gratings of practical interest to display and optical element holography, the two theories produce predictions that are extremely close.

12.6 Slanted Reflection Gratings in Three Dimensions

The PSM model for the slanted grating at oblique incidence can easily be extended to the case of oblique incidence in three dimensional space.

We start by recognising that the PSM equations for an unslanted grating in three dimensions are no different from their two-dimensional counterparts. We use the same index profile as before (Equation 12.58) and simply generalise the coordinates from the (x,y) Cartesian system to an (x,y,z) Cartesian system. The reference and image rays are therefore, as before, in the (x,y) plane. We now introduce a primed system of Cartesian coordinates which has been rotated first about the z -direction (as before) and then about the x -direction. Accordingly, we have

$$\begin{aligned} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} &= R_x R_z \mathbf{r} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \cos \varphi \sin \psi & \cos \varphi \cos \psi & -\sin \varphi \\ \sin \varphi \sin \psi & \sin \varphi \cos \psi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned} \quad (12.93)$$

The primed system of coordinates is then, as before, the real-world coordinates of the grating. In these coordinates, the grating has a slant in two directions defined by the angles ψ and φ .

The PSM equations (Equations 12.43 and 12.44) can then be written in vector notation as

$$\frac{\mathbf{k}'_c}{\beta} \cdot \nabla' R = \frac{R}{2} \left\{ 2i\beta - \frac{1}{2n \cos \theta_c} \frac{\partial n}{\partial y} \right\} - \frac{S}{2n \cos \theta_c} \frac{\partial n}{\partial y} \quad (12.94)$$

and

$$\frac{\mathbf{k}'_i}{\beta} \cdot \nabla' S = \frac{S}{2} \left\{ 2i\beta + \frac{1}{n \cos \theta_c} \frac{\partial n}{\partial y} \right\} + \frac{R}{2n \cos \theta_c} \frac{\partial n}{\partial y} \quad (12.95)$$

where the reference and image k -vectors in the primed frame are defined as

$$\mathbf{k}'_c = R_x R_z \mathbf{k}_c = \beta \begin{pmatrix} \sin(\theta_c - \psi) \\ \cos \varphi \cos(\theta_c - \psi) \\ \sin \varphi \cos(\theta_c - \psi) \end{pmatrix}; \quad \mathbf{k}'_i = R_x R_z \mathbf{k}_i = \beta \begin{pmatrix} \sin(\theta_c + \psi) \\ -\cos \varphi \cos(\theta_c + \psi) \\ -\sin \varphi \cos(\theta_c + \psi) \end{pmatrix} \quad (12.96)$$

These equations may now be solved in exactly the same way as the two-dimensional equations using

$$\begin{aligned} R &\rightarrow Re^{i\mathbf{k}_c' \cdot \mathbf{r}'} \\ S &\rightarrow Se^{i\mathbf{k}_c' \cdot \mathbf{r}'} \end{aligned} \quad (12.97)$$

giving

$$\begin{aligned} \frac{\mathbf{k}_c' \cdot \nabla' R}{\beta} &= -\frac{S}{2n \cos \theta_c} \frac{\partial n}{\partial y} e^{i(\mathbf{k}_c' - \mathbf{k}_c) \cdot \mathbf{r}'} \\ &= -\frac{S}{2} \left\langle \left\{ \frac{n_1 i \beta \alpha (\cos \theta_r)}{n_0 \cos \theta_c} \left\{ e^{2i\beta\alpha(\cos \theta_r)y} + \dots \right\} \right\} e^{-2i\beta \cos \theta_c y} \right\rangle \\ &= -i\kappa\alpha \frac{\cos \theta_r}{\cos \theta_c} Se^{2i\beta(\alpha \cos \theta_r - \cos \theta_c) \left(\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \right)^T \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}} \\ &= -i\kappa\alpha \frac{\cos \theta_r}{\cos \theta_c} Se^{2i\beta(\alpha \cos \theta_r - \cos \theta_c)(-\sin \psi x' + \cos \psi \cos \varphi y' + \cos \psi \sin \varphi z')} \end{aligned} \quad (12.98)$$

and

$$\frac{\mathbf{k}_c' \cdot \nabla' S}{\beta} = -\frac{i\beta n_1}{2n_0} \alpha \frac{\cos \theta_r}{\cos \theta_c} Re^{-2i\beta(\alpha \cos \theta_r - \cos \theta_c)(-\sin \psi x' + \cos \psi \cos \varphi y' + \cos \psi \sin \varphi z')} \quad (12.99)$$

We then choose

$$\begin{aligned} \hat{S} &= S(y') e^{2i\beta(\alpha \cos \theta_r - \cos \theta_c)(-\sin \psi x' + \cos \psi \cos \varphi y' + \cos \psi \sin \varphi z')} \\ \hat{R} &= R(y') \end{aligned} \quad (12.100)$$

at which point the three-dimensional PSM equations reduce to the standard form of Kogelnik's equations (e.g., Equation 12.87) with coefficients

$$\begin{aligned} c_{R(\text{PSM})} &= \frac{\cos \varphi \cos \theta_c \cos(\theta_c - \psi)}{\alpha \cos \theta_r} \\ c_{S(\text{PSM})} &= -\frac{\cos \varphi \cos \theta_c \cos(\theta_c + \psi)}{\alpha \cos \theta_r} \\ \vartheta_{\text{PSM}} &= 2\beta \left(1 - \frac{\cos \theta_c}{\alpha \cos \theta_r} \right) \cos^2 \theta_c \end{aligned} \quad (12.101)$$

We therefore see that the effect of tilting the grating in the second dimension is simply to multiply the obliquity constants, c_R and c_S , by a factor of $\cos \varphi$. Note that we have treated the σ -polarisation above with respect to the grating planes. The polarisation with respect to the grating surface is, in general, mixed for finite φ .

At Bragg resonance, the diffraction efficiency for the three-dimensional case reduces to

$$\eta_{\text{PSM}/\sigma} = \tanh^2 \left(d\kappa \sec \varphi \sqrt{\sec(\theta_c - \psi) \sec(\theta_c + \psi)} \right) \quad (12.102)$$

Likewise,

$$\eta_{\text{PSM}/\pi} = \tanh^2 \left(d\kappa \cos 2\theta_c \sec \varphi \sqrt{\sec(\theta_c - \psi) \sec(\theta_c + \psi)} \right) \quad (12.103)$$

12.7 Transmission Gratings with Slanted Fringes

In this chapter we have concentrated on the reflection hologram. However, the PSM model can be applied to the transmission hologram by simply using the appropriate boundary conditions to solve the PSM equations in a rotated frame. Alternatively, the analysis of Section 12.4 can be repeated for Figure 12.10 rotated by 90°. The first strategy is very simple and works in two or three dimensions. For example, in two dimensions and with a single slant angle, ψ , we use the transmission boundary conditions

$$\begin{aligned} R(0) &= 1 \\ S(0) &= 0 \end{aligned} \quad (12.104)$$

to solve Equation 12.87, which, at Bragg resonance, results in the standard formula given by Kogelnik's theory.

$$\eta_{\sigma T/PSM} = \sin^2\left(\kappa d / \sqrt{c_R c_S}\right) = \sin^2\left(\kappa d / \sqrt{-\cos(\theta_c - \psi)\cos(\theta_c + \psi)}\right) \quad (12.105)$$

12.8 Comparison of the PSM Theory with Kogelnik's Theory for Slanted Gratings

Figure 12.15 shows typical plots comparing diffraction efficiency versus replay angle according to the PSM model and Kogelnik's model for the two-dimensional slanted reflection grating. The plots refer

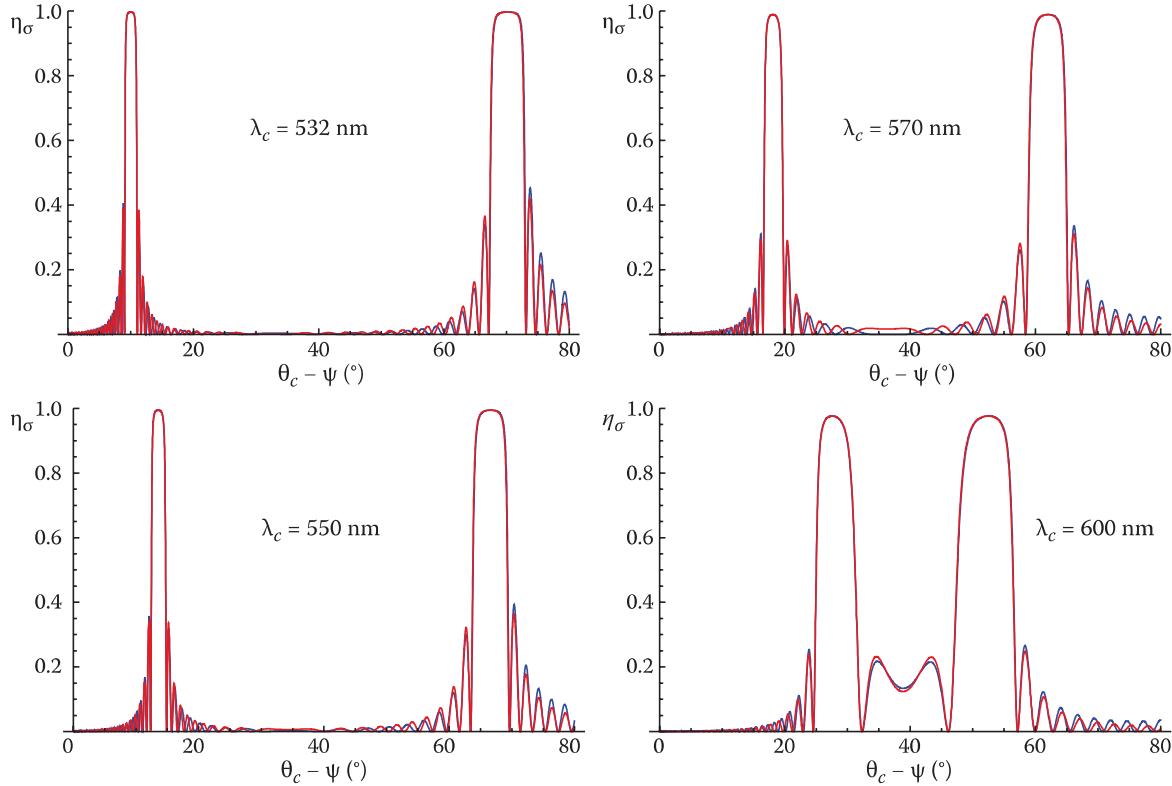


FIGURE 12.15 Diffraction efficiency (σ -polarisation) against replay angle in degrees for a 12 μm grating recorded at 532 nm with an index modulation of $n_1 = 0.03$ ($n_0 = 1.5$) at a recording reference angle of $\theta_r = 30^\circ$ and with a fringe slant of $\psi = -40^\circ$. The grating is replayed at four different wavelengths. Note that, as expected, when the replay and recording wavelengths are the same, reconstruction is observed at 70° and 10° (PSM, blue lines; Kogelnik, red lines). Note that all angles are interior angles.

to a 12 μm grating recorded at 532 nm, with an index modulation of $n_1 = 0.03$ ($n_0 = 1.5$) at a recording reference angle of $\theta_r = 30^\circ$ and with a fringe slant of $\psi = 40^\circ$. Clearly, the two models give *extremely* similar predictions. This includes an almost exact replication of the asymmetry between the two reconstruction maxima. A survey of a wide range of typical scenarios reveals that there are some differences between the two models, but these are at larger interior grating angles than permitted by Snell refraction when gratings are used in air or at very high index modulation ($n_1 > 0.25$)—the overall behaviour of the two models is thus very similar for practical purposes. In Appendix 8 rigorous coupled wave theory is reviewed and its predictions compared to those of both Kogelink's theory and the PSM theory.

12.9 Polychromatic Slanted Reflection Gratings

The discussion in Section 12.4.5 applies identically to the case of the polychromatic slanted reflection grating. The PSM model yields simple ordinary differential equations for both the two- and three-dimensional cases and these can be solved very easily on a PC. Alternatively, if the primary diffractive peaks in a polychromatic grating do not overlap, approximate formulae such as Equations 12.74 and 12.75 may be used very effectively with the coefficients in Equations 12.88 or 12.101.

12.10 Extending PSM to Describe Spatially Multiplexed Monochromatic Gratings

The fundamental idea of the PSM model is to describe the permittivity distribution of a grating as an infinite sum of discontinuous step functions. The Fresnel reflections of an incident plane wave can then be summed to calculate the properties of the grating. We started our discussion of the PSM model at finite incidence angle by considering the following simple index distribution:

$$n = n_0 \left(1 + \frac{\kappa}{\beta} \left\{ e^{2i\alpha\beta \cos \theta_{r,y}} + e^{-2i\alpha\beta \cos \theta_{r,y}} \right\} \right) \quad (12.106)$$

We shall now extend our discussion to consider the case of a spatially multiplexed monochromatic grating of the form

$$n = n_0 \left(1 + \sum_{j=1}^N \frac{\kappa_j}{\beta} \left\{ e^{2i\alpha_j \beta \cos \theta_{rj} (\sin \psi_j x' - \cos \psi_j y')} + e^{-2i\alpha_j \beta \cos \theta_{rj} (\sin \psi_j x' - \cos \psi_j y')} \right\} \right) \quad (12.107)$$

where we have used the same notation as in Sections 12.5 and 12.6.

We can imagine that this grating has been created by the holographic interference of a single plane reference wave with many plane object waves, each incident at a different angle $\Phi_{oj} \equiv -(\theta_{rj} + \psi_j)$ to the physical substrate normal. The reference wave is incident at an angle $\Phi_r \equiv \theta_{rj} - \psi_j$ for all j . In this way, the grating can be regarded as being composed of N unslanted gratings each tilted by a different angle ψ_j and each having a common angle of incidence with respect to the substrate normal. We assume that interference only occurs between the reference and each of the object waves and not between the object waves themselves. This is equivalent to the assumption of sequential recording or to the assumption that the reference wave is of a much larger amplitude than each of the object waves.

The PSM model can now be used to treat Equation 12.107 by summing the Fresnel reflections of a single incident plane wave from N infinite series of stacked mirrors. Each of these infinite mirror sets is characterised by a different tilt. To see how this works, we need to imagine a single plane reference wave, R , which is incident onto the multiplexed grating. This wave interacts with each grating, generating a separate signal wave, which we can label S_j .

The analysis [3] follows Section 12.5 except that now Equations 12.84 and 12.85 become

$$\begin{aligned}
 \cos\Phi_c \frac{\partial R}{\partial y'} &= -i\alpha \sum_{j=1}^N \kappa_j \frac{\cos(\Phi_r + \psi_j)}{\cos(\Phi_c + \psi_j)} S_j \left\langle e^{2i\beta\alpha\zeta_j(\Phi_r, \psi_j)} e^{i(\mathbf{k}'_j - \mathbf{k}'_c) \cdot \mathbf{r}'} \right\rangle + \langle \dots \rangle \\
 &= -i\alpha \sum_{j=1}^N \kappa_j \frac{\cos(\Phi_r + \psi_j)}{\cos(\Phi_c + \psi_\mu)} S_j e^{2i\beta\alpha\zeta_j(\Phi_r, \psi_j)} e^{i(\mathbf{k}'_j - \mathbf{k}'_c) \cdot \mathbf{r}'} \\
 -\sin\Phi_{ij} \frac{\partial S_j}{\partial x'} - \cos\Phi_{ij} \frac{\partial S_j}{\partial y'} &= -i\alpha\kappa_j \frac{\cos(\Phi_r + \psi_j)}{\cos(\Phi_c + \psi_j)} R \left\langle e^{-2i\beta\alpha\zeta_j(\Phi_r, \psi_j)} e^{-i(\mathbf{k}'_j - \mathbf{k}'_c) \cdot \mathbf{r}'} \right\rangle + \langle \dots \rangle \\
 &= -i\alpha\kappa_j \frac{\cos(\Phi_r + \psi_j)}{\cos(\Phi_c + \psi_j)} R e^{-2i\beta\alpha\zeta_j(\Phi_r, \psi_j)} e^{-i(\mathbf{k}'_j - \mathbf{k}'_c) \cdot \mathbf{r}'}
 \end{aligned} \tag{12.108}$$

where

$$\left. \begin{array}{l} \theta_{ej} - \psi_j = \Phi_c \\ \theta_{rj} - \psi_j = \Phi_r \\ \theta_{ej} + \psi_j = -\Phi_{ij} \\ \theta_{rj} + \psi_j = -\Phi_{oj} \end{array} \right\} \quad \forall j \leq N \tag{12.109}$$

and where

$$\zeta_j(\theta_{rj}, \psi_j) \equiv \cos\theta_{rj}(y' \cos\psi_j - x' \sin\psi_j) \tag{12.110}$$

We then define different pseudo-fields for each signal wave according to the following rule:

$$S_j = \hat{S}_j(y') e^{-2i\beta\alpha\zeta_j(\theta_{rj}, \psi_j) - i(\mathbf{k}'_j - \mathbf{k}'_c) \cdot \mathbf{r}'} \tag{12.111}$$

This leads directly to the N-PSM equations for the spatially multiplexed monochromatic grating

$$\begin{aligned}
 \frac{\partial R}{\partial y'} &= -i \sum_{j=1}^N \frac{\kappa_j}{c_{Rj}} \hat{S}_j \\
 c_{Sj} \frac{\partial \hat{S}_j}{\partial y'} &= -i\vartheta_j \hat{S}_j - i\kappa_j R
 \end{aligned} \tag{12.112}$$

where for the σ -polarisation

$$\begin{aligned}
 c_{Rj} &= \frac{\cos\theta_{ej} \cos(\theta_{ej} - \psi_j)}{\alpha_j \cos\theta_{rj}} \\
 c_{Sj} &= -\frac{\cos\theta_{ej} \cos(\theta_{ej} + \psi_j)}{\alpha_j \cos\theta_{rj}}
 \end{aligned} \tag{12.113}$$

$$\vartheta_j = 2\beta \left(1 - \frac{\cos\theta_{ej}}{\alpha_j \cos\theta_{rj}} \right) \cos^2\theta_{ej}$$

Equations 12.112 may now be solved using the boundary conditions appropriate for a reflection multiplexed grating, that is,

$$\begin{aligned} R(0) &= 1 \\ \hat{S}_j(d) &= 0 \quad \forall j \leq N \end{aligned} \tag{12.114}$$

At Bragg resonance, c_R becomes a constant

$$c_{Rj} = \frac{\cos \theta_{cj} \cos(\theta_{cj} - \psi_j)}{\alpha_j \cos \theta_{rj}} = \cos \Phi_c \tag{12.115}$$

and Equation 12.112 then gives the following expression for the diffractive efficiency of the j th grating:

$$\eta_j \equiv \frac{1}{c_R} |c_{Sj}| \hat{S}_j(0) \hat{S}_j^*(0) = \frac{1}{c_{Sj}} \frac{\kappa_j^2}{\sum_{k=1}^N \frac{\kappa_k^2}{c_{Sk}}} \tanh^2 \left\{ d \sqrt{-\frac{1}{c_R} \sum_{k=1}^N \frac{\kappa_k^2}{c_{Sk}}} \right\} \tag{12.116}$$

The total diffraction efficiency of the entire multiplexed grating is likewise found by summing the diffractive response from each grating:

$$\eta \equiv \sum_{j=1}^N \eta_j = \tanh^2 \left\{ d \sqrt{\frac{1}{\cos \Phi_c} \sum_{k=1}^N \frac{\kappa_k^2}{\cos \Phi_{ik}}} \right\} \tag{12.117}$$

Here, Φ_c is the incidence angle of the replay reference wave and Φ_{ik} is the incidence angle of the k th signal wave. These results are of course none other than the expressions that we obtained in Chapter 11 using the N -coupled wave theory (see Equation 11.165)! At Bragg resonance, the PSM model of the multiplexed grating gives an identical description to the corresponding N -coupled wave theory just as the simple PSM theory gave an identical description at Bragg resonance to Kogelnik's theory. Here again, however, the advantage of the PSM model over the N -coupled wave theory is that we can immediately extend the result to the polychromatic spatially multiplexed grating in the limit that the recording wavelengths are sufficiently separated. Here, the diffraction efficiency at the Bragg resonance associated with the j th replay angle at the m th recording wavelength is given by

$$\eta_{mj} \equiv \frac{1}{c_{Sj}} \frac{\kappa_{mj}^2}{\sum_{k=1}^N \frac{\kappa_{mk}^2}{c_{Sk}}} \tanh^2 \left\{ d \sqrt{-\frac{1}{c_R} \sum_{k=1}^N \frac{\kappa_{mk}^2}{c_{Sk}}} \right\} \tag{12.118}$$

Again, the total diffractive response at the m th wavelength is then simply

$$\eta_m \equiv \sum_{j=1}^N \eta_{mj} = \tanh^2 \left\{ d \sqrt{\frac{1}{\cos \Phi_c} \sum_{k=1}^N \frac{\kappa_{mk}^2}{\cos \Phi_{ik}}} \right\} \tag{12.119}$$

In the limit that $N \rightarrow \infty$ this then leads to formulae for the diffractive efficiency of the polychromatic hologram

$$\begin{aligned} \eta_m(\Phi_c, \Phi_i) &= \frac{\kappa_m^2(\Phi_i)}{L_m \cos \Phi_i} \tanh^2 \left\{ d \sqrt{\frac{L_m}{\cos \Phi_c}} \right\} \\ \eta_m &= \frac{1}{\Delta \Phi} \int \frac{\kappa_m^2(\Phi')}{L_m \cos \Phi'} \tanh^2 \left\{ d \sqrt{\frac{L_m}{\cos \Phi_c}} \right\} d\Phi' = \tanh^2 \left\{ d \sqrt{\frac{L_m}{\cos \Phi_c}} \right\} \end{aligned} \tag{12.120}$$

where

$$L_m = \frac{1}{\Delta\Phi} \int \frac{\kappa_m^2(\Phi)}{\cos\Phi} d\Phi \quad (12.121)$$

and where Φ is the replay image angle and $\Delta\Phi$ is the total reconstructed image angle range. If we assume that we record the hologram such that

$$\kappa_m^2(\Phi) \rightarrow \kappa_m^2 \cos\Phi \quad (12.122)$$

then this reduces to the simpler form

$$\eta_m = \tanh^2 \left\{ \frac{d\kappa_m}{\sqrt{\cos\Phi_c}} \right\} \quad (12.123)$$

This tells us that a lossless polychromatic hologram is therefore theoretically capable of perfect diffractive replay at each and every wavelength. In other words, if it is co-illuminated by P co-propagating reference plane waves, each of a different wavelength, then a hologram is capable of producing a perfect diffractive response to each of these P waves simultaneously. However, the all-important parameter here is κ_m —the effective index modulation achievable with a given material at the m th wavelength. If this is low, then a larger grating thickness will be needed to achieve a bright hologram. The PSM model can be extended to model lossy holograms by consideration of a complex index. In this case, the strategy of simply making a hologram thicker to produce a brighter diffractive response does not work because the further the reference wave penetrates into the hologram, the more it is absorbed. This leads to the lossy polychromatic hologram being characterised by a maximum attainable diffractive response, which is innately dependent on the achievable modulation in a given material.

Another insight to take away from Equation 12.120 is that if we record a spatially multiplexed grating that has a diffuse object beam spanning a finite but small angle range, the usual formula for a single polychromatic grating applies but now with an effective modulation and an average replay angle. To see this, we make the approximation

$$L_m = \frac{1}{\Delta\Phi} \int \frac{\kappa_m^2(\Phi_i)}{\cos\Phi_i} d\Phi_i \sim \frac{1}{\langle \cos\Phi_i \rangle \Delta\Phi} \int \kappa_m^2(\Phi) d\Phi \sim \frac{1}{\cos\langle\Phi_i\rangle} \langle \kappa_m^2(\Phi) \rangle \quad (12.124)$$

which then leads to a total diffraction efficiency at the m th wavelength of

$$\eta_m = \tanh^2 \left\{ d \langle \kappa_m^2(\Phi) \rangle^{1/2} \sqrt{\sec\Phi_c \sec\langle\Phi_i\rangle} \right\} \quad (12.125)$$

Figure 12.16 shows graphically the results of Equation 12.123 for several typical three-colour recording parameter sets. In Figure 12.16a we assume an index modulation of $n_1 = 0.01$ for each wavelength. This is typical of a Silver Halide emulsion such as PFG-03CN which has a thickness of between 9 and 10 μm . We can therefore expect a diffraction efficiency for each colour in the region of 25% for reflection holograms made in this type of material. In Figure 12.16b we use an index modulation of 0.02 for each chromatic component. This is more characteristic of a modern photopolymer. For a 12 μm emulsion we would therefore expect a diffraction efficiency for each colour component of around 70%. Clearly these numbers are for the lossless case and loss can be expected to diminish the actual efficiencies somewhat.

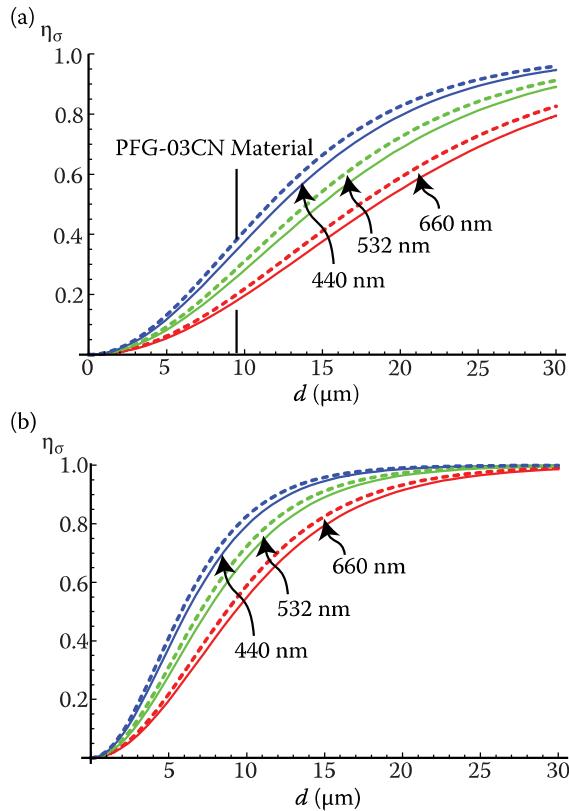


FIGURE 12.16 Diffractive efficiency versus grating thickness for typical lossless three-colour reflection volume phase holograms according to the N-PSM theory (σ -polarisation). All holograms are recorded at 660 nm, 532 nm and 440 nm. Dashed lines indicate a 45 degree reference illumination beam (in air). Solid lines indicate the case of normal incidence illumination. (a) $n_l(660) = n_l(532) = n_l(440) = 0.01$ and (b) $n_l(660) = n_l(532) = n_l(440) = 0.02$. Average index of grating assumed to be $n_0 = 1.5$.

12.11 Coupled Wave Theory, PSM and the Rigorous Coupled Wave Theory

In Section 12.10 we have generalized the PSM model to describe the spatially multiplexed volume phase grating. Just as the simple PSM model provides a simple and intuitive method of understanding and calculating diffraction in simple volume phase gratings, so the N-PSM model builds on this simplicity to construct an intuitive model of diffraction capable of accurately describing spatially multiplexed volume phase gratings and colour reflection holograms.

The basic idea behind PSM is that all volume phase gratings, and *in particular* reflection phase gratings, can be described very well by the process of Fresnel reflection of the illuminating wave from the fringe planes. To understand why this should be so is simple. The rigorous coupled wave theory of Moharam and Gaylord [11], which is reviewed in Appendix 8, demonstrates that almost always reflection gratings are dominated by the first order “+1” diffractive mode. But this is precisely the mode which the process of Fresnel reflection describes. The basic PSM equations are essentially rigorous for the un-slanted reflection grating and here Fresnel reflection describes the grating perfectly. PSM simply assumes that Fresnel reflection continues to operate as the sole diffractive process even when the fringe planes are tilted. Since only the “+1” mode actually is of any importance this is a rather good approximation. Snell’s law also helps in practice as this acts to steepen all incidence angles and lessen grating slants. In a tilted grating at Bragg resonance the Fresnel description is most accurate at Bragg resonance. Here it is hardly surprising that Kogelnik’s theory [1] agrees identically to PSM. In fact Kogelnik assumes that the signal ray propagates according to the formula

$$\mathbf{k}_i = \mathbf{k}_c + \mathbf{K} \quad (12.126)$$

where \mathbf{k}_i represents the wavevector of the diffracted signal ray, \mathbf{k}_c is the illuminating wavevector and \mathbf{K} is the grating vector.* PSM, on the other hand assumes that

$$\mathbf{k}_i = \mathbf{k}_c - \frac{2\mathbf{K} \cdot \mathbf{k}_c}{|\mathbf{K}|^2} \mathbf{K} \quad (12.127)$$

Finally we show in Appendix 8 that rigorous coupled wave theory implies that actually the “+1” mode must propagate according to the law

$$\mathbf{k}_i = (k_x + K_x)\hat{\mathbf{x}} - \sqrt{\beta^2 - (k_x + K_x)^2}\hat{\mathbf{y}} \quad (12.128)$$

All these expressions are the same at Bragg resonance but all three differ away from resonance.

In Appendix 8 we compare the N-PSM theory with rigorous coupled wave theory for a variety of simple monochromatic and polychromatic spatially multiplexed gratings. Here we see that N-PSM provides a good description of the spatially multiplexed grating—and for precisely the same reasons we have discussed above. When rigorous coupled wave theory is applied to the spatially multiplexed grating once again only first order “+1” type modes turn out to be of importance. Cross-modes which are driven by several gratings at once are nearly always extremely small unless index modulations are extremely high. This then means that grating cross-coupling can be effectively ignored and a single signal wave can be attributed to a single grating in the multiplex.

12.12 Lippmann Photography

The PSM theory can be used to analyse Lippmann photography in the limit of a large focal length Lippmann camera. Here, we can approximate the Lippmann photograph as a normal-incidence broadband polychromatic reflection grating. As such, we can use Equation 12.27 to model playback. Of course, Lippmann photographs in reality are rather different from holograms and gratings because each of the many object rays in a finite focal length Lippmann camera act as reference beams to all other object beams. As such, the grating structure is rather more complex than the standard hologram, where it is usual to make the approximation that the individual object rays only interfere with the reference wave. This evident complexity is neglected in the present analysis. Nevertheless, the analysis is still instructive as it serves to clearly demonstrate the concept of interferometric spectral recording.

We start by assuming that the Lippmann photograph in the limit of infinite focal length recording can be described by the following index distribution:

$$n = n_0 + \frac{1}{2(\hat{\beta}_1 - \hat{\beta}_0)} \int_{\hat{\beta}_0}^{\hat{\beta}_1} \tilde{n}(\hat{\beta}) \left\{ e^{2i\hat{\beta}y} + e^{-2i\hat{\beta}y} \right\} d\hat{\beta} \quad (12.129)$$

where

$$\hat{\beta} \equiv \frac{2\pi n_0}{\lambda_r} = \alpha\beta \quad (12.130)$$

We imagine that this distribution has been created by a polychromatic wave of the form

$$E(y) = \int_{\hat{\beta}_0}^{\hat{\beta}_1} E_o(\hat{\beta}) e^{i\hat{\beta}y} d\hat{\beta} \quad (12.131)$$

* Note that we have used a different sign convention for the grating vector, \mathbf{K} and the grating vector, \mathbf{G} used elsewhere.

The spectral information in the wave $E_o(\hat{\beta})$ has therefore been transferred to the polychromatic grating in the form of the index spectrum $\tilde{n}(\hat{\beta})$. We can now generalise Equation 12.27 to write

$$\begin{aligned}\frac{dR}{dy} &= -\frac{iS}{\beta(\hat{\beta}_1 - \hat{\beta}_0)} \int_{\hat{\beta}_0}^{\hat{\beta}_1} \hat{\beta} \kappa(\hat{\beta}) e^{2i(\hat{\beta}-\beta)y} d\hat{\beta} \\ \frac{dS}{dy} &= \frac{iR}{\beta(\hat{\beta}_1 - \hat{\beta}_0)} \int_{\hat{\beta}_0}^{\hat{\beta}_1} \hat{\beta} \kappa(\hat{\beta}) e^{-2i(\hat{\beta}-\beta)y} d\hat{\beta}\end{aligned}\quad (12.132)$$

where

$$\kappa(\hat{\beta}) = \frac{\beta \tilde{n}(\hat{\beta})}{2n_0} \quad (12.133)$$

If the index modulation spectrum is flat between the start and end wavelength, that is, $\kappa(\alpha) = \kappa_1$, then these equations can be solved numerically using a Runge–Kutta integration subject to the boundary conditions

$$\begin{aligned}R(0) &= 1 \\ S(d) &= 0\end{aligned}\quad (12.134)$$

We shall also consider an index profile which possesses m peaks between the two limits, $\hat{\beta}_0$ and $\hat{\beta}_1$, where

$$\kappa(\hat{\beta}) = \kappa_1 \sin^2 \left\{ m \pi \frac{(\hat{\beta} - \hat{\beta}_0)}{\hat{\beta}_1 - \hat{\beta}_0} \right\} \quad (12.135)$$

The integrals in Equation 12.129 can then be calculated analytically and the resulting equations solved using a numerical Runge–Kutta method. Figure 12.17 shows four plots of the index distribution, $\kappa(\lambda)$, and the associated spectrum on replay, $\eta(\lambda) \equiv S(y = 0, \lambda) S^*(y = 0, \lambda)$. If we assume that the index modulation on recording is proportional to the square of the electric field, then $\kappa(\lambda)$ represents the recording spectrum and $\eta(\lambda)$, the reconstruction spectrum. The first diagram (a) illustrates the case of a boxcar spectrum. Here, the response is roughly flat in the middle of the spectrum, but there are two peaks at each extremity. The blue peak is rather bigger as might be anticipated by realising that the blue radiation will “feel” a deeper grating due to its smaller wavelength. The remaining three diagrams (b–d) show the cases of $m = 5, 6$ and 7 in Equation 12.135. These cases pertain to a $2 \mu\text{m}$ -thick grating with a value of $\kappa_1 = n_1 \pi / \lambda = 0.03 \pi / \lambda$. As can be seen, the spectrum is reproduced well until the case of $m = 7$. Here, the grating becomes too thin to reconstruct properly and the spectral peaks are averaged. If one increases the thickness to $5 \mu\text{m}$, then the case of Figure 12.17d becomes almost identical to that of Figure 12.17c with the spectral peaks being well reproduced. If even thicker emulsions are used, cases of $m = 30$ and beyond lead to good reconstruction spectra.

There are two other points to note from Figure 12.17. The first is that all the graphs show a larger diffractive response in the blue spectrum than in the red spectrum. Again, this can be understood by the ratio of the wavelength to grating thickness, which favours diffraction at shorter wavelengths. The second point is that some peaks are shifted slightly on reconstruction with respect to their maximum on recording. This phenomenon is due to interactions between the different fundamental gratings.

The PSM equations (Equation 12.132) are an approximation to the exact equations (Equation 12.6). However, numerical integration of Equation 12.6 using the index profile in Equation 12.129 produces substantially identical results to the above analysis. This is because, even in the fairly extreme case of a Lippmann photograph, the approximation (Equation 12.10) is still valid. As we have seen, Equation 12.6

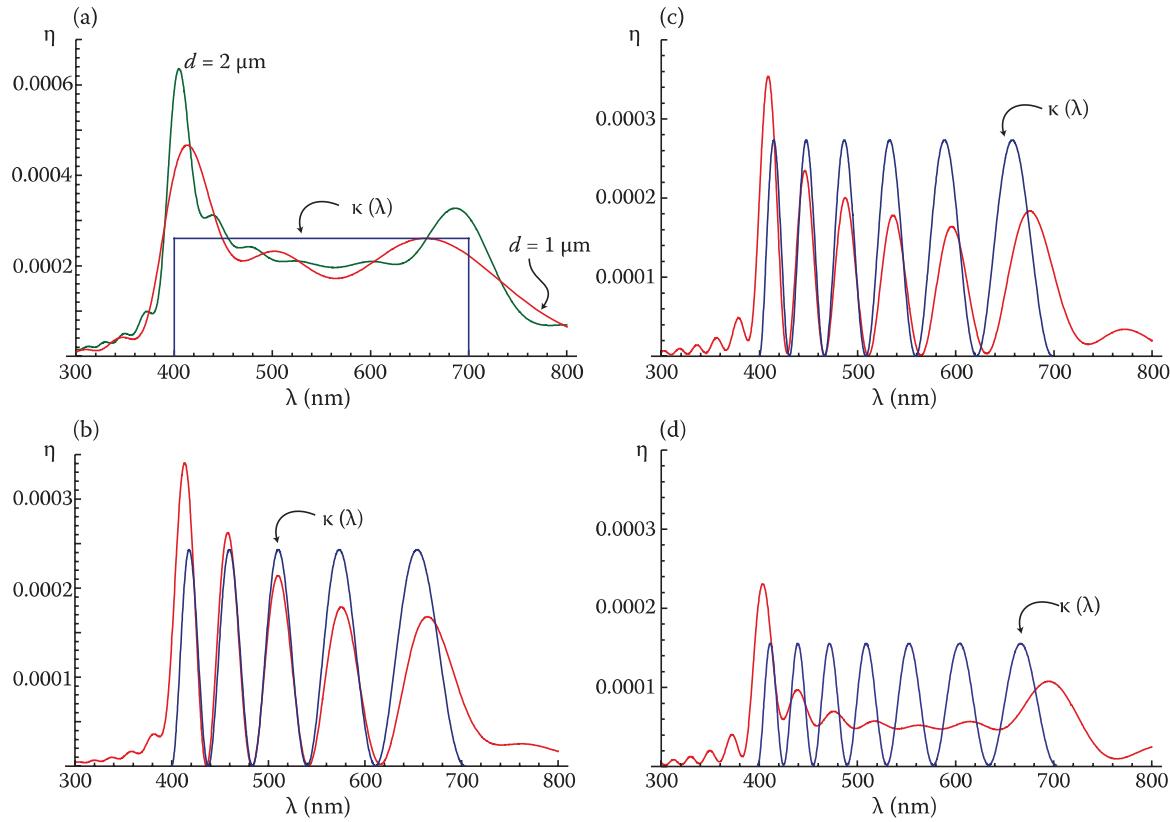


FIGURE 12.17 PSM spectral calculations of the Lippmann photograph in the limit of long focal length. (a) A boxcar recording spectrum and the diffractive response of 1 μm -thick and 2 μm -thick Lippmann photographs. (b) Different recording and (c) replay spectra for a 2- μm -thick Lippmann emulsion ($n_1 = 0.03$, $n_0 = 1.5$ for all graphs).

constitutes a differential representation of the chain matrix method [6,7]. In 1991, Nareid and Pedersen [12] used the chain matrix method to analyse Lippmann photographs and compared their numerical solutions to a solution of the Helmholtz equation in the first Born approximation. In a recent publication on Lippmann photography, Kellerbauer [13] explains that when viewing a Lippmann photograph, there is a slight distortion of the recorded spectrum in which the most strongly pronounced parts of the original spectrum are additionally amplified.

12.13 Discussion

The main result of this chapter is that a general volume holographic grating may be conveniently and accurately described as an infinite set of infinitesimal parallel-stacked mirrors. At each of these mirrors, the classical laws of Fresnel reflection and transmission can be applied. In this way, a reference wave illuminating the grating gives rise to an infinite sum of secondary waves. By adding the waves together in a consistent manner, the electric field distribution within the grating and the diffractive response of the grating can be simply calculated.

Summing the waves from finite dielectric stacks is of course nothing new. For instance, Rouard [14] described this in 1937. In addition, the numerical chain matrix method [6,7] has been used with great success to calculate the optical properties of stratified media. What is important here, however, is that by using a *differential formulism*, the picture of parallel-stacked mirrors operating under Fresnel's laws offers a completely alternative description of the general holographic grating. Not only do we find that the PSM model is capable of largely reproducing, with very good accuracy, the results of Kogelnik's coupled wave theory through an abundance of analytical expressions for the diffractive efficiency, but

the model is also naturally capable of treating multicolour holographic gratings, spatially multiplexed gratings and even full-colour holograms!

REFERENCES

1. H. Kogelnik, "Coupled wave theory for thick hologram gratings," *Bell Syst. Tech. J.* **48**, 2909–2947 (1969).
2. D. Brotherton-Ratcliffe, "A treatment of the general volume holographic grating as an array of parallel stacked mirrors," *J. Mod. Optic.* **59**, 1113–1132 (2012).
3. D. Brotherton-Ratcliffe, "Analytical treatment of the polychromatic spatially multiplexed volume holographic grating," *Appl. Opt.* **51**, 7188–7199 (2012).
4. J. E. Ludman, "Approximate bandwidth and diffraction efficiency in thick holograms," *Am. J. Phys.* **50**, 244 (1982).
5. A. Heifetz, J. T. Shen and M. S. Shariar, "A simple method for Bragg diffraction in volume holographic gratings," *Am. J. Phys.* **77**, 623–628 (2009).
6. F. Abeles, "Recherches sur la propagation des ondes électromagnétiques sinusoïdales dans les milieux stratifiés. Application aux couches minces," *Ann. Phys. (Paris)* **5**, 596–640 (1950).
7. O. S. Heavens, "Optical properties of thin films," Reports on Progress in Physics, Vol XXIII, (1960) p. 1.
8. R. Guenther, *Modern Optics*, John Wiley and Sons, Hoboken, NJ (1990).
9. M. G. Moharam and T. K. Gaylord, "Chain-matrix analysis of arbitrary-thickness dielectric reflection gratings," *J. Opt. Soc. Am.* **72**, 187–190 (1982).
10. D. W. Diehl and N. George, "Analysis of multitone holographic interference filters by use of a sparse Hill matrix method," *Appl. Opt.* **43**, 88–96 (2004).
11. M. G. Moharam and T. K. Gaylord, "Rigorous coupled wave analysis of planar grating diffraction," *J. Opt. Soc. Am.* **71**, 811–818 (1981).
12. H. Nareid and H. M. Pedersen "Modeling of the Lippmann color process," *J. Opt. Soc. Am.* **8**, 257–265 (1991).
13. A. Kellerbauer, "Farbbilder aus gefrorenem Licht," *Phys. Unserer Zeit* **41**, 16–22 (2010).
14. M. P. Rouard, "Etudes des propriétés optiques des lames métalliques très minces," *Ann. Phys. (Paris)* Ser. II **7**, 291–384 (1937).

13

Illumination of Colour Holograms

13.1 Introduction

In this chapter, we discuss how holograms can be displayed. We shall pay particular attention to colour holograms. Often, the term *reconstruction* is used to describe the display of a hologram. This refers to the fact that if the reference beam, which is used to record the hologram, is used to also illuminate the processed hologram plate, the wavefront emitted from the object during recording will be faithfully *reconstructed*, generating the holographic image. This underlines the critical fact that to obtain a faithful holographic reconstruction, the properties of the reference recording and reconstruction beams must usually be identical. Indeed, it is well established in the literature that proper illuminating sources must be used to display holograms; this was already stressed as early as 1971 by Collier et al. [1].

The illumination of holograms is nevertheless a topic that is seldom included in books on holography; but it is as important as the techniques to record holograms. Remember that holography is a *two-part process* in which the holographic plate, with its recorded interference pattern, constitutes only one of the two parts needed to create the holographic image. The second part is the reference light, which is used to both record and replay the hologram. The characteristics of this reference light fundamentally control the display process. Only if the properties of the reference and replay lights are identical (spatial coherence, ray divergence, angle of incidence and wavelength) will a distortion-free, correct holographic image be generated.

13.1.1 Chromatic and Source-Size Blurring

In most practical situations, it is rare that the replay light is *absolutely* identical to the recording light. In particular, conventional broadband white-light sources are frequently used to illuminate reflection holograms. Here, one uses the wavelength-selective properties of the reflection hologram to filter out light having a different wavelength from that of the recording light. However, this comes at a price, as we have studied in Chapter 11. Illumination by a broadband source will bring with it chromatic blurring and, in general, the clear image depth of the hologram will be compromised.

One of the most important properties of any light source used to illuminate a hologram is spatial coherence. When a hologram is recorded, the reference laser light emanates from a spatial filter having a diameter of between 10 and 25 µm. This is several orders of magnitude smaller than the source size of current popular illumination sources. For example, halogen spotlights commonly used to display reflection holograms have source sizes from one to several centimetres. As we discussed in Chapter 11, a finite spatial coherence* in the reconstruction source leads to source-size blurring (Figures 13.1 and 13.2), which limits the clear image depth of the hologram. If our aim is to be free of such source-size blurring, a rough rule that can be applied is that the reconstruction source size must be less than 1 mm in diameter for every 1 m of diagonal distance that separates the illumination source and any point on the hologram. This ensures that any residual source-size blurring will be below the perception level of the standard human observer—as the average human eye can resolve image details of approximately 1 mm at a distance of 1 m.

* Strictly speaking, if the illumination source is temporally coherent, one should talk about the étendue of the source rather than its spatial coherence.

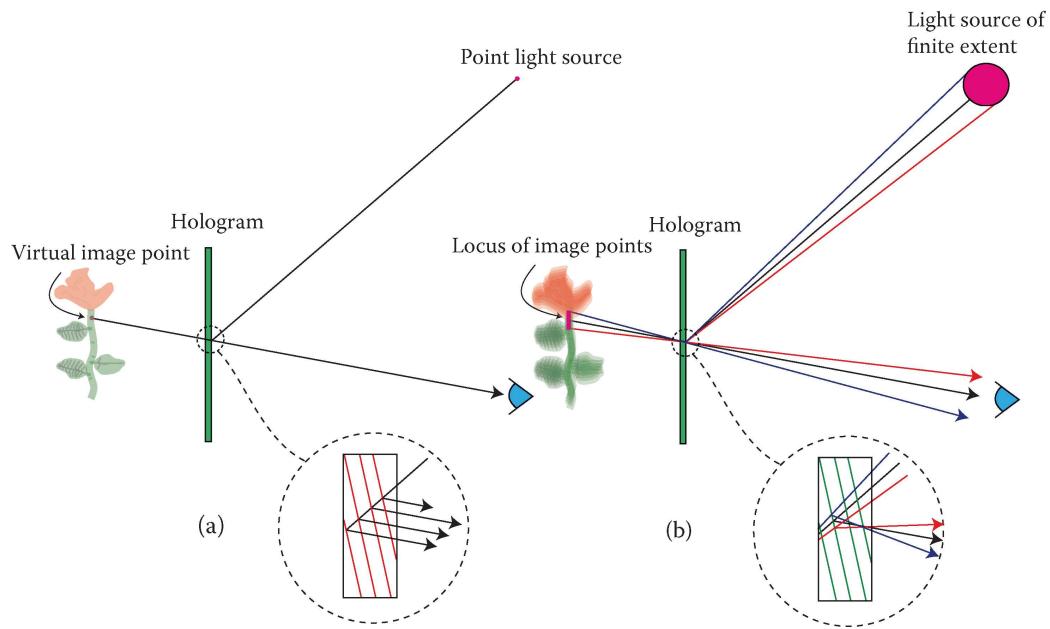


FIGURE 13.1 Source-size blurring in a reflection hologram causes the reflection of rays from the grating planes emanating from different parts of the source. (a) The case for a point source. Here, a single ray from the source is reflected from the grating to create a unique virtual image point. (b) The corresponding case for a source of finite size. Now, there exist various rays connecting the original point on the hologram to the light source. Each of these rays is reflected by the same grating at slightly different angles (angle of incidence = angle of reflection). This causes a blurred image.

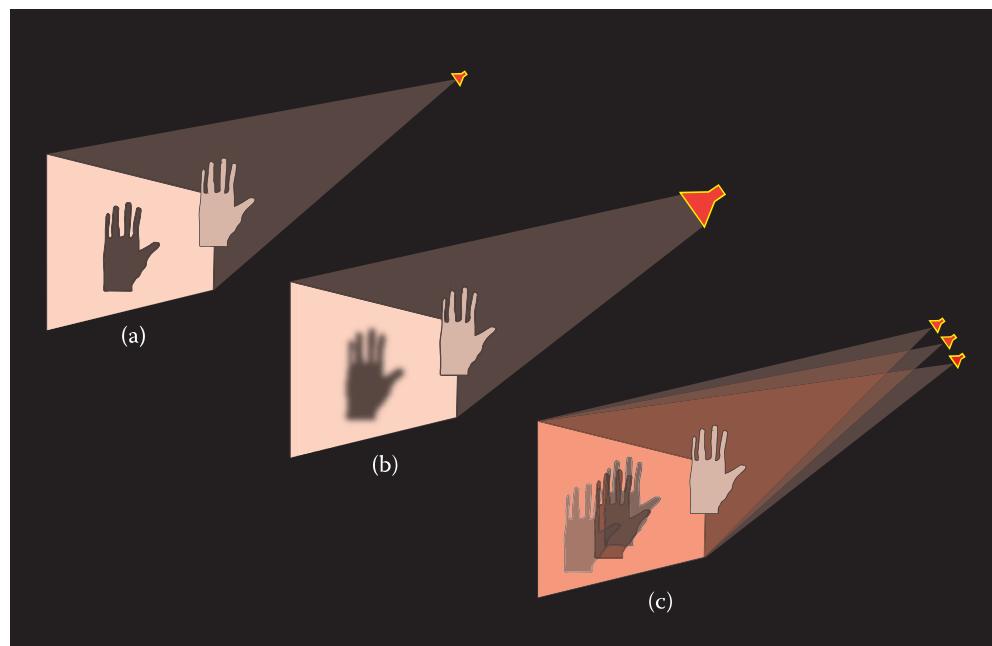
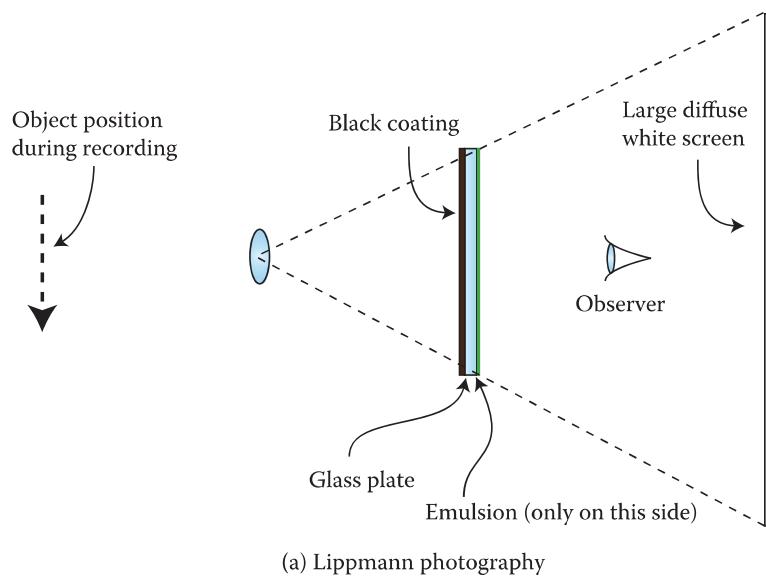


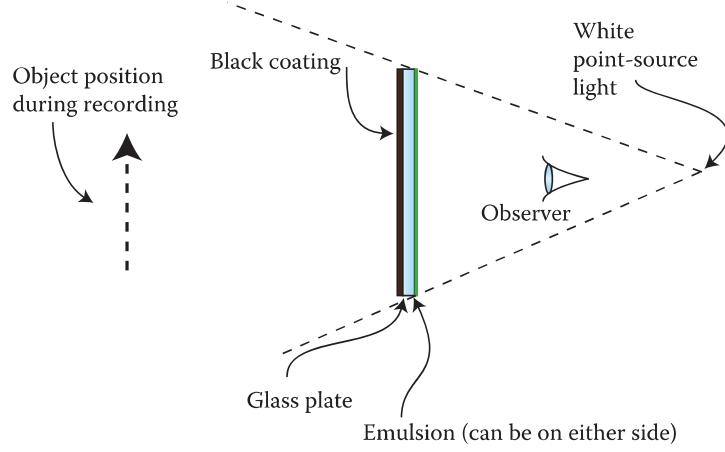
FIGURE 13.2 Layperson's diagram showing the importance of using a point source light to illuminate a hologram. Here, we make an analogy to the shadow cast by a spotlight on a screen. (a) The spotlight emits light from a small point, resulting in a sharp image. (b) A large-area diffuse spotlight results in a diffuse image. (c) Many spotlights illuminate the hologram, resulting in multiple images.

It is interesting to note that Lippmann photographs present a very different behaviour from holograms with regard to the illuminating source size. Here, one needs a fundamentally large and diffuse white-light source for their illumination (Figure 13.3).

The bandwidth of the light source also exerts a fundamental influence on the displayed image quality. Reflection holograms act as discriminating filters, but this intrinsic Bragg filter is usually rather broad. Certainly, for popular photosensitive materials, it is narrow enough to effectively stop any cross-talk between the primary colours of a three-colour reflection hologram, but it is very unusual for it to be narrow enough to annul chromatic blurring. For most commercial silver halide materials, the full-width half-maximum (FWHM) of the Bragg filter is approximately 15 to 25 nm. When a broadband source such as a halogen spotlight is used to illuminate a hologram, it is this filter that determines the clear depth of the hologram if source-size blurring is less than chromatic blurring. As we discussed in Chapter 11, a rough rule is that to be visually free of chromatic blurring, the bandwidth of the illuminating source must be less than 1 to 2 nm. As we shall see in later sections, a new paradigm is likely to operate in the future concerning the illumination of ultra-realistic reflection holograms such as high virtual volume



(a) Lippmann photography



(b) Reflection holography

FIGURE 13.3 The difference between displaying (a) a Lippmann photograph and (b) a reflection hologram, showing that a large-area diffuse light is needed to display the Lippmann photograph, but the light source for a hologram has to come from a point source which can be located at infinity.

(HVV) displays. Currently, if source-size blurring is controlled, it is the hologram, through its wavelength discriminating Bragg grating, which determines the clear image depth. However, with new laser diode illumination sources, which have a bandwidth of 1 to 2 nm, it is now the intrinsic properties of the source and not the hologram that are likely to determine the clear image depth.

13.1.2 Geometry Matching

The illumination reference source may be different from the recording reference source in its ray geometry. For example, it is often the case that holograms are recorded with collimated reference beams, but it is rather rare that a hologram is illuminated with such a beam. More likely, an approximation is sought to the original recording geometry and a point illumination source is used at some distance from the final hologram. For small holograms, this is usually a workable solution because the difference in ray angles is rather small if the light is placed at a distance of, for instance, five to ten times the diameter of the hologram. However, for larger displays, it clearly becomes difficult to place the light at a great distance from the hologram. Take, for example, a $1\text{ m} \times 1\text{ m}$ reflection hologram recorded with a collimated reference beam. Ideally, one would like to situate the light at a diagonal distance from the hologram centre of at least 5 m, but this is often just not practical. As we have discussed in Chapter 7, the issue is with variable reference beam systems in digital printers. Here, it is not much more difficult to implement the proper recording reference beam geometry such that the hologram can be replayed accurately with a close point source. This is also the *great advantage* of the direct-write digital technique. On the other hand, for certain analogue holograms, this is definitely not the case: for example a $1\text{ m} \times 1\text{ m}$ analogue H₂ reflection hologram would require a very expensive off-axis parabolic mirror to produce the correct reference beam.

If the illuminating reference beam differs substantially from the recording reference beam, aberration is introduced into the hologram. Such aberration is both chromatic and geometric in that it distorts the colours and the geometry of the recorded image. The human observer can usually tolerate a fair amount of geometric aberration without this creating too much of a problem. However, this does depend on the image in question. For example, an average observer is fairly resilient to aberration in holographic por-

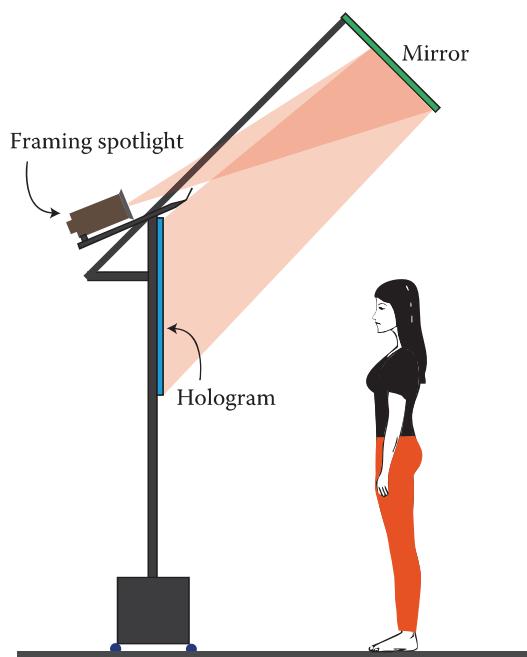


FIGURE 13.4 For large holograms which have been recorded with a collimated reference beam, there is often no other choice than to use a folded illumination beam path to attain a greater distance from the light source to the hologram. The diagram shows a large reflection hologram illuminated by a high-power framing spotlight.

traits, but where a hologram contains the image of a geometrical shape, such as a sphere, the eye will very quickly perceive the replayed image as an ellipsoid.

It is important to realise that induced aberration is always worse at the sides of a hologram. It is usually a fairly simple matter to match the ray directions in the centre by simply locating the spotlight at the correct position and angle. The induced aberration due to reference beam disparity can nonetheless have serious consequences regarding the entire image. If the light source is too close on replay, the image projection will be pushed out and magnified and the image behind the plate will be squashed. Large-depth holograms will therefore exhibit unnaturally small depths, and large-projection holograms will exhibit anomalous projections that can potentially make the image undecipherable. Colour holograms are particularly susceptible to reference beam disparity because the induced chromatic aberration means that colours will be altered with the observation location. Although this may be regarded as acceptable in a monochromatic reflection hologram, this is almost always not the case with a full-colour hologram.

Matching the ray geometry on recording and replay should therefore be taken extremely seriously; even more so for full-colour reflection holograms. For digital holograms, the best solution is to use a printer fitted with a variable reference beam system. For small holograms, either analogue or digital, a distant light provides a workable solution. For large analogue holograms, there are two solutions. The first is to use relay mirrors on replay to attain a larger distance from the light to the hologram (Figures 13.4 and 14.27). The second is to record the holograms in smaller panels (each with its own reference beam) and to use individual framing spotlights to illuminate each panel. We shall discuss this second option a little later in the context of HVV displays.

13.1.3 Illumination of HPO Holograms

Horizontal-parallax only (HPO) holograms may be illuminated by more than one spotlight as illustrated in Figures 13.5 and 13.6. Because these displays do not encode parallax in the vertical dimension, arrays of vertically stacked spotlights may be used to increase both the brightness of the holographic image and its vertical field of view without the introduction of blurring. In such an arrangement, each of the

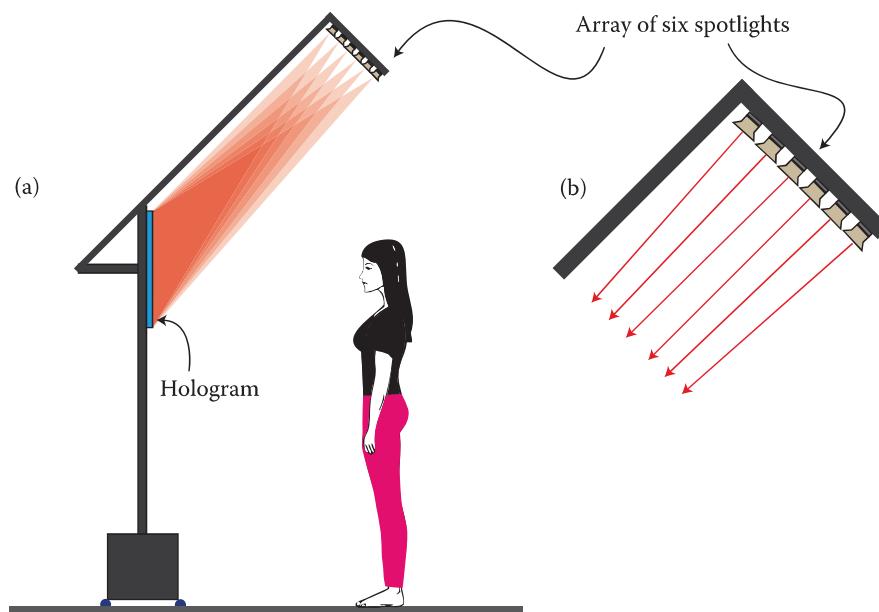


FIGURE 13.5 HPO reflection and rainbow transmission holograms can be illuminated by vertical stacks of broadband spotlights. Each of the lamps is directed towards the centre of the hologram. The centre lamp is usually arranged such that it corresponds to the reference beam geometry at recording. The lamps to either side therefore have an angle of incidence either above or below the recording angle. The Bragg condition is fulfilled for these lamps at a slightly different wavelength leading to a desaturation of the image. There is no induced geometrical distortion as the parallax information is exclusively in the horizontal dimension. (a) Display of HPO hologram with vertically stacked spotlights. (b) Detail drawing of the spotlights.



FIGURE 13.6 DWDH hologram and mounting stand made by Geola illuminated by a spotlight array comprising six individual 75 W MR16 halogen lamps.

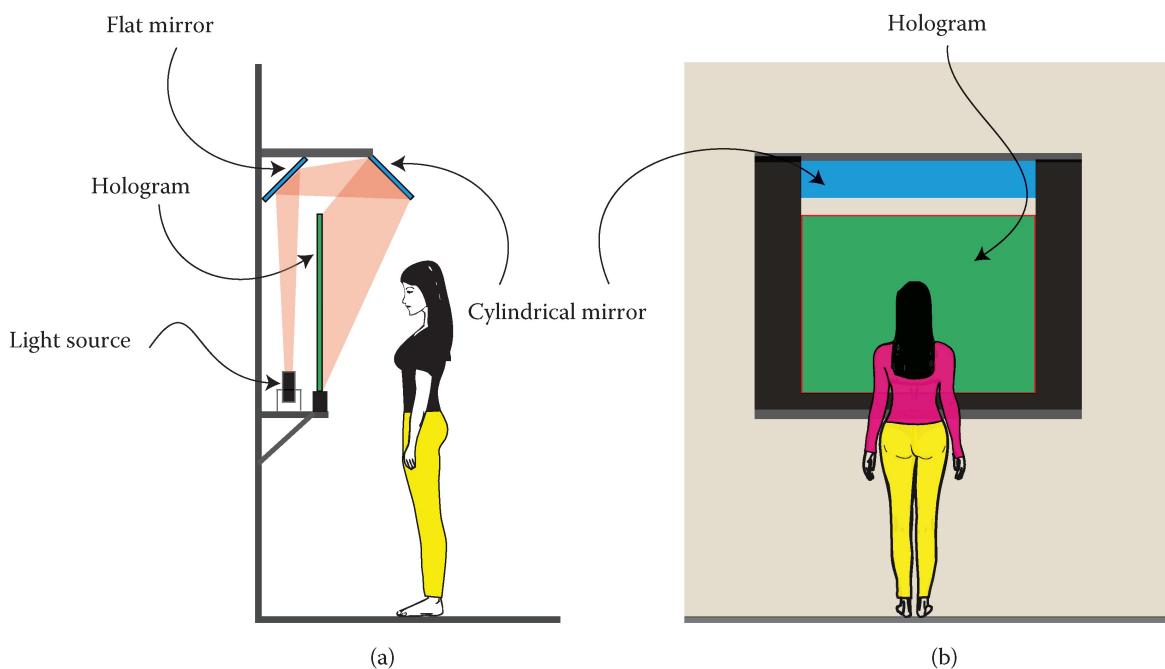


FIGURE 13.7 Compact illumination of an HPO hologram using a cylindrical mirror. Often, the angle of incidence of recording and illumination is a little larger than usual such that the mirror width can be minimised. Seen from (a) side and from (b) behind the viewer.

spotlights is directed towards the centre of the hologram. The central lamp is usually arranged such that it corresponds to the reference beam geometry at recording. The lamps to either side therefore possess an angle of incidence either above or below the recording angle. The Bragg condition is fulfilled for these lamps at a slightly different wavelength than the recording wavelength leading to a desaturation of the image and a greater vertical range of angles in which the hologram is visible. There is no induced geometrical distortion, as the parallax information exists exclusively in the horizontal dimension. In addition to arrays of lights, one-dimensional extended sources may also be used. The technique is applicable to all HPO holograms, whether of the transmission or of the reflective variety, and is extremely effective. One must be careful, however, not to use light of too small a bandwidth. Arrays of laser diodes, for example, will not work. Arrays of light-emitting diodes (LEDs) can be used, but the array length will be limited by their smaller inherent bandwidth.

Another technique that can be successfully applied to HPO holograms is the use of a cylindrical mirror. This is a version of the relay mirror technique in which the beam is collimated only in the horizontal direction, thus perfectly preserving the parallax information. The technique can be used to produce relatively compact displays (Figure 13.7).

13.1.4 Illumination of Large Rainbow and Rainbow-Achromatic Displays

Large rainbow-type holograms are particularly sensitive to the illumination geometry. As far as source-size blurring goes, they are no different. However, chromatic blurring operates in a rather different way in these holograms. The rainbow technique, if applied properly, allows large depth images to be attained by encoding the parallax information in the horizontal dimension and simply letting the vertical dimension chromatically disperse. This strategy works very well if there is a collimated reference beam on replay; however, if this is not present, then chromatic dispersion also occurs in the horizontal dimension, as we underlined in Chapter 11. Of course, large-depth HPO reflection holograms also suffer from chromatic blurring in the horizontal dimension under such circumstances—but there is very little Bragg wavelength discrimination in the transmission hologram, and this can make it more sensitive than the reflection hologram. However, more importantly, with the reflection hologram, there is always the option (at least in principle) to match the recording and replay reference beam geometries. The clear image depth (assuming a point source) is then determined by either Bragg wavelength discrimination or the bandwidth of the illuminating source. With rainbow-type holograms, however, one is forced to replay with a collimated beam—as only with a collimated beam will chromatic dispersion operate exclusively in the vertical direction. Transmission rainbow holograms are then, by their nature, rather more sensitive to the constraint of a properly collimated replay reference beam.

13.2 Illumination of Holograms by Laser Sources

In the early days of display holography, lasers were often used to illuminate monochrome transmission holograms. This resulted in spectacular images characterised by large clear image depths; in no uncertain way, this illumination technique led in large part to the creation of the public fascination for holography. The deep and detailed images available with laser illumination are due to the high temporal and spatial coherence characteristic of the laser source. The helium–neon laser was often used to display small or mid-size holograms. In early public exhibitions, other more powerful continuous wave (CW) lasers, such as argon–ion lasers, krypton–ion lasers, metal–vapour lasers and even tunable dye lasers were also employed. Soon enough, however, laser safety regulations were introduced, making it difficult for high-power lasers to be used for the display of holograms to the public.

Today, there are many new types of small, powerful and rather inexpensive semiconductor and solid-state lasers in the market that can be used to illuminate holograms. Provided that laser safety regulations can be observed, these sources promise to quite simply create a revolution in the illumination of display holograms. The main problem with laser safety is not the output power level itself but the spatial coherence (i.e., the fact that the light is emitted from a point source)—as it is this property of the light

that determines the size of the focussed spot at the retina. Of course, in one sense, a very high spatial coherence is desirable because one eliminates the most ubiquitous source of blurring in any hologram—source-size blurring. However, one only needs to get below a source size of approximately 1 mm per 1 m distance from lamp to hologram for source-size blurring to become imperceptible to the average human observer. As such, most lasers have an effective source size of several orders of magnitude smaller than required. This leads to an irradiance at the human eye, in the case of direct ocular exposure, of some hundreds of thousands of times greater than that needed.

Using lasers to display reflection holograms can completely circumvent chromatic blurring if the bandwidth of the laser is chosen to be less than approximately 1 to 2 nm. However, if the laser radiation is too (temporally) coherent, the problem of laser speckle arises. This is caused by the intrinsically coherent nature of laser light and results from interference at the surface of the retina. Generally, this becomes a problem at a bandwidth smaller than several nanometres [2]. Lee et al. [3] have reported that the appearance of speckle also depends on the brightness of the image, the illumination level of the room, the viewing distance and whether the image is moving or stationary. Furthermore, the perception of speckle depends on the wavelength, which can be explained by the sensitivity and the resolution of the eye: speckle is less visible in blue, than in the green and red spectra. The contrast in the image also influences the perception of speckle: in contrast-rich images, a given speckle contrast is less annoying than in a uniform image.

In addition to wavelength diversity, there are additional techniques that may be used to efficiently reduce speckle noise. The introduction into the laser beam of a moving diffuser is perhaps the best-known solution. This creates a dynamic ray angle diversity that averages out the speckle noise. For example, the company Optotune, Dietikon, Switzerland [4] now offers a unique commercial product based on this principle. Electroactive polymers (so-called artificial muscles) are used to laterally oscillate a small diffuser at high frequency, averaging out the speckle. The device is extremely compact, completely free of mechanics and has low power consumption. Optotune's laser speckle reducer (LSR) can be customised in terms of size, frequency, coatings and diffuser structure (Figures 13.8 and 13.9). The LSR consists of a diffuser bonded onto a polymer membrane that includes four independent dielectric elastomer actuators. When activated, the surface area of the electrodes increases and causes a motion of the rigid diffuser in the membrane plane. Four independent electrodes are used to obtain displacement of the diffuser in both lateral directions. The moving frequency is optimal when the mechanical resonance frequency of the system is attained and this provides the highest speckle reduction.

As well as reducing speckle, LSRs such as Optotune's reduce the spatial coherence of the source. For example, the LSR-3005 increases the effective source size at a 10° divergence angle to just under 5 mm. This corresponds to a fairly ideal value of étendue for a hologram illumination source, given the dual concerns of source-size blurring and laser eye safety.



FIGURE 13.8 A Laser speckle reducer from Optotune. (Photo courtesy of Optotune.)

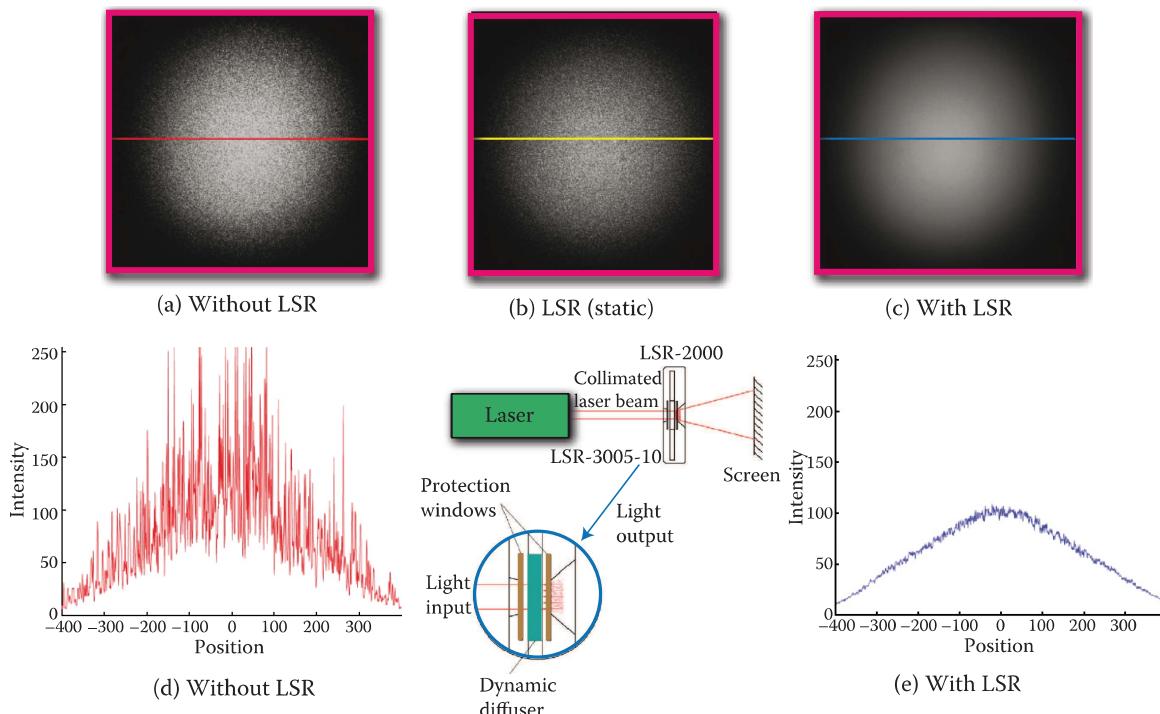


FIGURE 13.9 Optotune laser speckle reduction system. (a) Typical laser speckle patterns without LSR, (b) with LSR in passive mode and (c) with LSR switched on. Intensity profiles from an LSR-3005-10 device (5 mm aperture, 10° diffusion angle, 20 µm speckle; d and e). The schematic in the centre of the diagram illustrates how the LSR is used.

13.2.1 Importance of Wavelength Matching

An inconvenience of illuminating a reflection hologram with a laser source is that the laser wavelengths on recording and replay must usually be matched quite precisely. The precision of matching depends on the level of induced aberration that can be tolerated in a given image and also on the brightness reduction due to non-compliance with Bragg resonance. With some of the more modern photosensitive materials, very high index modulation can be achieved (see Chapter 4) and, as a result, high diffraction efficiency is possible from a thin emulsion with a very broad Bragg filter. With digital holography techniques such as DWDH, this opens up the possibility of actually accepting a small mismatch between the wavelengths at recording and replay but correcting for the induced aberration by image predistortion. Because the Bragg grating is broad in this case, its only function is then to avoid chromatic cross-talk, and as long as the difference in wavelengths is not so great, the diffraction efficiency of the hologram will be largely unaffected. Analogue holography does not allow for such a possibility and here one really is obliged to rigorously match the wavelengths at recording and replay.

Another point to mention about laser illumination is the deformation of the emulsion and the change of its refractive index on processing. Again, to maintain freedom from aberration and a maximum diffractive response, both the refractive index and the emulsion thickness should be invariant between the phases of recording and playback. Once again, digital holography does have the innate flexibility of image predistortion, which can be used to compensate for unavoidable changes in index and thickness. (See Appendix 4.)

13.2.2 Illumination of Full-Colour Transmission Holograms

Full-colour transmission holograms may be illuminated effectively by several lasers if each laser illuminates the hologram at a substantially different angle. Transmission holograms are generally more discriminating in reference beam angle than reflection holograms as we saw in Chapter 11. This makes

them more suitable for angle multiplexing just as the reflection hologram, having greater wavelength discrimination, is more suitable for wavelength multiplexing. The three different reference beams required for the illumination of a trichromatic transmission hologram can be organised at different altitudinal angles or different azimuthal angles depending on the amount of Bragg angle discrimination available from a given photosensitive material. The trichromatic or polychromatic transmission hologram is generally more difficult to illuminate than the corresponding reflection hologram. However, in certain circumstances, these holograms may offer several advantages. For example, the illuminating laser sources are always situated behind the hologram and so they are essentially hidden from view. In addition, the zeroth order beam may be prevented from passing through the display by a variety of techniques such as the inclusion of an additional reflective grating in front of the transmission grating. Trichromatic digital HGV displays have the advantage that they can be integrated into false walls, giving the impression that the display is completely self-illuminating.

Finally, we should mention that laser illumination is not suitable for rainbow-type displays. These types of displays require a broadband illumination.

13.2.3 Gas Lasers

In Chapter 3, different types of gas lasers, such as helium–neon lasers, argon–ion, krypton–ion lasers and metal-vapour lasers were described. With the possible exception of the He–Ne laser, and the copper-vapour laser these lasers are almost never used today to display holograms. This is because there are now new laser sources available that are much smaller and much cheaper. These are the semiconductor and solid-state lasers which we shall discuss below.

13.2.4 Semiconductor Laser Diodes and Solid-State CW Lasers

High-power semiconductor laser diodes are now becoming available at many wavelengths. The typical bandwidths of these ultracompact sources are in the order of several nanometres, making such diodes ideal for the illumination of holograms. In addition, ultracompact diode-pumped solid-state (DPSS) lasers, typically based on Nd:YVO₄ crystals with frequency doubling using potassium dihydrogen phosphate, lithium triborate or beta barium borate (see, for example, Figure 13.10) are now available, providing watt level TEM₀₀ emissions in red, green and blue. Light from these miniature lasers may be combined into a compact red, green, and blue (RGB) source far more easily than the lambertian light produced by LEDs. With prices at the time of writing ranging from \$50 to several hundred dollars per unit (depending on output power), compact laser sources are now a definite option for hologram installations requiring the best quality of illumination.



FIGURE 13.10 Example of a DPSS laser pointer producing watt level TEM₀₀ emission at 457 nm from an ultracompact low-power package (beam diameter, 5 mm; divergence, 1.5 mRad). (Photograph courtesy of Wicked Lasers.)

Solid-state lighting based on lasers offers significant advantages for illuminating holograms—most notably, the complete absence of source-size and chromatic blurring, leading to the real possibility of essentially infinite clear-image depths. However, it has also been suggested by Neumann et al. [5] that white light produced by a set of lasers of different colours could well be suitable for general illumination. In this recent study, an RYGB (red–yellow–green–blue) white laser light source was used to illuminate various objects and an experimental survey was conducted which compared people’s perception of the laser-illuminated objects with that of the same objects illuminated by other sources. The source was composed of the following four discrete laser lines with the following specifications:

- Red: 635 nm, 800 mW maximum power
- Yellow: 589 nm, 500 mW maximum power
- Green: 532 nm, 300 mW maximum power
- Blue: 457 nm, 300 mW maximum power

The study concluded that the four-colour white laser illuminant studied was *virtually indistinguishable* from high-quality state-of-the-art white reference illuminants. The major perceived advantage of a laser source is, of course, that it is likely to be rather more efficient than current LED sources. This is good news for holography, as we may realistically expect that a strong research effort directed at laser illumination sources for general illumination will also inevitably generate a much larger range of laser light sources suitable for holographic illumination. Another area in which RGB laser sources are being developed and which may be expected to drive the laser lighting industry is the field of laser projection for large-screen television and cinema [2].

13.3 Non-Laser Light Sources Used for Hologram Illumination

Today, most holograms are displayed by non-laser sources. This is both because of cost—the non-laser sources are usually substantially cheaper—but also because there are still relatively few groups capable today of making the type of large-depth holograms that can really benefit from laser illumination. The narrow-beam halogen lamp has historically been the most popular source for display holography and is still commonly selected for exhibitions. However, today there are new types of non-laser light sources that are more suitable for illuminating holograms. First among these sources and by far the most promising is the LED. It is interesting to note that modern LEDs can use as little as 15% of the energy required by incandescent lights (see Table 13.1). At this point in time, solid-state lighting based on blue-emitting InGaN LEDs and phosphors has demonstrated the highest luminous efficacy of any white-light source (265 lm/W), although this is only for small current densities. One of the key features of LED lighting devices is that many may last 10 to 20 years before needing replacement. The long LED lifetime is very important when illuminating holograms—remember that a hologram without illumination means no visible image!

Issues that are all too-often overlooked when selecting a light source to illuminate a hologram are

- Electrical efficiency of the source
- Generation of heat and infrared radiation

TABLE 13.1

Efficacy and Lifetime of Different Light Sources

Light Source	Efficacy (lm/W)	Lifetime (h)
Incandescent lamp	10–18	1500
Halogen lamp	12–24	2000
Compact fluorescent lamp (CLP)	40–50	5000
LED	70–90	40,000

- Temperature of source and associated fire risk
- Requirement for ventilation
- Generation of UV radiation

New light sources such as LEDs often have a positive influence on many of these aspects.

13.3.1 Halogen Lights

Halogen lights have, for a long time, been the light source of choice used to illuminate both monochrome and colour reflection holograms. The halogen lamp contains a small quantity of an active halogen gas such as bromine. The inert gas suppresses the evaporation of the tungsten filament, while the halogen gas acts to reduce the amount of tungsten that plates the interior wall of the lamp. The halogen gas reacts with the tungsten that has evaporated, migrated outward and has been deposited on the lamp wall. When the lamp wall temperature is sufficient, the halogen reacts with this tungsten to form tungsten bromide, which is then freed from the wall of the lamp and migrates back to the filament. The tungsten bromide compound reacts at the filament of the lamp where temperatures close to 2500°C cause the tungsten and halogen to dissociate. The tungsten is deposited on the filament and the cycle repeats. Unfortunately, the tungsten is not always deposited in the same zone at which evaporation initially took place and so the filament becomes thinner in places and eventually fails.

The light output of a halogen lamp is more stable than a non-halogen gas lamp due to the cleaning action of the halogen gas on the lamp envelope. This feature, coupled with the high colour temperature of the light and long-life, makes these lamps very desirable for many industrial and scientific applications.

Halogen lamps with aluminium reflectors, dichroic reflectors or focussed ellipsoidal dichroic reflectors are suitable for illuminating holograms. Most often narrow-beam 20, 50 and 75 W, 12 V MR11 or MR16 lamps are used. A typical 50 W lamp is shown in Figure 13.11a, and its spectrum is illustrated in Figure 13.11b.

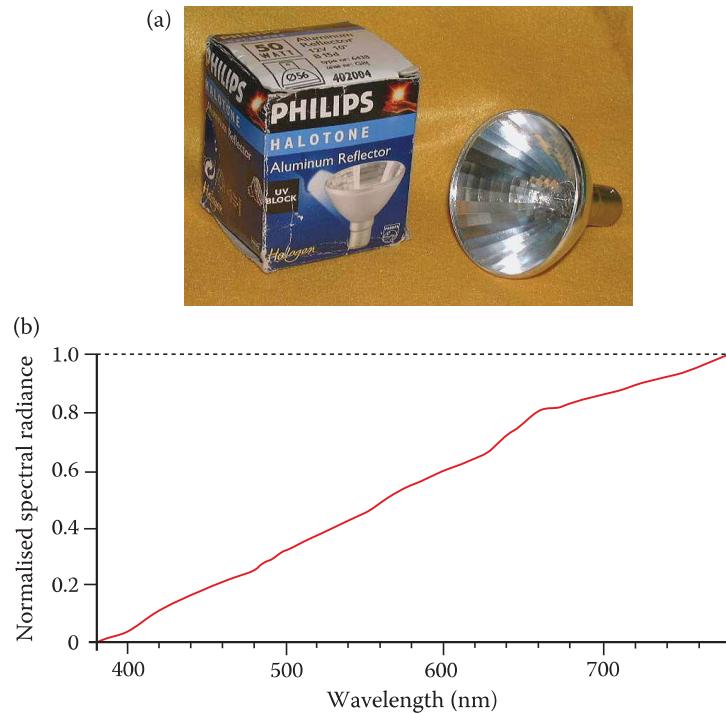


FIGURE 13.11 (a) A typical 50 W halogen lamp suitable for hologram illumination. (b) Spectrum of the emitted light.

13.3.2 Special Lamps

International Light Technologies, Peabody, MA [6] have many special types of lamps suitable for hologram illumination. Miniature lamps have advantages and should be considered for those applications requiring higher light output and better spatial coherence. Reflectors and lenses may be incorporated to shape the emitted light distribution.

For the illumination of larger holograms, the Electronic Theatre Controls Source Four Junior Zoom (also known unofficially as Source 4 or S4), which is an ellipsoidal reflector spotlight used in stage lighting, can be used quite effectively. This is an excellent, if bulky, source suitable for square-metre plus colour reflection holograms. It is a “framing projector” in the sense that the incorporated baffles and zoom may be used to arrange exclusive illumination of the rectangle of the hologram, eliminating any overspill of light. First released in 1992, the Source 4 features an improved lamp and reflector compared with previous ellipsoidal reflector spotlight designs, tool-free lamp adjustment, and a rotating, interchangeable shutter barrel. It also uses a faceted dichroic borosilicate reflector behind the lamp. The proprietary high-performance lamp uses four compact filaments and an ellipsoidal reflector system. It is rated at 575 W but produces light equivalent to a 1000 W spotlight using older technology. High-performance lamps are available at two colour temperatures: 3250 K (300–400 h lifetime) and 3050 K (1500–2000 h lifetime).

13.3.3 Mercury Lamps

As an alternative to lasers, in the early days of holography, short-arc mercury-vapour lamps were used to display off-axis monochrome transmission holograms. Combined with corresponding interference filters, the green (546.1 nm) or yellow-orange (578.2 nm) lines could be used (Figure 1.11, which shows a mercury lamp display cabinet). There are other similar lamps, such as metal-halide lamps, which use various compounds in an amalgam with the mercury in addition to lamps based on sodium iodide and scandium iodide. All these lamps can potentially be used for the illumination of holograms, but all suffer from severe problems. In addition, the use of mercury-vapour lamps for lighting purposes will be banned in the European Union in 2015 (the United States banned them in 2008).

13.3.4 Arc Lamps

Miniature arc lamps such as the Sōlarc lamps (Figure 13.12) available from Welch Allyn, Skaneateles Falls, NY [7] constitute an alternative to the common halogen lamp. The Sōlarc lamp is a metal halide light source in the class of high-pressure, high-intensity-discharge lights, which differs in a fundamental way from halogen, incandescent, fluorescent or LED illumination sources. Light is emitted from an arc discharge between two closely spaced electrodes, hermetically sealed inside a small quartz glass envelope. During operation, small amounts of metal are heated to a liquid state that provides the needed vapour to create a desired light colour. The light emitted from this arc tube is intense and generates more



FIGURE 13.12 The Sōlarc lamp from Welch Allyn. (a) Packaging, (b) lamp and (c) close-up showing arc gap.

than 60 lm/W. Sõlarc lamps feature a small, typically 1.2 mm arc gap, the smallest gap available in a metal halide arc lamp at the time of writing. Combined with elliptical reflectors, this arc gap allows one to focus illumination with laser-like precision into very small areas. The low-voltage (9–16 V) miniature arc lamp emits more blue light than halogen lamps. The colour temperature is between 6500 and 13,700 K. It has a very short arc and is equipped with a parabolic reflector, providing a narrow beam. The small intrinsic source size and a well-designed reflector translate into a somewhat sharper holographic image. The drawback is the rather short lamp life (350–700 h).

13.3.5 Plasma Lamps

Continuous full-spectrum lighting can be generated by a sulphur plasma light source. A typical emission spectrum from this type of illuminant is shown in Figure 13.13. The spectrum is very similar to the light from the sun. Plasma lamps use an electrode-less bulb design. Externally generated high-frequency electromagnetic energy is used to generate plasma within the bulb, converting the high-frequency energy into light in an efficient manner. The high-frequency energy (microwaves) is generated by a magnetron, which is powered from an adjustable stabilised power source. A prototype sulphur plasma lamp is shown in Figure 13.14. This unit, although bulky, can emit light with better spatial coherence than is normally

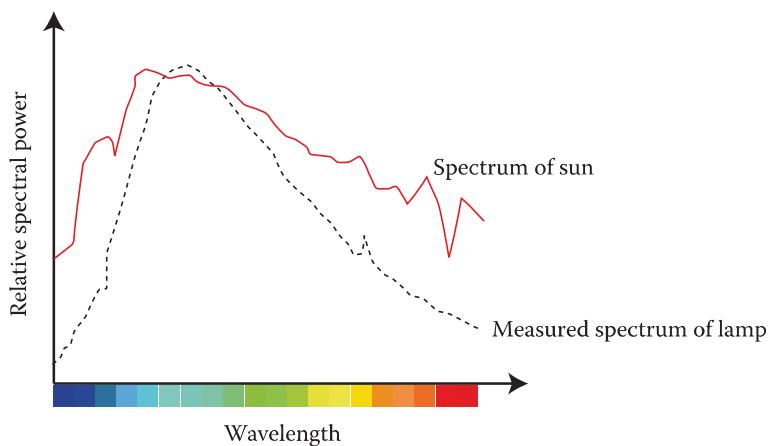


FIGURE 13.13 Typical sulphur plasma lamp spectrum.

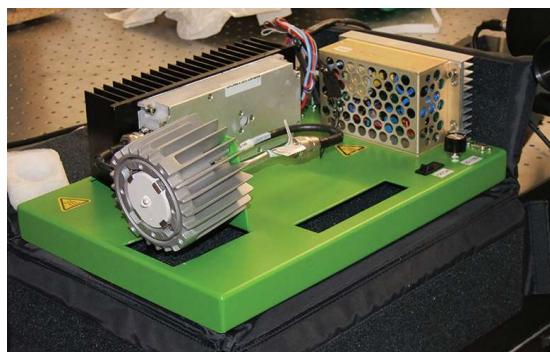


FIGURE 13.14 A prototype sulphur plasma lamp.

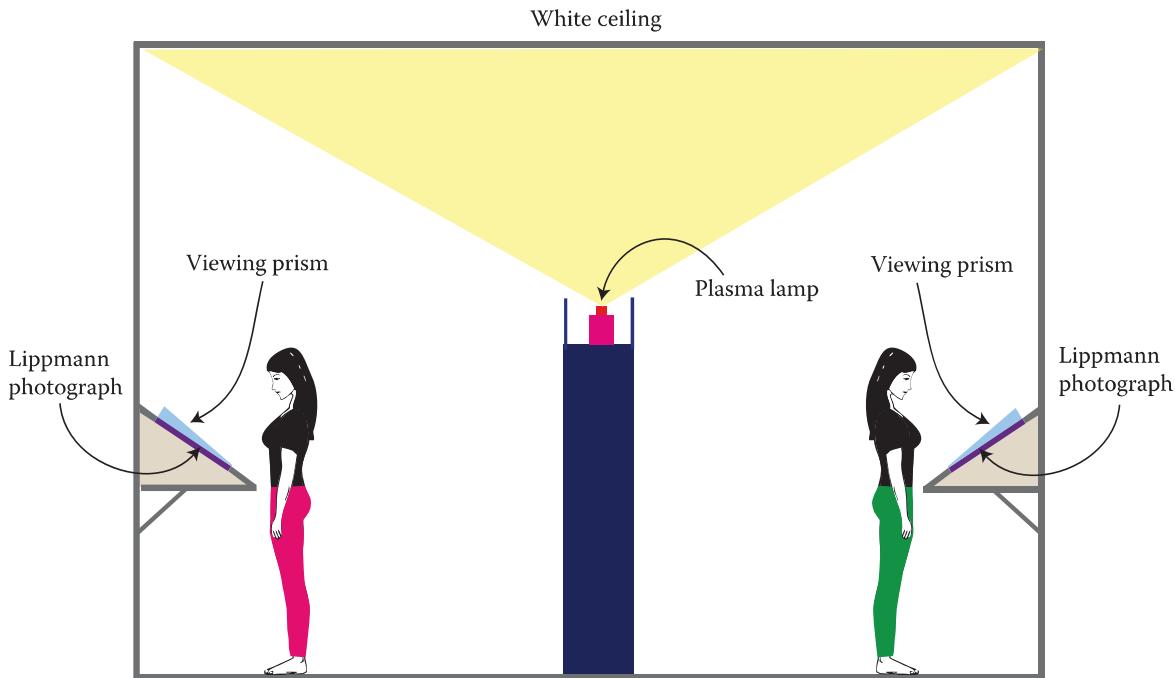


FIGURE 13.15 Display of Lippmann photographs using a sulphur plasma lamp. Most Lippmann photographs are recorded in sunlight and, as such, sulphur lamps, with their solar-like spectrum, provide an ideal illumination source, assuring optimal reconstruction of colours.

the case from many other commercial lighting sources. The colour rendering index (CRI) of the source is between 98 (60 lm/W) and 80 (140 lm/W).

One manufacturer of such light systems is Plasma International Lighting Systems [8] in Germany which markets two variations of the plasma lamp: the Standard Sulphur lamp (SS0) and the Triple A class Solar Simulator lamp (AAA). The lamp power can be set by software to a desired level between 500 and 1300 W, and the light spectrum remains almost unchanged between 600 and 1300 W power setting.

This type of light is not really merited (principally due to its high cost, bulky size and its broadband spectrum) for displays of reflection holograms, but it is an excellent (if expensive) source for large-format rainbow holograms. Its most interesting application is, however, for displaying Lippmann photographs. If the light source is made to illuminate the white ceiling in a gallery, many Lippmann photographs can be displayed in a truly excellent fashion according to the arrangement shown in Figure 13.15.

13.3.6 LEDs

The past few decades have seen a continuing and rapidly developing race among manufacturers of LEDs to produce ever cheaper and ever more efficient illuminants. This process is now producing a fundamental transformation in the field of general lighting. These miniature semiconductor devices will undoubtedly lead to the obsolescence of the common incandescent light bulb in the near future.

The rapid progress in solid-state LED lighting has opened up new possibilities to illuminate colour reflection holograms. A significant advantage of LEDs is that they possess a much smaller bandwidth than broadband white-light sources. Although typical bandwidths are much larger than those commonly associated with lasers and laser diodes, LED light sources should nevertheless be matched to the recording laser wavelengths (or vice versa). This guarantees that only the white light from the LED source (which is a mixture of the primary LED wavelengths) contributes to creating the holographic image. Using a halogen spotlight, a large part of the light spectrum emitted illuminates the surface of the plate without having any effect on the intensity of the image. Instead, this light is scattered, lowering the image

contrast. The lack of this scattered light in LED illumination can lead to significantly higher image fidelity. In addition, LED light sources have considerable advantages over halogen and other traditional lighting sources, such as

- Long life (20,000 to 100,000+ h)
- Small size
- Small étendue
- High durability and robustness to thermal and vibration shocks
- Low energy usage/high energy efficiency
- No infrared or UV in beam output
- Directional light output
- Digital dynamic colour control—white point tunable
- Relatively low cost

13.3.6.1 Theory of LED Operation

LEDs are semiconductors that convert electrical energy into light energy. The colour of the emitted light depends on the semiconductor material and on its composition. LEDs are formed from various doped semiconductor materials in the form of a PN junction. When an electrical current passes through the junction in the forward direction, the electrical carriers give up energy in the form of photons at a level proportional to the forward voltage drop across the diode junction. The amount of energy is relatively low for infrared or red LEDs. Although LEDs are semiconductors and need a minimum voltage to operate, they are still diodes and need to be operated in a current mode. The band-gap of the semiconductor determines the wavelength of the emitted light. Shorter wavelengths equate to greater energy and therefore higher band-gap materials emit at shorter wavelengths. High-efficiency LEDs can be produced in many wavelength ranges. Each material technology has a peak efficiency within the operational wavelength range, but it is difficult to make high-efficiency LEDs that operate at the edge of the material technology.

The semiconductor material typically takes the form of a very small chip or die, which is mounted onto a lead frame and encapsulated in a clear or diffused epoxy. The shape of the epoxy and the amount of diffusing material in the epoxy control the light output angle of emission. The construction of a common LED package is illustrated in Figure 13.16.

The output of high-power LEDs is typically expressed either in terms of luminous flux, which is measured in lumens, or in terms of radiant flux in watts. Alternatively, it is sometimes expressed as luminous

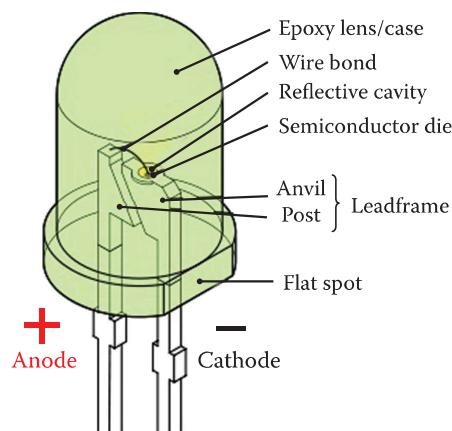


FIGURE 13.16 Common LED package.

intensity (in candela) or radiant intensity (in watts per steradian). If specified as an intensity, the value is usually measured along the projection axis of the device. The viewing angle for LEDs is specified as the included angle between the half-intensity points on either side of the output beam (Figure 13.17).

The wavelength of light emitted is determined by the difference in energy between the recombining electron-hole pair of the valence and conduction bands. The approximate energies of the carriers correspond to the upper energy level of the valence band and the lowest energy of the conduction band because of the tendency of the electrons and holes to equilibrate at these levels. Consequently, the wavelength, λ , of an emitted photon can be approximated by the following expression:

$$\lambda = \frac{c}{h} E_{\text{bg}} \quad (13.1)$$

where h represents Planck's constant, c is the velocity of light, and E_{bg} is the band-gap energy. To change the wavelength of emitted radiation, the band-gap of the semiconducting material used to fabricate the LED must be changed. Gallium arsenide is a common diode material and may be used as an example, illustrating the manner in which a semiconductor's band structure can be altered to vary the emission wavelength of the device. Gallium arsenide has a band-gap of approximately 1.4 eV, and emits in the infrared at a wavelength of 900 nm. To increase the frequency of emission into the visible red region (~650 nm), the band-gap must be increased to approximately 1.9 eV. This can be achieved by mixing gallium arsenide with a compatible material having a larger band-gap. Gallium phosphide, having a band-gap of 2.3 eV, is a candidate for this mixture. LEDs produced with the compound GaAsP (gallium arsenide phosphide) can be customised to produce band-gaps of any value between 1.4 and 2.3 eV, through adjustment of the content of arsenic to phosphorus. A major development occurred in the late 1980s, when LED designers borrowed techniques from the rapidly progressing laser diode industry, leading to the production of high-brightness visible light diodes based on the indium–gallium–aluminium–phosphide (AlGaInP) system. Again, this material allows changes in the emission colour by adjustment of the band-gap. The same production techniques can be used to produce red, orange, yellow, and green LEDs. More recently, blue LEDs have been developed based on gallium nitride and silicon carbide materials. One of the most important aspects of a blue LED is that it completes the RGB primary colour family to provide an additional mechanism of producing solid-state white light, through the mixing of these component colours. In particular, the indium gallium nitride (InGaN) system is the leading candidate for the production of blue LEDs and is also a primary material in the development of white LED lights.

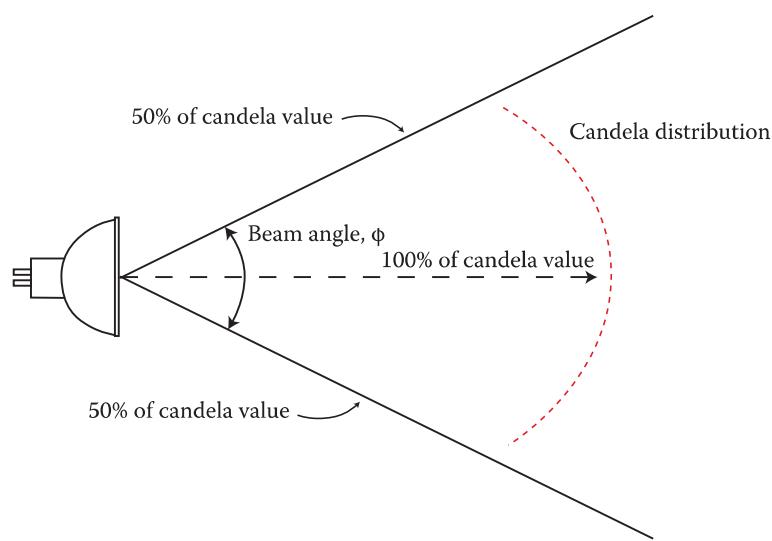


FIGURE 13.17 When luminous or radiant intensity is used to specify the emission characteristic of an illuminant, the candela value on-axis is quoted. The beams angle is then twice the angle from the axis to the half-intensity line.

13.3.6.2 Blue to Green LEDs: 450 to 530 nm

The material for this wavelength range of products is InGaN. Most large suppliers concentrate on creating blue (450–475 nm) LEDs for making white light with phosphors, and green LEDs that fall into the 520 to 530 nm range for green traffic signal lighting. Rapid advancements and improvements in efficiency are noted in the blue wavelength range especially as the race to create brighter white illumination LEDs continues.

13.3.6.3 Yellow-Green to Red LEDs: 565 to 645 nm

Aluminium indium gallium phosphide (AlInGaP) is the semiconductor material used for this wavelength range. It is predominately used for amber traffic signals (590 nm) and red (625 nm) lighting. The lime green (or yellowish-green 565 nm) and orange (605 nm) are also available from this technology, but they are somewhat limited. The technology is rapidly advancing, for the red wavelength in particular, because of the growing commercial interest in making red–green–blue white LED lights.

13.3.6.4 LEDs and Hologram Illumination—Wavelength Matching

In applications in which LEDs are used for illuminating colour holograms, the peak wavelengths can be as important—or even more important—than the output in lumens or candela. Many of the current photosensitive materials have a relatively important Bragg wavelength discrimination, meaning that if the wavelengths of the LED and the recording lasers are not matched fairly well, then the hologram will suffer from poor diffraction efficiency. The matching problem is certainly much less severe than for laser sources (the comments we made in Section 13.2.1 are valid here too), but nonetheless, it is still an important constraint that is just not present when a hologram is illuminated by a broadband source.

13.3.6.5 White Phosphor LEDs

Most white-light diodes use a semiconductor chip emitting at a short wavelength (blue, violet or ultra-violet) and a wavelength converter (usually a yellow phosphor) that absorbs light from the diode and produces secondary emission at a longer wavelength. Such diodes, therefore, emit light at two or more wavelengths; combined, these appear as white light. The phosphors typically used are composed of an inorganic host substance containing an optically active dopant. Yttrium aluminium garnet is a common host material, and for diode applications, it is usually doped with one of the rare earth elements or a rare earth compound. Cerium is a common dopant element in yttrium aluminium garnet phosphors designed for white LEDs.

Viewed directly, the LED will appear to be white as the blue and yellow wavelengths are mixed together. This product is ideal for general lighting. However, for illumination of holograms in which colour rendering is important, this type of LED is not the most suitable. Visible LED wavelength bands are listed in Table 13.2 with their corresponding general applications.

TABLE 13.2

Visible LED Wavelength Bands and Applications

Wavelengths (nm)	Main Application
430–470 (Blue)	White LEDs using phosphor, blue for RGB white lights
520–530 (Green)	Green traffic signal lights, green for RGB white
580–590 (Amber)	Amber traffic signal lights, amber for RGBA white lights
630–640 (Red)	Red signal lights, red for RGB white lights

13.3.6.6 Multicolour LEDs

By mixing together a variety of semiconductor, metal and gas compounds, the following list of LEDs can be produced.

- Gallium arsenide (GaAs)—infrared
- Gallium arsenide phosphide (GaAsP)—red to infrared, orange
- Aluminium gallium arsenide phosphide (AlGaAsP)—high-brightness red, orange-red, orange, and yellow
- Gallium phosphide (GaP)—red, yellow and green
- Aluminium gallium phosphide (AlGaP)—green
- Gallium nitride (GaN)—green, emerald green
- Gallium indium nitride (GaInN)—near ultraviolet, bluish-green and blue
- Silicon carbide (SiC)—blue as a substrate
- Zinc selenide (ZnSe)—blue
- Aluminium gallium nitride (AlGaN)—ultraviolet

The most popular type of tricolour LED comprises a single red and a single green LED combined in one package with their cathode terminals connected together, producing a three-terminal device. They are called tricolour LEDs because they can give out a single red or a green colour by turning “on” only one LED at a time. They can also generate additional shades of colours (the third colour) such as orange or yellow by turning “on” the two LEDs in different ratios of forward current.

13.3.6.7 LED Colour Rendering

Colour rendering is very important when LED lights are used in museum illumination of paintings and other works of art. Poor colour rendering is often evident when a hologram is illuminated by white LEDs. This is usually because the green and red components are too weak. With phosphor-pumped white LEDs, there is no smooth output-versus-wavelength behaviour, and this leads to errors in the colours perceived by the eye. Here, one uses the colour rendering index (CRI) as a quantitative measure of the ability of a light source to reproduce the colours of various objects in comparison with a natural light source. Getting the correct CRI is very important for such applications as photography, cinema, television and holography. The CRI on many high-quality lamps is between 85 and 90. This translates into the human eye actually seeing the true colour of the object being lit. The CRI of most traditional light sources can be as low as 25 to 50, with high-pressure sodium lamps being the worst with a CRI of approximately 25. In Table 13.3, the CRI for different LEDs show the high value of the trichromatic LED.

People who spend a lot of time in museums and galleries have become very accustomed to viewing art under halogen lights. Halogens have a colour temperature somewhere between 2700 and 3000 K, but this technology definitely adds warmth (yellow) to the light. With LEDs, control over the colour temperature of the light becomes much more precise. LEDs hold their colour temperature even though the lamps can be dimmed to approximately 5% of their light output. Traditional lighting in museums and galleries has often appeared rather dim because of the need to reduce damage to the art over time

TABLE 13.3

LED Efficiency and CRI

LED type	Luminous Efficiency (lm/W)	General CRI
Dichromatic LED	33.6	10
Broadened output dichromatic LED	30.6	26
Trichromatic LED	28.3	60
Phosphor-based LED	28.0	57

owing to the UV and heat generated by halogen lamps. With LEDs, it is now possible to light such fragile art a bit brighter owing to the fact that LEDs do not emit UV and will not cause heat damage to the art. This is also important when illuminating holograms. UV light can cause printout in silver halide holograms. And heat, over an extended period, can create emulsion shrinkage affecting the colour of the holographic image.

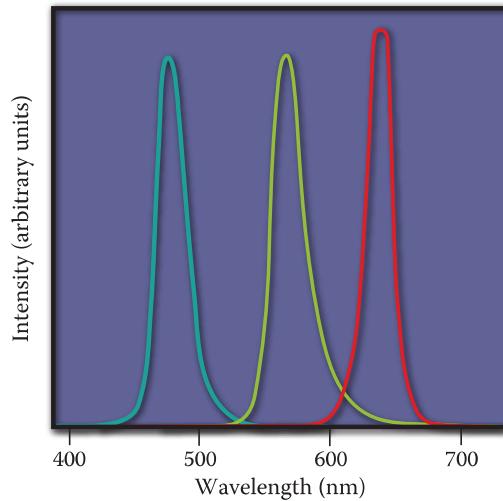


FIGURE 13.18 RGB LEDs combined technology: combined spectral curves for blue, yellow-green and high-brightness red solid-state semiconductor LEDs. The FWHM spectral bandwidth is approximately 24 to 27 nm for all three colours.

TABLE 13.4

Current LEDs Suitable for Holographic Illumination

	Power	Bandwidth (FWHM)	Dominant Wavelength	Emission Area
<i>Red Diode</i>				
Luxeon Rebel	720 mW	30 nm	655 nm	1.4 mm × 1.4 mm
Luxeon Rebel	102 lm	30 nm	627 nm	1.4 mm × 1.4 mm
Luminous SBT 16	155–300 lm	18 nm	623 nm	1.6 mm × 1.0 mm
Luminous PT39	390–750 lm	18 nm	623 nm	2.09 mm × 1.87 mm
Luminous PT54	600–1050 lm	19 nm	623 nm	2.7 mm × 2.0 mm
Luminous PT120	1360–2225 lm	19 nm	623 nm	4.6 mm × 2.6 mm
<i>Green Diode</i>				
Luxeon Rebel	161 lm	20 nm	530 nm	1.4 mm × 1.4 mm
Luminous SBT 16	315–600 lm	40 nm	525 nm	1.6 mm × 1.0 mm
Luminous PT39	815–1650 lm	36 nm	525 nm	2.09 mm × 1.87 mm
Luminous PT54	1275–2150 lm	36 nm	525 nm	2.7 mm × 2.0 mm
Luminous PT120	2700–4300 lm	36 nm	525 nm	4.6 mm × 2.6 mm
<i>Blue Diode</i>				
Luxeon Rebel	70 lm	20 nm	470 nm	1.4 mm × 1.4 mm
Luxeon Rebel	1120 mW	20 nm	448 nm	1.4 mm × 1.4 mm
Luminous SBT 16	60–110 lm	20 nm	460 nm	1.6 mm × 1.0 mm
Luminous PT39	1.5–2.75 W	20/25 nm	460/462 nm	2.09 mm × 1.87 mm
Luminous PT54	275–480 lm	20/25 nm	460/462 nm	2.7 mm × 2.0 mm
Luminous PT120	550–970 lm	20 nm	460/462 nm	4.6 mm × 2.6 mm

13.3.6.8 Preferred LEDs for Colour Holography

An alternative white LED technology to the phosphor-pumped LEDs is offered by RGB or RGBA LEDs. These combine red, green and blue or red, green, blue and amber chips onto one discrete package allowing the generation of white light or any of 256 colours by utilising circuitry that drives the three diodes independently. In applications requiring a full spectrum of colours from a single point source, this type of RGB diode format is the preferred type and is the most suitable for illuminating colour holograms. Currently, one of the best methods for the generation of white light is the “RGB LEDs combined” technology—that is, the generation of white light using a combination of red, green and blue. Note, however, that this form of white light relies on the electrical control of three LED chips. A spectrum from a typical RGB combined LED device is illustrated in Figure 13.18. Typical FWHM spectral widths are 24 to 27 nm.

When different-colour LED lights were first marketed in 2003, the OptiLED chip, a commercial LED spotlight, was introduced to illuminate monochrome reflection holograms. The 4° red (627 nm, linewidth 20 nm), green (530 nm, linewidth 35 nm) and amber (590 nm, linewidth 14 nm) versions provided suitable illumination for such holograms. The 2.5 W LED spotlight had an electric circuit allowing them to be operated at 90 to 240 V AC. Different interface types were offered: E27, E14, B22 and MR16, for example. These new LED lights were promoted and distributed by Laser Reflections in San Francisco, CA [9,10]. A list of some current LEDs suitable for holographic illumination is given in Table 13.4.

13.3.6.9 Lenses for LED Spotlights

It is important when illuminating holograms to be able to adjust the divergence of the illuminating source such that the illuminated area is best matched to the hologram size. Optotune has developed a new type of adaptive lens for LED lights [4], which is shown in Figure 13.19. The adaptive lens is a liquid polymer in a flexible container. The focal length of the lens can be controlled by electrically changing the thickness of the lens. The design includes an LED, a secondary optic and an adaptive condenser lens for the tuning of the spotlight. The secondary optic (usually a total internal reflection or TIR lens) defines the maximum angle of the spotlight. The secondary optic may have a diffuse layer on top, preventing the display of the lens or LED structure, but this may not always be warranted because it can increase the étendue of the source, thereby increasing source-size blurring in the hologram. The adaptive condenser lens is added as a tertiary optic. By tuning the adaptive condenser lens, the divergence of the spotlight can be changed.

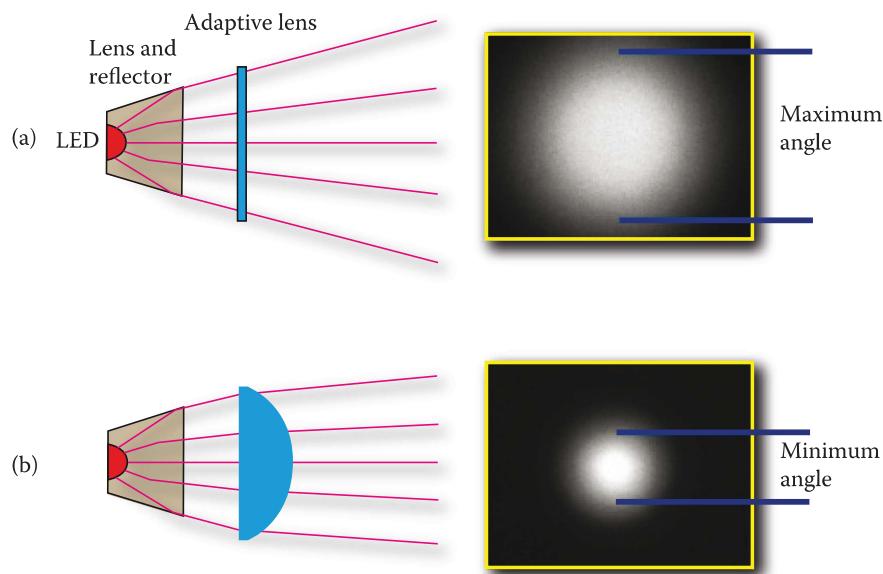


FIGURE 13.19 Adaptive lighting principle using a tunable condenser lens.

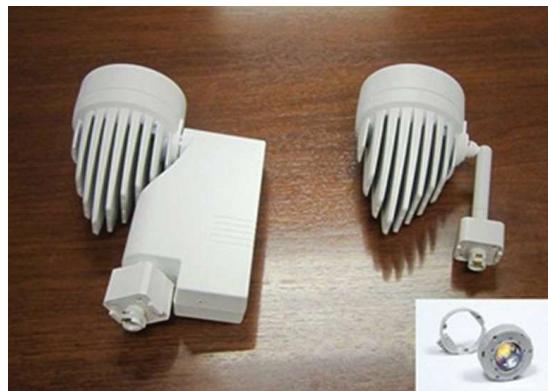


FIGURE 13.20 The Westport track lights based on Bridgelux/Molex Helieon modules.

Another company that offers special lenses for LED lights is Aether Systems Inc., Taipei, Taiwan [11]. Aether Systems designed and produced a digital lens for LEDs. The Aether Digital Lens is a high-performance lens for LEDs and produces excellent light shape.

The concept of focussing the light from one or more LEDs with a lens can actually be taken a few steps further. In Appendix 6, we summarise a design study of an RGB LED illuminant of high spatial coherence that uses a specially moulded aspheric lens of high index to achieve an almost perfect irradiance distribution at the hologram surface (see, for example, Figure A6.11). The technique allows an effectively top hat distribution of irradiance to be created at a defined distance from the illuminant with a defined boundary. For example, an LED illuminant of high spatial coherence can be constructed to accurately illuminate a tilted rectangular hologram with very little overspill of the light. To design such lens systems, one must use a computational procedure that optimises a desired boundary, a flat irradiance distribution and the final spatial coherence of the source. The problem is complicated by its non-linear nature, but the results are quite impressive. Clearly, this is a technique that can be applied to laser illumination as well. In fact, as will be clear from Appendix 6, rather more accurate framing of the hologram may be expected in this case.

13.3.6.10 White LED Spotlights

Not all white LEDs are suitable for the illumination of holograms. In particular, we have already discussed white phosphor-type LED spotlights. These are not really suitable for illuminating colour holograms. The spectrum of these devices is broadband with normally a low-intensity red region. The manufacture of these white LEDs involves coating the LEDs of one colour (mostly blue LEDs made of InGaN) with phosphor of different colours to form white light. If several phosphor layers of distinct colours are applied, the emitted spectrum is broadened. Due to the simplicity of manufacturing, the phosphor method is currently still the most popular method for making high-intensity white LEDs.

As we have already mentioned, LED spotlights suitable for illuminating colour reflection holograms should be constructed from separate RGB LED chips. This way of making white LED lamps is particularly interesting because of the inherent flexibility of colour mixing; the mechanism also has a higher quantum efficiency in producing white light. We shall review commercial spotlights based on this concept in Section 13.3.6.11.

There are, however, many white phosphor-type LED spotlights on the market that can be used for the illumination of monochromatic and rainbow holograms.* For example, the Bridgelux/Molex Helieon modules, Livermore, CA [12], which are compatible with Westport tracks, are shown in Figure 13.20. There are also modules already on the market that are based on the Zhaga standards [13].† One example

* One should nevertheless be mindful of the potential low red intensity of many of these illuminants.

† Zhaga is an industry-wide cooperation aimed at the development of standard specifications for the interfaces of LED light engines.

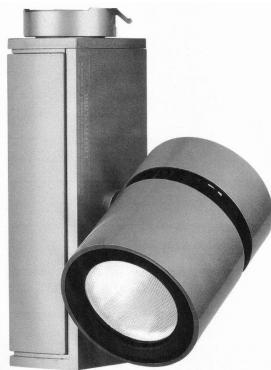


FIGURE 13.21 Lytespan Mini LED. (Photo courtesy of Philips Lightolier.)



FIGURE 13.22 Definity PAR38 LED lamp. (Photo courtesy of Lighting Science Group.)

is the socket-able LED light engine, which fits GE's Infusion module and Phillips' Fortimo spotlight module. Another example of a white phosphor-type LED spotlight is the Lytespan Mini LED manufactured by Philips Lightolier [14] (Figure 13.21), which offers excellent beam control. Interchangeable optics also provide flexibility to adjust the lighting. Finally, Lighting Science Group, Satellite Beach, FL [15] manufactures a new family of 18 to 24 W LED lamps, the Definity lamps, which are intensely bright (lumen output, 840–1460 lm) all in one package shown in Figure 13.22.

13.3.6.11 Special LED Spotlights for Colour Holograms

In Greece, the Hellenic Institute of Holography has developed a special LED spotlight to illuminate colour holograms. The HoLoFoS LED spotlight, based on Cree LEDs, is manufactured at AutoTech, Athens, Greece [16] and is commercialised by TAURUS SecureSolutions Ltd., Athens, Greece [17]. Through proper choice of the component LEDs in terms of bandwidth and wavelength, the HoLoFos LED spotlight is capable of achieving high-quality reproduction of deep full-colour reflection holograms.

The device consists of an illuminating head, extending arms and a mounting base. The illuminating head contains the RGB LEDs, mixing optics, lenses and heat sinks. The system has an embedded microcontroller for intensity control of each LED with DMX protocol decoding and a miniature wireless receiver. Remote adjustment of the colour mixing by DMX protocol communication is achieved by a handheld wireless remote control. A small switching power supply provides the power needed for EU or US mains.

The optics incorporated in the unit provide for an axial mixing of the LED beams resulting in a homogeneous colour mixing over the full extent of the projected beam. The small footprint of the LED die (~2 mm) is small enough to produce clear and deep holograms even at small illuminating distances.

The illuminating head can be fitted with a variety of LEDs at selected wavelengths and more than three different LEDs can be fitted to match various recording wavelengths. For example, an RRGGBB configuration can be achieved. This is important for colour holograms that will be recorded with four or five laser wavelengths to obtain more or less perfect colour rendering.

The current prototype unit uses three LEDs with the following spectral characteristics for the red, green and blue LEDs, which correspond to the lasers of the Z3 Holographic Camera (see Chapter 14).

- Red 620–630 nm
- Green 520–535 nm
- Blue 450–465 nm

The LED spectrum of the HoLoFoS illuminating system is shown in Figure 13.23 and the LED spotlight with its colour control box is shown in Figure 13.24.

To demonstrate the advantage of using this LED light for displaying colour holograms, the same hologram was illuminated with a conventional halogen spotlight and the new HoLoFoS LED light; the results are shown in Figure 13.25a and b. Note the increased contrast obtainable with the new LED light. This is also a good illustration of the vital importance of the illumination source in holography. Pulsed digital holograms, which are often printed using a red wavelength of 660 nm, require a version of the HoLoFoS light with a deeper red diode.

LED spotlights such as the HoLoFoS light can also be mounted in stacks for the illumination of HPO reflection holograms as described in Section 13.1.3.

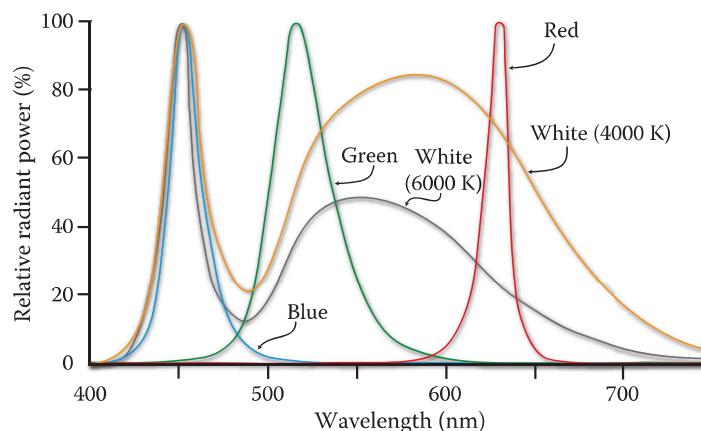


FIGURE 13.23 The LED spectrum of the HoLoFoS illuminating system.



FIGURE 13.24 The HoLoFoS LED spotlight for the illumination of colour holograms.

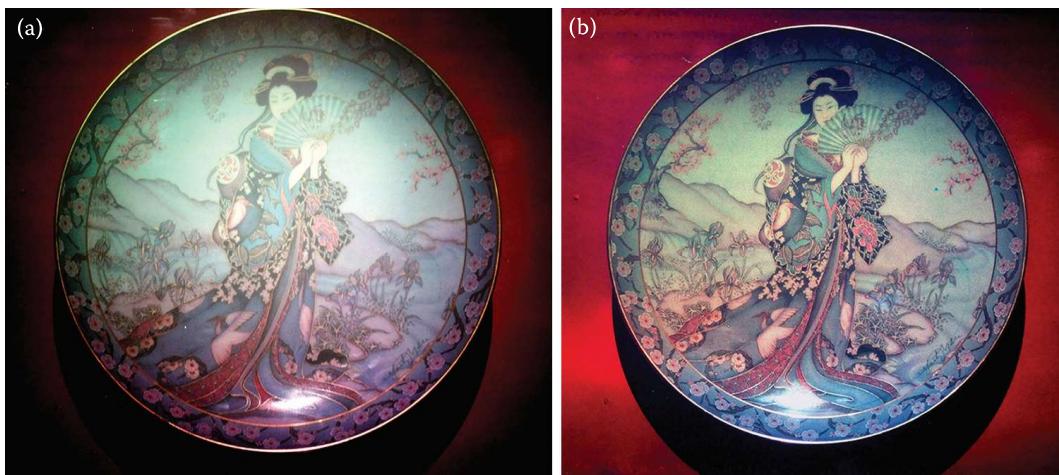


FIGURE 13.25 The same colour hologram illuminated (a) with halogen light and (b) LED light.

13.3.6.12 LED and Lamp Safety

It is important to be aware of the safety aspects of operating powerful light sources such as lasers and LEDs. If these sources are used for the public display of holograms, it is even more necessary to be aware of existing safety regulations. Laser safety regulations have existed for a long time, but because the new very bright LED lights also pose a risk hazard to people, new safety regulations for these products have been introduced. Initially, these standards were intended to be incorporated with the safety regulations for lasers, but the International Electrotechnical Commission (IEC) committee finally decided to remove LEDs from the scope of the IEC 60825-1:2007 laser safety standard. This was, in fact, an acknowledgement of the existing standard CIE S 009:2002, which addresses the photobiological safety of lamps and lamp systems, including LED sources. This standard was prepared by the International Commission on Illumination (CIE). The safe exposure limits in the CIE lamp safety standard are, like those in the IEC laser safety standard, based on the underpinning International Commission on Non-Ionizing Radiation Protection (ICNIRP) safety data and guidelines. Also, the measurement methods in CIE S 009:2002 are similar in format to those contained in the IEC laser safety standard. Unfortunately, the CIE standard does not provide detailed information on how to measure a source configured within an array, nor does it include the hazards posed by aided optical viewing, and there is also no information on product safety labelling requirements unlike the case for lasers.

Furthermore, the exposure measurement is undertaken at a fixed distance of 200 mm from the (apparent) source location and this approach may not be universally applicable to the evaluation of LED and LED array sources. The standard does, however, provide a clear delineation of the various ocular hazard bands and its methodology is directly applicable to broadband sources (such as a white HB-LED). To accommodate the removal of LED sources from IEC 60825-1, the IEC published a lamp and LED standard (IEC 62471:2006) in 2006, which was harmonised with (and indeed directly based on) the CIE S 009:2002 lamp safety standard. This relatively new IEC standard, which was released in 2008, is already undergoing revision to better reflect the needs of LED source hazard assessment and labelling. Although there is no specific information on labelling requirements, the lamp safety standard does contain its own risk classification scheme for potentially hazardous lamps, namely, exempt, low risk, moderate risk and high risk, which follows a similar reasoning with the IEC laser product classification. In summary, for the purposes of international trade of LED products, manufacturers and vendors of LED products should consider IEC 62471:2006 to be the currently applicable product standard for LED safety. Again, this is notwithstanding any specific national regulatory requirements or directives that refer to an alternative standard or assessment method such as the IEC 60825-1:2001 laser (and LED) safety standard.

Caution needs to be advised that a fixed assessment distance of 200 mm may not be sufficient to fully ascertain the maximum optical radiation hazard posed by the source and consideration might also need to be given to the effect of aided viewing (e.g., magnifiers and telescopes) on the hazard assessment.

13.4 Exhibition Facilities and Galleries Suitable for Displaying Holograms

A disturbing fact often associated with hologram exhibitions is the presence of many bright spotlights, as illustrated in Figure 13.26. If possible, one should try to avoid this situation when arranging hologram exhibitions. In addition, one should try to arrange exhibition facilities for holograms in such a way that it is easy to change and mount different holograms. The first task is to obtain a suitable type of spotlight, preferably with a lens system for the control of beam divergence. In addition, the spotlight tracks should be positioned perpendicularly to the walls (not parallel with) so that the illuminating angle of each spotlight can be easily adjusted. The tracks in existing galleries are often too close to the walls to be used for illuminating holograms properly.

In a specialised hologram gallery, the following design makes it possible to hide all the spotlights so the visitor cannot see them. The spotlights are arranged in the ceiling with a track system as mentioned previously. A suspended ceiling, underneath the ceiling on which the spotlights are mounted, should be



FIGURE 13.26 Many bright spotlights in the ceiling at a hologram exhibition can distract the viewer and detract from his or her viewing experience.

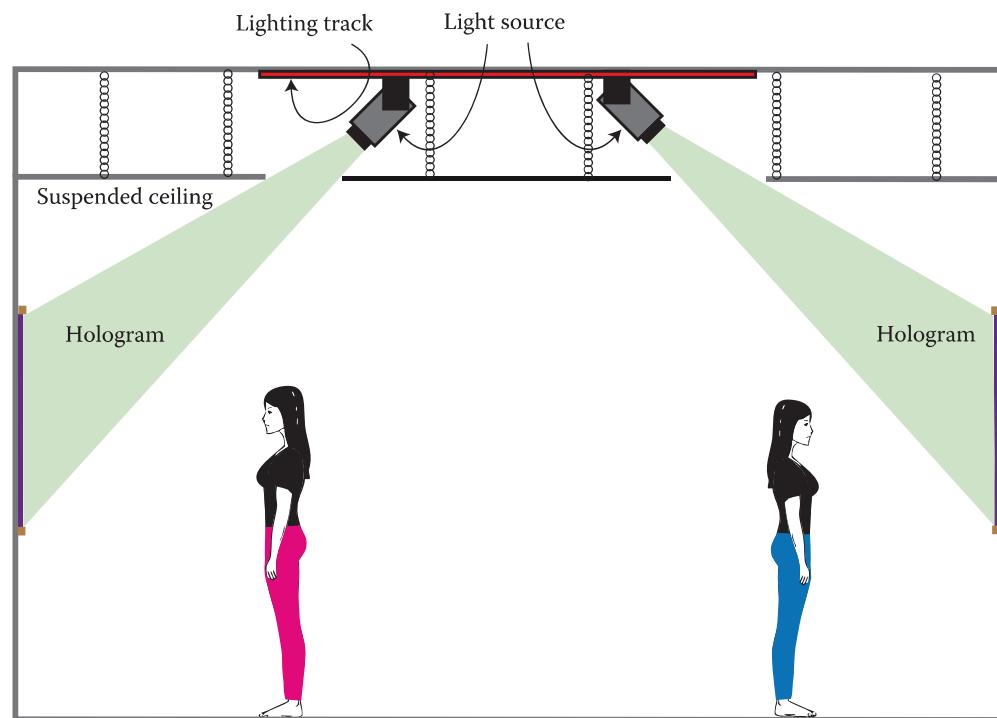


FIGURE 13.27 Layout of a hologram exhibition with hidden spotlights.

constructed with openings for the light to pass through it. This means that the light from the spotlights can pass through openings in the ceiling (a grid system) to illuminate the holograms without the spotlights being visible. The visitor is therefore not disturbed by many bright light sources in the ceiling and, in the best case, is not even aware that spotlights are used to illuminate the holograms on display. This makes a radical improvement to the exhibition (Figure 13.27).

Another way of hiding the spotlights is to arrange the holograms on 45° shelves along the walls with shielded spotlights mounted overhead on the ceiling (Figure 13.28). This type of installation is very suitable for the display of artefact holograms in museums. A further advantage of this arrangement is that both tall and short people (as well as children) can easily see the holograms by moving closer or further away from the hologram. A Swedish museum used this method of displaying an early artefact hologram, which was mentioned in Chapter 1 (Figure 1.31).

Horizontally mounted, overhead-illuminated holograms can be displayed on a table with a spotlight directly above them. This technique is used to display holographic maps, which we will describe in Chapter 14. An example of a horizontally mounted, overhead-illuminated hologram mounted as a tabletop and made by Laser Reflections [9] is shown in Figure 13.29.

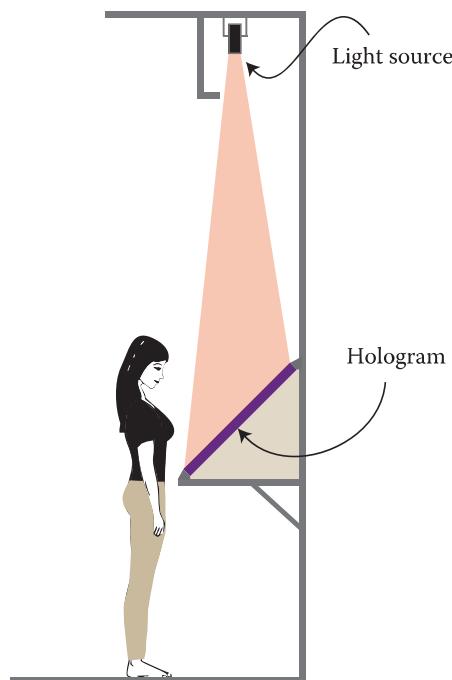


FIGURE 13.28 Hologram on 45° shelf with spotlight mounted overhead.



FIGURE 13.29 Tabletop hologram by Laser Reflections.

When arranging exhibitions of display holograms, it is recommended that spotlights of the type used in theatres be utilised—so-called framing projectors. Such spotlights can project an adjustable frame around the hologram, which means that only the hologram area is illuminated with no illumination outside the hologram frame, eliminating any disturbing light on the wall. Figure 13.30 illustrates typical illumination of holograms using spotlights in which the walls are illuminated around the hologram frames. This is both visually disturbing and also reduces the contrast of the holographic image. Illumination of a large-format hologram using an Electronic Theatre Control Source 4 Junior Zoom is shown in Figure 13.31. As we mention in Appendix 6, LEDs in conjunction with high-index aspheric lenses can be expected to dramatically decrease the size and efficiency of framing projectors.



FIGURE 13.30 Spotlights illuminating the wall surrounding the holograms.



FIGURE 13.31 A spotlight of the framing projector-type is shown illuminating a large-format hologram.

13.5 Edge-Lit Holograms

Considering all the problems associated with the correct illumination of holograms, it would be rather nice if it were possible to integrate the light with the hologram to create an essentially self-luminous display. One way to do this is by introducing the illuminating light into the edge of the hologram plate. Such holograms are referred to as edge-illuminated or edge-lit holograms.

One type of edge-lit hologram is based on total internal reflection (TIR) and was described in 1968 by Nassenstein [18]. In the recording arrangement for this type of hologram, the incidence angle of the reference wave must be greater than the critical angle. In fact, the critical angle is not strictly defined here as there is weak light absorption by the photosensitive medium—this is the case of attenuated TIR. In 1969, Bryngdahl [19] described how TIR holographic recordings could be understood in terms of evanescent waves. Stetson [20] published the first experimental investigations of TIR gratings. Edge-lit TIR gratings and holograms can only be used for the recording of small-size plates—for example, holograms used in fingerprint detectors.

A lot of research has been undertaken to investigate how edge-lit holograms can be successfully recorded. Various solutions have been suggested, but so far, it has only been possible to apply these to displays for small holograms. Lin [21] was the first to introduce the concept of the edge-lit technique to display holography.

Edge-lit holograms not using the TIR principle are illuminated at a very steep angle by a light source positioned nearby. One solution here is to attach the holographic plate to a glass or plastic block. Light is introduced through the edge of the block and directed onto the holographic emulsion. There are two types of such holograms: one is the reflection edge-lit hologram and another is the transmission edge-lit hologram, as illustrated in Figure 13.32. Several variants are possible. For example, in a colour transmission display, the red, green and blue primary reference beams can be brought in from three sides of the block. Reflection holograms can also be treated in this way, although it is more usual to use a single white reference beam.

Holograms somewhat larger than TIR holograms have been displayed successfully in this fashion. However, no medium- or large-format displays have thus far been reported. From an absolute point of view, it should be possible to create large displays using this principle. However, the larger the display, the thicker the glass or plastic block must be. Because the edge of this block must be optically polished, the larger blocks are much more expensive. They are also heavy. This is certainly an area that needs more work. We should, in addition, point out that digital holograms may have some advantages here, as distortions in the reference beam (such as non-collimation and even astigmatism) may be corrected when the hologram is written using a variable reference beam system.

Unlike the standard types of transmission and reflection holograms, edge-lit holograms are not susceptible to interference and image blurring from extraneous light in the viewing environment. This allows such holograms to be displayed in well-lit areas with little regard for other light sources that might otherwise affect the holographic image. However, the main advantage is the fact that the hologram forms an effectively compact self-luminous display with no zeroth order reference beam reaching the observer's side of the hologram. Over many years, research on edge-lit holograms has been undertaken by several groups [18–33]. The ability to create full colour edge-lit holograms has also been considered and this has been described by Ueda et al. [27].

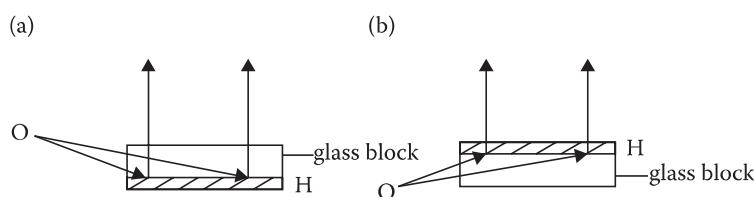


FIGURE 13.32 (a) Reflection and (b) transmission edge-lit holograms.

13.6 Illumination of Large Displays

Large displays pose particular problems for illumination. With digital holographic printing technologies improving all the time, it is now becoming realistically possible to print high-definition HVV displays. These types of displays require illumination sources of a small bandwidth and high spatial coherence if chromatic and source-size blurring are to be avoided. From a practical point of view, this means that HVV displays must be illuminated using laser technology.

Full-colour HVV displays can be made using reflection holograms or transmission holograms. Transmission holograms have the advantage that the zeroth order illumination beam may be suppressed by an additional grating and an in-wall installation arranged. The inconvenience with colour transmission holograms is that angle multiplexing is required and this does represent a significant complication. Reflection-type HVV displays are therefore likely to be the first type to appear commercially. However, unless upscaled edge-lit technologies are used, it is not really possible to suppress the zeroth order illumination beam here. This places a constraint on the maximum size that a single HVV reflection display can have due to laser safety concerns.

Essentially, an observer must be able to stand next to a laser-illuminated hologram and look directly into the laser illumination source without incurring eye damage. Key elements here will be limiting the spatial coherence of the RGB laser source (no doubt using an LSR) to no greater than that required to abolish source-size blurring, maximising the diffractive efficiency of the hologram and, if possible, restricting the field of view of the display. With these measures, quite large panels could be lit using compact RGB laser sources. Nevertheless, above a certain size (which will depend on the above points), too powerful a laser will be required for ocular safety. One particularly good solution is therefore to break-up the display into rectangular panels, each with their own illumination source (Figure 13.33). Of course, there can be very little overspill of light from one panel to the next as multiple images will then result. However, the spatially coherent light from an RGB laser source can be easily focussed into an almost perfect achromatic rectangle on the wall using only simple compact optics.

We are fundamentally used to making windows, mirrors and even pictures out of panels. There is therefore little problem in including thin borders between the holographic panels and this may conveniently cater to some overspill from the individual illuminating light sources if necessary. The one disadvantage of splitting an HVV display into panels is the relatively high precision (~ 0.1 mm per metre distance between light and hologram) that will be required in the alignment of the light sources such that the individual images all align!

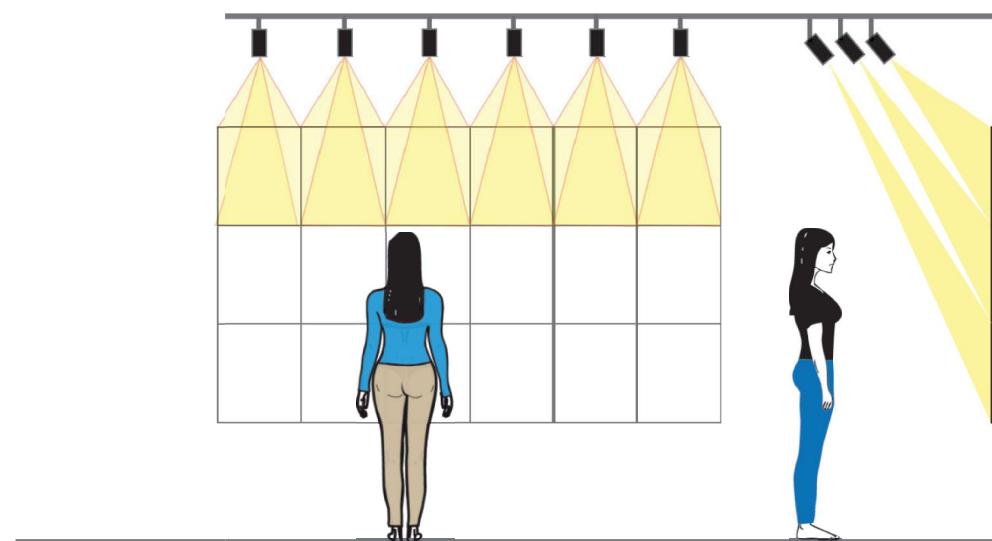


FIGURE 13.33 Illumination of HVV reflection hologram panels using individual RGB laser framing projectors for each panel.

REFERENCES

1. R. J. Collier, C. B. Burckhardt and L. H. Lin, *Optical Holography*, Academic Press, New York (1971) pp. 170–174.
2. P. Janssens and K. Malfait, “Future prospects of high-end laser projectors,” in *Emerging Liquid Crystal Technologies IV*, L.-C. Chien and M. H. Wu eds., Proc. SPIE **7232**, OY-1–12 (2009).
3. Y. M. Lee, J. M. Park, S. Y. Park and S. G. Lee, “A study of the relationships between human perception and the physical phenomenon of speckle,” SID Int. Symp. Digest of Tech. Papers P-45, 1347–1350 (2008).
4. Optotune AG, Switzerland; www.optotune.com (Sept. 2012).
5. A. Neumann, J. J. Wierer, Jr., W. Davis, Y. Ohno, S. R. J. Brueck and J. Y. Tsao, “Four-color laser white illuminant demonstrating high color-rendering quality,” *Opt. Expr.* **19**, A982–990 (2011).
6. International Light Technologies, USA; www.intl-lighttech.com (Sept. 2012).
7. Welch Allyn, USA; www.welchallyn.com
8. Plasma International Lighting Systems, Germany; www.plasma-i.com (Sept. 2012).
9. Laser Reflections, USA; www.laserreflections.com (Sept. 2012).
10. H. A. Jones-Bey, “LEDs illuminate holographic displays,” *Laser Focus World*, February 1, (2003).
11. Aether Systems Inc., Taiwan; www.aether-systems.com (Sept. 2012).
12. Molex Corporation, USA; www.molex.com (Sept. 2012).
13. The Zhaga Consortium, USA; www.zhagastandard.org (Sept. 2012).
14. Philips Lightolier, USA; www.lightolier.com (Sept. 2012).
15. Light Sciences Group, USA; www.lsgc.com (Sept. 2012).
16. AutoTech Wireless Automation, Greece; www.autotech.gr (Sept. 2012).
17. Taurus SecureSolutions Ltd. Greece; www.taurus.com.gr (Sept. 2012).
18. H. Nassenstein, “Holographie und Interferenzversuche mit inhomogenen Oberflächenwellen,” *Phys. Lett. A* **28**, 249–251 (1968).
19. O. Bryngdahl, “Holography with evanescent waves,” *J. Opt. Soc. Am.* **59**, 1645–1650 (1969).
20. K. A. Stetson, “Holography with total internally reflected light,” *Appl. Phys. Lett.* **11**, 225–227 (1967).
21. L. H. Lin, “Edge-illuminated holograms,” *J. Opt. Soc. Am.* **60**, 714A (1970).
22. S. Benton, S. Birner and A. Shirakura, “Edge-Lit Rainbow Holograms,” in *Practical Holography IV*, T. H. Jeong, ed., Proc. SPIE **1212**, 149–157 (1990).
23. N. J. Phillips, C. Wang and T. E. Yeo, “Edge-illuminated holograms, evanescent waves and related optical phenomena,” in *Int'l Symp on Display Holography*, T. H. Jeong, eds., Proc. SPIE **1600**, 18–25 (1991).
24. Q. Huang and H. J. Caulfield, “Edge-lit reflection holograms,” in *Int'l Symp on Display Holography*, edited by Jeong, T. H., Proc. SPIE **1600**, 183–186 (1991).
25. J. Upatnieks, “Edge-illuminated holograms,” *Appl. Opt.* **31**, 1048–1052 (1992).
26. N. J. Phillips, C. Wang and Z. Coleman, “Holograms in the edge-illuminated geometry—new materials developments,” in *Practical Holography VII: Imaging Materials*, S. A. Benton, ed., Proc. SPIE **1914**, 75–81 (1993).
27. H. Ueda, K. Taima, and T. Kubota, “Edge-illuminated color holograms,” in *Holographic Imaging and Materials*, T. H. Jeong, ed., Proc. SPIE **2043**, 278–286 (1993).
28. H. Ueda, E. Shimizu, and T. Kubota, “Image blur of edge-illuminated holograms,” *Opt. Eng.* **37**, 241–246 (1998).
29. S. Sainov and R. Stoycheva-Topalova, “Total internal reflection holographic recording in very thin films,” *J. Opt. A: Pure Appl. Opt.* **2**, 117–120 (2000).
30. Y. J. Wang, M. A. Fiddy and Y. Y. Teng, “Preshaping of reference beam in making edge-illuminated holograms,” in *Diffractive and Holographic Device Technologies and Applications VI*, I. Cindrich and S. H. Lee eds., Proc. SPIE **3291**, 190–198 (1998).
31. W.-C. Su, C.-C. Sun and N. Kukhtarev, “Multiplexed edge-lit holograms,” *Opt. Eng.* **42**, 1871–1872 (2003).
32. M. Metz, “Edge-lit holograms,” in *Holography for the New Millennium*, Chapter III:3, J. Ludman, H. J. Caulfield and J. Riccobono eds., Springer-Verlag (2002) pp. 59–78.
33. W. Farmer, “Edge-lit holography,” in *Holographic Imaging*, Chapter 18, S. A. Benton and V. M. Bove Jr., Wiley-Interscience, John Wiley & Sons Inc. (2008) pp. 193–205.

14

Applications of Ultra-Realistic Holographic Imaging

14.1 Introduction

This book is predominantly targeted at visual imaging applications that have arisen because of new and improved holographic methods. In particular, during the course of this book, we have described how ultra-realistic full-colour three-dimensional (3D) images may now be generated using analogue and digital holographic techniques. In this final chapter, using several examples, we illustrate how such ultra-realistic 3D images can be used today. Along these lines, we discuss how high-quality holographic reproductions of priceless artefacts and items of interest to cultural heritage are gaining interest with museums, how new holographic techniques offer a very effective method of reproducing oil paintings to a level of detail which is impossible using any other technique, and how digital holographic printing has already produced stunning large-format 3D posters that are miles ahead of other technologies, such as lenticular displays. We conclude the chapter with a brief discussion of such future applications as holographic windows and real-time holographic displays.

Before embarking on a presentation of the direct visual imaging applications of ultra-realistic holography, we will review several applications of scientific interest that fall somewhat to the side of our main theme. These are bubble chamber holography, holographic endoscopy and holographic microscopy. All are based on the capacity of holography to generate ultra-realistic structural images.

General applications of holography, conventional holographic optical elements, volume holographic gratings, holographic memory and holographic interferometry are covered in many books [1–6]. Here, we limit ourselves simply to making the comment that, today, holography is being adopted more and more in various “non-imaging” applications. For example, thick low-loss volume holographic gratings are leading to better and more readily available single-frequency laser sources [7–9]. Here, the hologram acts as a frequency-selective mirror. The applications of such single-frequency lasers are extremely wide. Another example is holographic data storage [2–4]. Finally, improved holographic optical elements are finding applications in an increasing number of areas. For example, high-quality gratings are now useful in visual astronomy for eliminating certain parasitic wavelengths [10,11], and very large holographic gratings are being made for chirped pulse amplification systems in the petawatt lasers that are required for inertial confinement nuclear fusion [12].

14.2 Some Scientific Applications of Holographic Imaging

One considerable advantage that holography has over photography is its ability to offer a high-resolution recording over a large volume—in particular, when using very short exposure times. In photography, one needs to know in advance where to focus one’s camera to record a high-resolution image.* Three applications that make use of this principle are bubble chamber holography, holographic endoscopy and holographic microscopy.

* One can of course use deconvolution techniques to deblur a photograph but only to a certain point—holography is intrinsically more stable than deconvolution as the hologram contains far more information than the photograph.

14.2.1 Bubble Chamber Holography

A bubble chamber holography project [13] took place in the mid-1980s at the Fermi National Accelerator Laboratory in Batavia, Illinois.* Nowadays, bubble chambers have largely been replaced by faster and more accurate solid-state electronic detectors. Nevertheless, the same principles used here may be applied to similar detectors. For example, large cloud chambers are currently being used to measure the effect of cosmic rays on precipitation levels in clouds for climate models [14].

For efficient detection of short-lived particles with lifetimes of the order of 10^{-13} s, bubble chambers must be equipped with high-resolution imaging systems. High-energy physics research interest in the 1980s was centred on particles containing heavy-flavour quarks (charm, bottom, etc.). These particles, as well as the heavy lepton (τ), have lifetimes of the order of 10^{-12} to 10^{-13} s, which was much shorter than those previously observed. The experimental difficulty in decay-vertex detection of short-lived particles with mean lifetimes of approximately 10^{-13} s is that their flight paths before decay are of the order of several hundred microns up to a few centimetres. This requires a vertex detector with a spatial resolution better than several microns. The most ideal vertex detector is one that tags short-lifetime particles by precisely measuring the impact parameter of their decay products[†] and efficiently assigning the bubble chamber tracks to their correct parents. Bubble chambers equipped with high-resolution imaging systems were the most prominent “visual” vertex detectors in the 1980s.

As a charged particle traverses the superheated metastable liquid in a bubble chamber, it produces bubbles along its path through ionisation. By improving the resolution in conventional photography, a rapid decrease in the depth of field occurs, thereby resulting in a smaller useful volume photographed. Using holography to supplement conventional photography remedies this problem. Holography separates resolution from depth of field and achieves a high resolution over a large depth of field.

The E-632 Physics Experiment used a dark-field holographic recording technique in the 15 ft. (4.6 m) bubble chamber at Fermilab [13,15]. A high-power pulsed ruby laser was used as the light source. The laser, which we reviewed in Chapter 6 along with the electronics required to lengthen its pulses, produced light pulses of 2 to 100 μ s duration with energies of more than 8 J. The Fermilab bubble chamber with the holographic recording system installed is shown in Figure 14.1.

The laser beam entered the 15 ft. bubble chamber through a small window at its bottom after traversing 30 m in an underground vacuum pipe. The laser beam was expanded using a special dispersing lens. In addition to the holographic recording camera, the bubble chamber was equipped with conventional cameras. All the cameras had concentric hemispherical lenses called *fisheye* windows. The centre of the expanding lens was on an axis passing through the centre of the holographic camera; the film was placed only 10 cm from the inner surface of the fisheye window. The laser pulses used were mostly of 40 μ s duration with approximately 600 mJ energy in a 2.5 cm-diameter beam entering the bubble chamber. The beam had a flat light intensity distribution before hitting the lens. The expanding lens distributed the laser light inside the chamber with increased intensity at large angles to compensate for the reduced light diffraction off bubbles illuminated at large angles. Rolls of 70 mm Agfa-Gevaert Holotest 10E75 film were used for recording the holograms. During the experiment, approximately 100,000 holograms were recorded. Conventional bubble chamber photographs were recorded 10 ms after each hologram. An example of what could be achieved with the holography system is shown in Figure 14.2. In the upper part of the figure is a conventional photograph of a neutrino event taken 12 ms after the particle beam entered the bubble chamber (the diameter of the bubbles is <1 mm). The lower part of the figure depicts the same event photographed from its reconstructed holographic image showing much smaller bubbles (190 μ m diameter) as the hologram was recorded only 2 ms after the neutrino beam injection. Whereas the large bubbles in the conventional photographs obstruct the details of a close-in interaction at a distance of approximately 2 cm from the primary vertex, the hologram clearly allows a detailed study of the secondary interaction.

* One of the authors (HB) was involved in this project.

[†] This is the distance by which the trajectory of a charged particle coming from a secondary misses the production vertex.

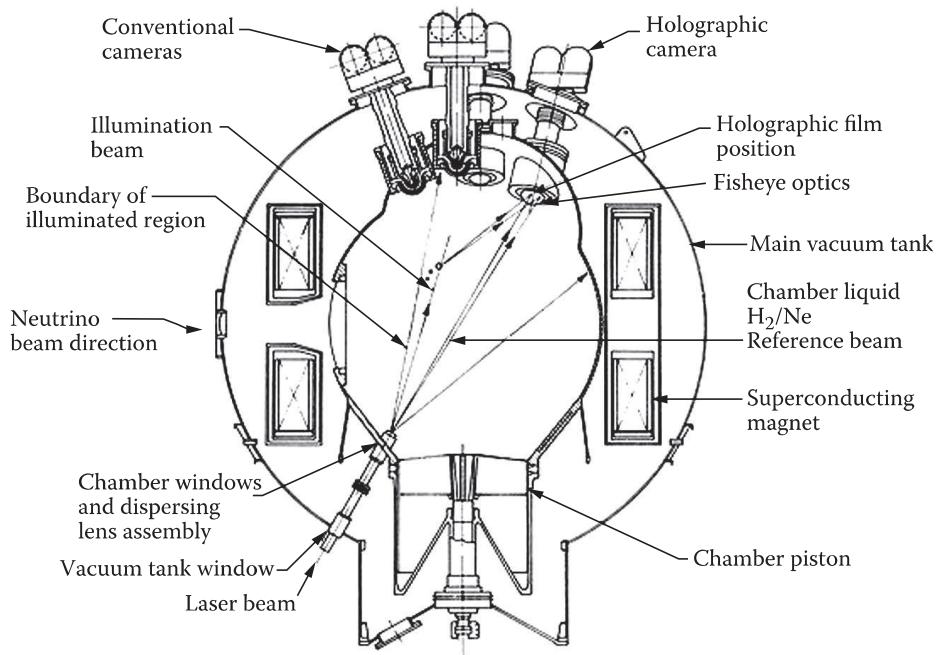


FIGURE 14.1 The 15 ft. Fermilab bubble chamber with holographic components shown (side view).

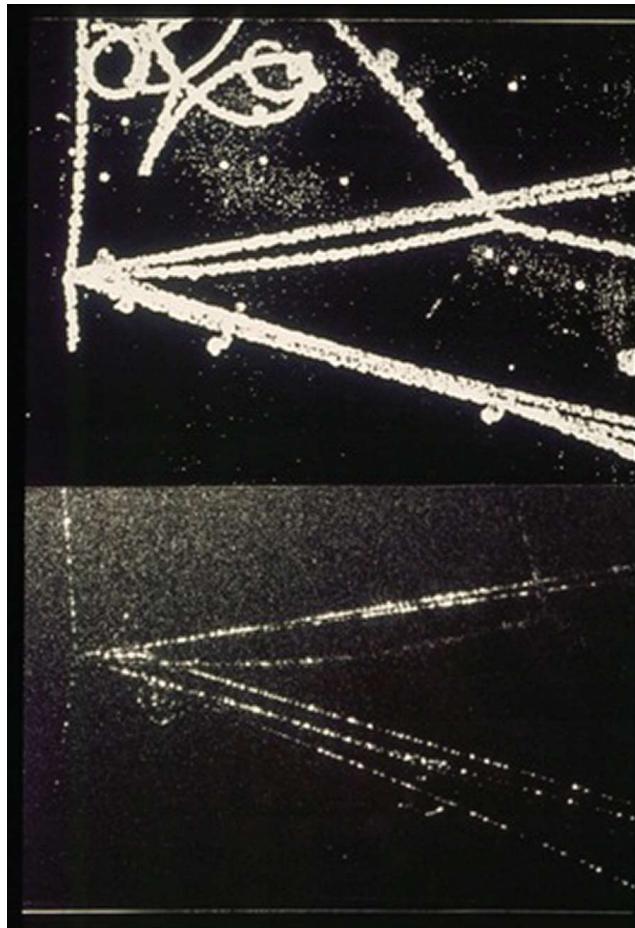


FIGURE 14.2 Comparison of the vertex region of a neutrino event imaged with a conventional photograph (top picture) and with the hologram of the same event, having a resolution of 150 μm (bottom picture).

14.2.2 Holographic Microscopy

Holography can be used to provide high image resolution over large microscopic volumes. There are many examples of early holographic microscopy applications that used holographic photosensitive plates or film. Today, it is possible to record holograms on electronic detector arrays such as charged coupled devices (CCDs) and to replay them numerically using fast Fourier transform (FFT) algorithms.* Digital holographic microscopy (DHM) has many important applications that were either not possible or not convenient with analogue recording. Digital holographic recording (DHR) eliminates the need for wet chemical processing and other time-consuming procedures, so recording and numerical reconstruction can be done in almost real time. The basic principles of DHR, holographic microscopy and state-of-the-art applications have been reported by Sang et al. [16].

An example of the application of DHM to real-time process control has been reported by Osanlou [17]. Here, holographic microscopy is used to monitor the production of various emulsions in specialised production vessels. This study is particularly relevant to online production monitoring and control in continuous manufacturing processes. The technique reported is capable of quantitative 3D mapping of moving fluids in a snapshot. It is non-invasive, high-resolution and precise. Important features of the methodology include quality of image capture and reconstruction without the use of a pulsed laser. Examples of reconstructed image quality are shown in Figure 14.3.

An ambitious and difficult application of DHR for recording marine particles has been studied by Watson [18]. Watson has constructed several holographic cameras for underwater imaging of plankton and other marine particles. These devices use both analogue and digital holography to record water volumes of up to a cubic metre. For in-water deployment, however, the weight and size of the early instruments, which were based on analogue holography, restricted their use on advanced observation platforms such as remotely operated vehicles and limited operational depth to a few hundred metres. Recent advances made in DHR on electronic sensors coupled with fast numerical reconstruction has led to the development of smaller underwater holographic recording equipment (Watson's eHoloCam). This has freed holography from many of its constraints, allowing rapid capture and storage of images including holographic video of moving objects. Digital holography's ability to record true 3D, full-field, high-resolution, distortion-free images *in situ* from which particle dimensions, distribution and dynamics can be extracted is hard to match. Figure 14.4 shows some examples of images of marine creatures calculated by the numerical reconstruction of the recorded interference patterns at 100 m depth.

DHM is very suitable for use in the fields of biological and medical research. Here, the in-line recording technique is preferentially used (see Xu et al. [19] for the properties of such digitally recorded holograms). For many years, DHM has been applied by the University of Münster, Germany, for studying living cells. An off-axis configuration has been introduced by Carl et al. [20] for studying tumourous human hepatocytes *in vivo*. It should be mentioned that DHM has significant advantages over standard phase-contrast microscopy in the very high resolution it can offer.

Another recent biological application of DHM was presented by Jourdain et al. [21], in which neuronal activity was observed in real time and in three dimensions, with a resolution of up to 50 times better than previously available. The new technique accurately visualises the electrical activity of hundreds of neurons simultaneously, at up to 500 images per second without damage by electrodes. In comparison other techniques are only able to record activity from a few neurons at a time.

Traditional microscopes are limited to a resolution of approximately 500 nm. However, DHM allows a resolution down to as low as 10 nm in some cases. DHM constitutes a fundamentally novel application for studying biological objects such as neurons, and the technique has many advantages over traditional microscopes. It is non-invasive, allowing extended observation of neural processes without the need for electrodes or dyes that inevitably damage cells. By inducing an electric charge in a culture of neurons using glutamate (the main neurotransmitter in the brain), charge transfer carries water inside the neurons and changes their optical properties in a way that can be detected only by DHM. A computer is then able

* This process is often referred to as *digital holography* and is closely related to computer-generated holography. Note that digital holographic printing techniques such as DWDH and MWDH, as described in Chapter 7, are often also referred to as digital holography. However, the two fields of digital holography are rather different.

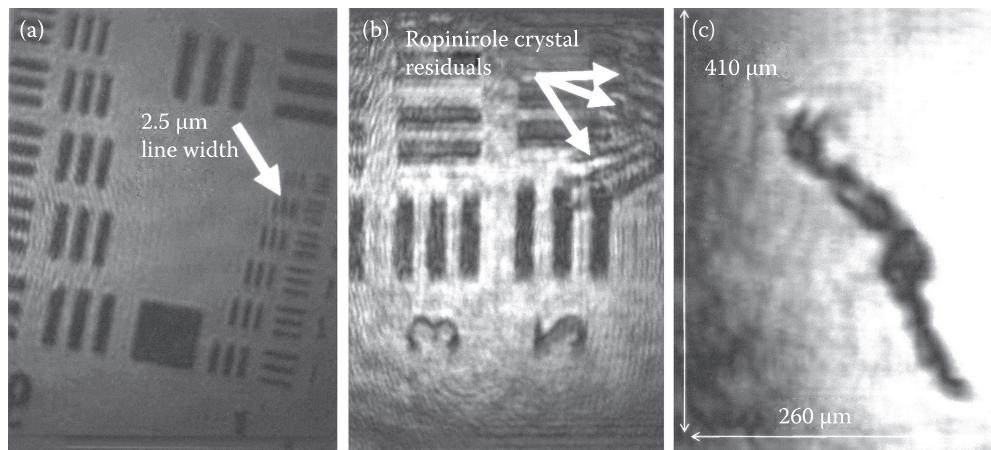


FIGURE 14.3 Reconstructed images (real-time holographic process control) showing (a) US Air Force test target (b) ropinirole crystal residuals (width, 6.2 μm) in front of the US Air Force test target and (c) benzoic crystal chains floating in the gently stirred fluid within the reactor vessel. (Photos courtesy of A. Osanlou.)

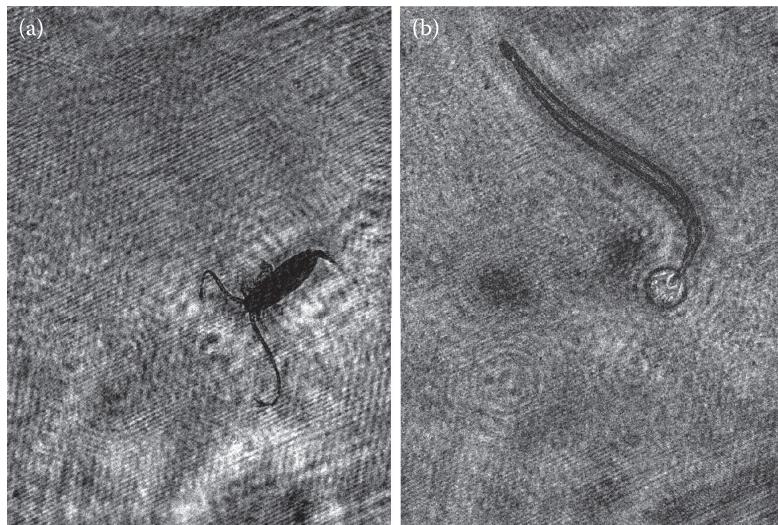


FIGURE 14.4 Reconstructed DHR images of (a) a calenoid copepod (3 mm body length) and (b) an arrow worm (1.75 mm length). (Photos courtesy of J. Watson.)

to numerically reconstruct a real-time 3D image of the neurons. DHM is the first imaging technique that has been able to monitor the activity of these cotransporters dynamically and *in situ* during physiological or pathological neuronal conditions.

There are many other applications of digitally recorded holograms and we can expect the field to expand considerably in the coming years. Today, the possibility to record digital holograms is limited to small angles between the object and reference beams. This is because current electronic recording devices only possess a certain finite resolution. This is why today's applications are limited primarily to microscopic recording techniques in which in-line holograms can provide the required image resolution.

14.2.3 Holographic Endoscopy

One of the authors (HB) has been involved in the 2006 *Holoendoscope* project [22] to develop high-resolution colour holographic endoscopy for cellular analysis. Endoscopic holography or *endoholography* combines the features of endoscopy and holography. The purpose of endoholographic imaging is to provide the physician with a unique means of extending diagnosis by providing a lifelike record of tissue. Endoholographic recording provides a means for microscopic examination of tissue and, in some

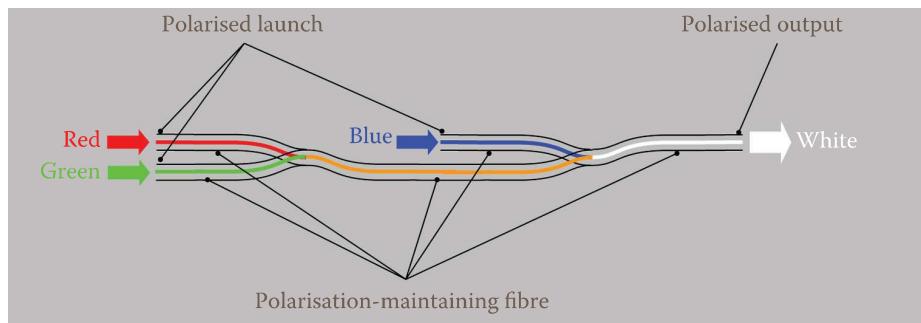


FIGURE 14.5 Fibre-coupler used to generate the “white” laser light for the endoscope. (Reproduced with permission from Gooch & Housego.)

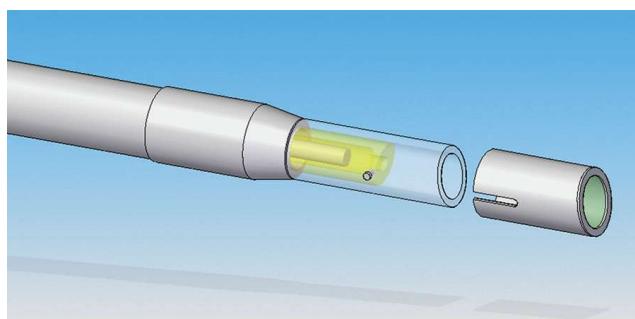


FIGURE 14.6 Prototype holoendoscope. The holographic film is inserted at the tip of the endoscope. (Reproduced with permission from Gooch & Housego.)

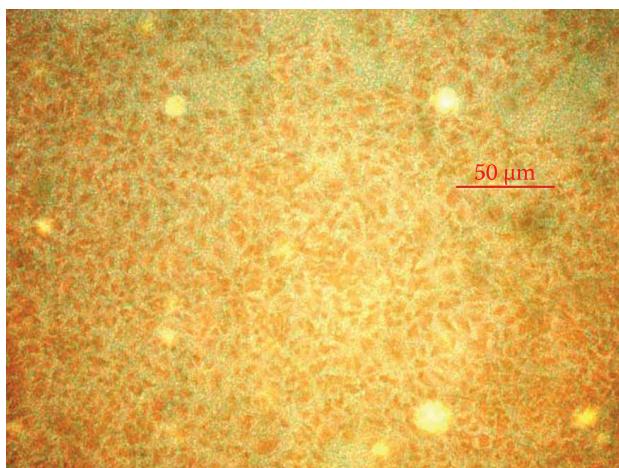


FIGURE 14.7 *In vitro* endoscopic colour hologram of human endothelial cells.

cases, may obviate the need to excise specimens for biopsy. In the project, *in vitro* colour holograms were recorded to demonstrate the feasibility of recording high-resolution tissue images. The prototype endoscope* used small RGB lasers with the laser light being combined using single-mode optical fibres to provide illumination through the endoscope. Denisyuk single-beam colour reflection holograms were recorded in close contact with the tissue at the distal end of the endoscope [23]. The recorded holograms were viewed and photographically documented under a microscope. Earlier, in 1988, a demonstration of a monochrome hologram recording using an earlier prototype holoendoscope (holographic

* The holoendoscope project partner Sifam-Gooch & Housego was responsible for the RGB lasers, the fibre-optic coupler, and the prototype holoendoscope.

sigmoidoscope) was performed to obtain *in vivo* holograms of the colon of an anaesthetised dog [24]. The principle of the fibre coupler colour laser combiner is shown in Figure 14.5, the prototype holoendoscope is illustrated in Figure 14.6, and a microscopic photograph derived from an *in vitro* endoscopic colour hologram of human endothelial cells is shown in Figure 14.7.

14.3 Visual Applications of Full-Colour Holographic Imaging

In this section, we focus on current applications of the technologies of full-colour analogue hologram recording (as described in Chapter 5) and digital holographic printing (Chapters 7–10). Despite holography now being more than 60 years old, many of these applications are in an early stage of market development. This is really because the technology improvements that we have presented in this book have, in many cases, only recently led to a positive reassessment of old ideas.

14.3.1 Holographic Copies of Museum Artefacts

In the days of the British Empire, museums filled their collections with exotic items from around the world. British museums today still possess large collections from foreign countries, but there is increasing pressure to repatriate priceless artefacts to their respective homes. Holography now offers the possibility to essentially duplicate such artefacts—and to a point where observers practically cannot tell whether they are looking at the real exhibit or at a holographic copy. Although such holographic reproduction can never match the value of actually possessing the real artefact, it can allow the museum to fulfil one of its most important functions—to maintain display of the exhibits. Of course, the real exhibit potentially allows future scientific tests to be performed on an artefact to verify scientific theories—such as CAT scans, material testing and non-visual spectral analysis—but analogue holography does offer a means to preserving a faithful visual recording of unprecedented microscopic detail.

We should mention that digital holographic printing can also be usefully applied to museum recordings. By using a two-dimensional tracking camera system,* high-resolution digital image data from a museum exhibit can be recorded from over a million different angles. This data may then be written onto a very high-resolution digital hologram producing a digital holographic copy. Although the resolution of this type of hologram is less than that of an analogue hologram, at the smallest hogel sizes now available ($\sim 250 \mu\text{m}$), it can nevertheless be very difficult for an observer to tell the difference between an analogue and a digital hologram with the unaided eye. The digital hologram also offers several sizeable advantages. For instance, new holograms can be generated from the digital data whenever required. This means that as long as the original data is stored securely, there are no image lifetime problems. In addition, the same image data can be used to produce holograms of small and very large sizes. Thus, museum exhibits can be displayed in any format. Finally, such digital images are not constrained to be behind a glass plate.

Holography can also help museums with travelling exhibitions, as we shall see in later sections. It is difficult for some people to travel to museums, and, as a result, there is pressure on museums to take exhibits to the people. Transporting priceless artefacts, however, is both hazardous and expensive. Transporting holographic copies on the other hand is not.

A final reason why ultra-realistic full-colour holograms are useful to museums is related to insurance costs of exhibiting within a museum. Most museums have large collections “downstairs” which they do not exhibit. The reason is that it generally costs more to exhibit something than to securely store it—as the risk of damage or loss is greater when an exhibit is on display. Again, holography can help solve this problem. If the hologram is indistinguishable from the real exhibit, why not just securely store the real item and display the hologram?

14.3.1.1 Virtual Museum Exhibitions

One interesting colour holography project,[†] which was recently carried out by the Centre for Modern Optics in North Wales, was a project funded by the Esmee Fairbairn Foundation entitled *Bringing the*

* See Chapter 10 for a discussion of holocam systems.

[†] One of the authors (HB) was involved in this project.

Artefacts Back to the People. The project involved collaborations with a number of major museums including the National Museum of Wales, the British Museum, the Maritime Museum in Liverpool, as well as the Royal Commission for Ancient and Historical Monuments in Wales. Full-colour holograms of various artefacts were recorded using the analogue techniques described in Chapter 5. The holograms were completed by the end of 2009, after which they were displayed as a travelling exhibition that toured North Wales and its borders. The exhibition first opened at Llangollen Museum in June 2010 and later at the museums of Grosvenor (Chester), Wrexham, Llandudno, Bangor and many others [25].

One of the recorded artefacts, supplied by the British Museum, was a 14,000-year-old decorated horse jaw bone from the Ice Age, or late glacial period of Britain [26]. The recording setup is shown in Figure 14.8 and the hologram in Figure 14.9. Another hologram recorded was the *Tudor Owl Jug* and *Sergeant at Arms Ring* shown in Figure 14.10. These artefacts were from the Grosvenor Museum in Chester, United Kingdom. In total, ten full-colour holograms of different artefacts were included in the touring museum exhibition (Figure 14.11).

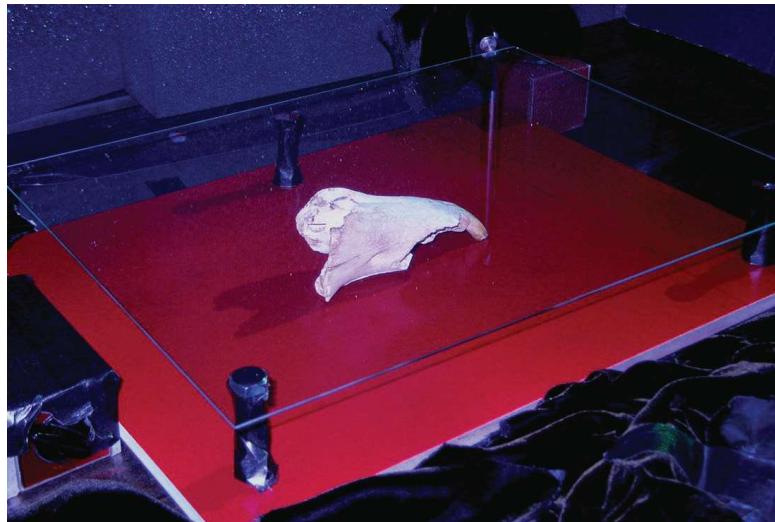


FIGURE 14.8 Horse jaw recording setup.



FIGURE 14.9 Decorated horse jaw hologram.



FIGURE 14.10 Hologram of the *Tudor Owl Jug* and *Sergeant at Arms Ring*.

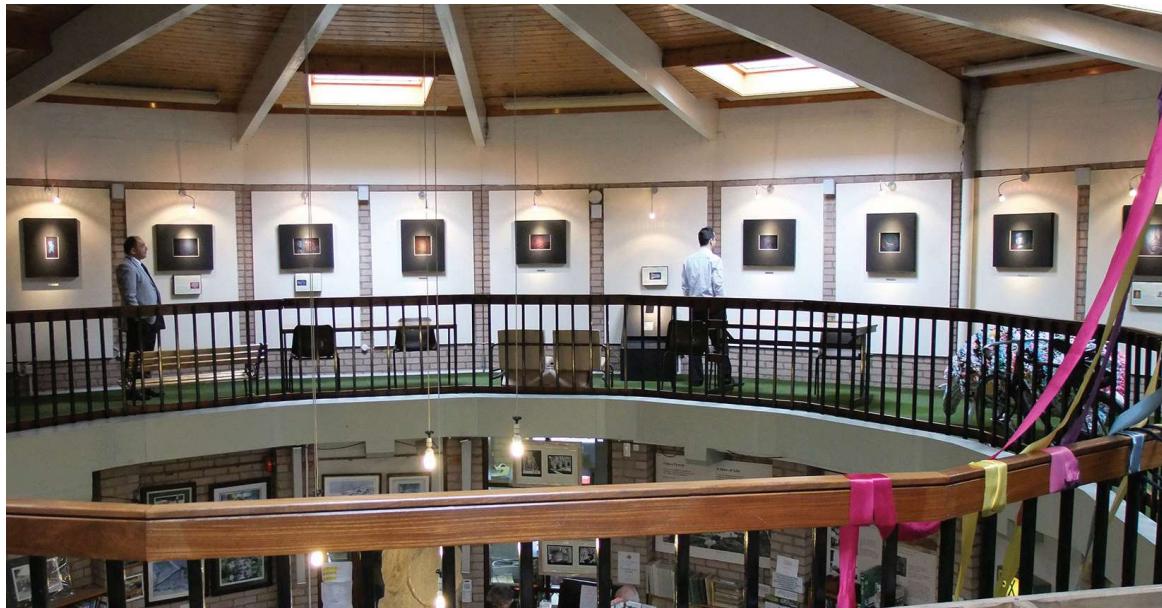


FIGURE 14.11 Exhibition of the artefact colour holograms at the Llangollen Museum in Wales, 2010.

14.3.1.2 Museum Holography in Greece

In Greece, the Hellenic Institute of Holography has recently invested in a full-colour recording facility and is currently promoting *Realistic Colour 3D Holography*, according to its director Alkis Lembessis [27]. The primary goal of the Hellenic Institute of Holography is to record Greek cultural artefacts through the “*HoloCultura: Applied Holography in Cultural Heritage*” project. The project consists of three parts:

- Phase A: Study necessary for implementation of the colour holography programme
- Phase B: Recording of experimental colour holograms
- Phase C: Pilot project involving the recording of cultural artefacts

The institute is active in a country with a unique cultural tradition of worldwide influence extending from classical ancient Greece to orthodox Byzantium Christianity. The use of display holography in the

preservation, recording and public visual dissemination of artefacts from this cultural heritage is at the core of the activities of the Hellenic Institute of Holography.

In 2011, the institute built a small, portable three-colour analogue holographic camera, the Z3 RGB Holography YSB1 prototype camera, which it is now using to record holograms of museum artefacts (Figure 14.12). The camera is a computer-controlled optomechanical device capable of exposing selected, commercially available or experimental, panchromatic silver halide emulsions to combined red, green and blue CW laser beams at appropriate energy levels. The device consists of a main camera unit (MCU) and a control electronics unit (CEU). The MCU is built on top of a lightweight aluminium honeycomb optical board to ensure portability, minimum spatial deformation under nominal temperature variation and fast damping of induced vibrations.

There are three lasers housed in the MCU with wavelengths selected to cover a broad triangle of hues in the Commission Internationale de l'Eclairage (CIE) chromaticity diagram.

- Red laser: 638 nm at an output power of 80 mW (*CrystaLaser* laser)
- Green laser: 532 nm at an output power of 100 mW (*Cobolt Samba* laser)
- Blue laser: 457 nm at an output power of 50 mW (*Cobolt Twist* laser)

The lasers produce TEM₀₀ emissions with coherence lengths of more than 5 m each, and the MCU contains suitable optics to generate a clean collinear mixed RGB beam.

The CEU houses all power supplies, A/D and D/A subsystems plus a specially designed F/P scanning interferometer for beam monitoring. The CEU also connects to an external PC running custom software to control all aspects of the holographic exposure plus full monitoring and tuning of each laser's stability and beam quality.

The Z3 has been successfully tested under wide ambient temperature and humidity ranges. The system becomes stable after 30 min to 1 h, depending on the ambient temperature. At 24°C, the system stabilises after 20 to 30 min. The Z3 camera is accompanied by auxiliary equipment for beam orientation and a flexible vibration-absorbent setup for the positioning of the object. One example of a hologram recorded with the camera is shown in Figure 14.13. The Hellenic Institute of Holography has also produced a portable darkroom for on-site processing (Figure 14.14).

In addition to recording their own holograms with the Z3, the Hellenic Institute of Holography has produced museum holograms in collaboration with Yves Gentet (Figure 14.15) and the Colour Holographic Company in London.

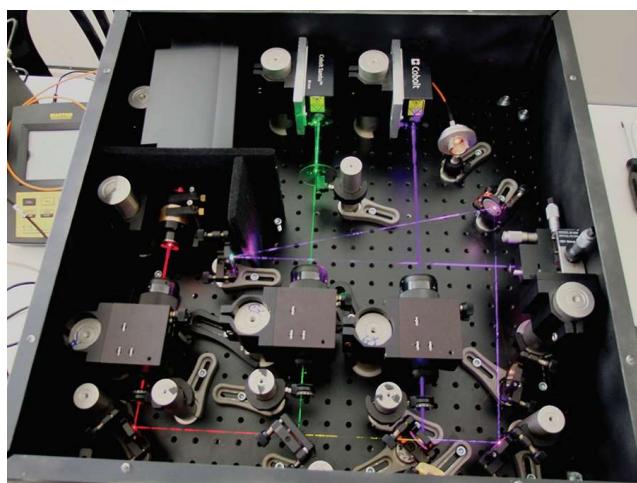


FIGURE 14.12 Z3 prototype portable RGB camera (model YSB1) made by the Hellenic Institute of Holography.

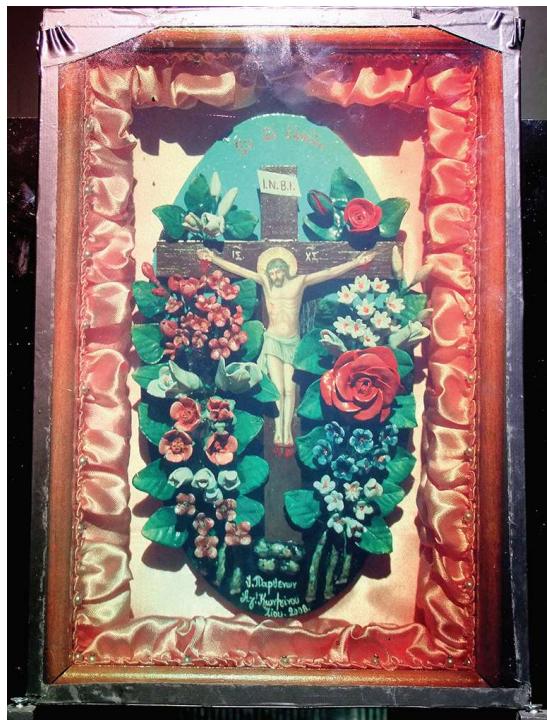


FIGURE 14.13 Full-colour Denisyuk reflection hologram made using the Z3 prototype portable RGB holography camera produced by the Hellenic Institute of Holography.



FIGURE 14.14 Andreas Sarakinos and Alkis Lembessis in front of the portable darkroom produced by the Hellenic Institute of Holography.



FIGURE 14.15 Full-colour Denisyuk reflection hologram of a Greek ceramic vase made by Yves Gentet. This analogue hologram was made in collaboration with the Hellenic Institute of Holography. The right-hand photograph shows an enlarged detail.

14.3.1.3 Museum Holography in England

In 2010, the Colour Holographic Company in London started to record colour holograms on their own ultrafine-grain panchromatic material. Some examples of the full-colour museum holograms that they have produced on this material are shown in Figures 14.16 and 14.17. A new product that the company has introduced is a nice wooden box for the display of colour holograms with an integrated LED light source built into the lid. When the box is opened, correct illumination of the hologram is provided. It is

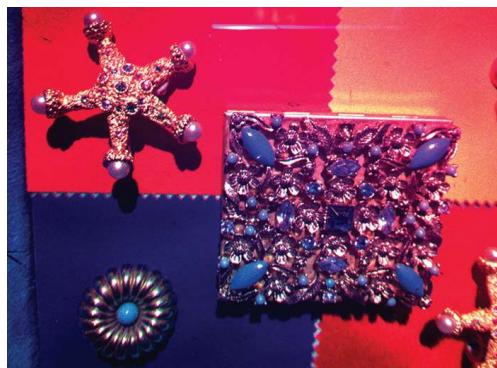


FIGURE 14.16 Full-colour Denisyuk reflection test hologram made by Colour Holographic Ltd.

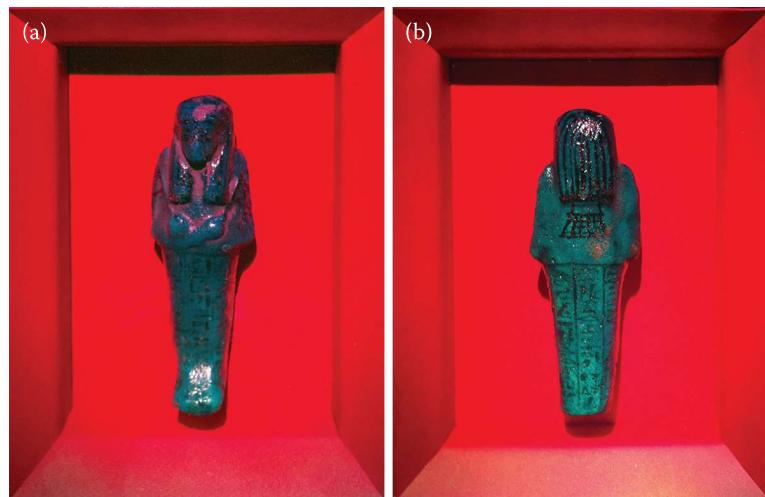


FIGURE 14.17 Full-colour Denisyuk reflection hologram made by Colour Holographic Ltd. The hologram shows (a) the front and (b) back of an Ushabti figure from the Theban cache of royal mummies found in 1881.



FIGURE 14.18 (a) Colour Holographic hologram box and (b) colour hologram.

possible to switch holograms and the company offers the box with different holograms for sale. The box and a colour hologram are shown in Figure 14.18.

14.3.1.4 Holographic Reproduction of Oil Paintings

Full-colour holographic copies of oil paintings offer another interesting application. Because the depth of the recorded image is essentially dictated by the thickness of the brushstrokes, this type of quasi two-dimensional recording can be rather easier to illuminate than conventional holograms. For example, source-size blurring is much less of an issue here. Holographic copying provides a method of producing copies of valuable or priceless paintings, which is unlike any other technique. An oil painting does not look the same from every angle and a photographic reproduction only records the view from straight ahead. A holographic reproduction, on the other hand, faithfully records how the light reflects at all angles, as well as accurately reproducing the relief of the brushstrokes.

One of the authors (HB), working with Dalibor Vukičević, introduced this potential application in 2000 [28]. An example of a *still life* oil painting (20 cm × 25 cm), which was copied using this technique, is shown in Figure 14.19a. The painting was selected mainly because it was painted on



FIGURE 14.19 (a) Oil painting and (b) full-colour analogue Denisyuk reflection holographic reproduction.

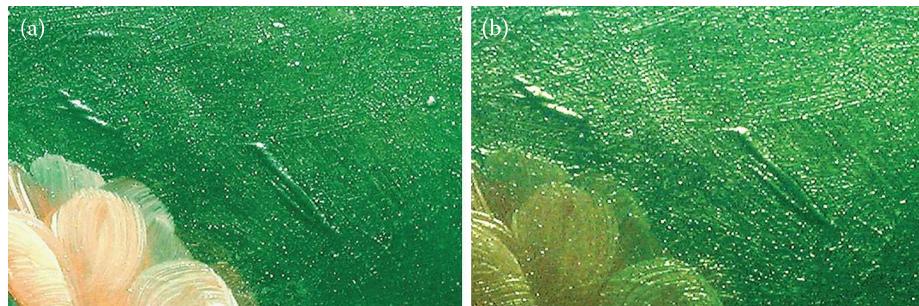


FIGURE 14.20 (a) Detail of brush strokes in the oil painting of Figure 14.19a and (b) in the holographic reproduction of Figure 14.19b.



FIGURE 14.21 Higher magnification of brush strokes in the hologram of Figure 14.19b.

wood* with a pronounced surface texture. A Denisyuk single-beam three-colour holographic setup (described in Chapter 5) was used to record the colour hologram onto a Slavich panchromatic PFG-03C silver halide plate (Figure 14.19b). A detail of the painting texture, visible when observed at a certain angle, is shown in Figure 14.20a. The same surface texture visible in the hologram is shown in Figure 14.20b. The holographic reproduction can be studied in more or less the same way as the real painting can be investigated. Brush strokes, visible in the hologram under high magnification, are illustrated in Figure 14.21.

Colour holograms of paintings have also been demonstrated by Yves Gentet, who recently took recording equipment to the Louvre in Paris to demonstrate the potential of this new reproduction technique. A transportable holographic camera is often required in applications regarding the recording of items of significant cultural heritage, including oil paintings—often because the item in question simply cannot be moved due to insurance or security reasons. Gentet and Shevtsov [29] were the first to develop a small, mobile full-colour analogue holographic camera system especially for this purpose. Much like the system currently used by the Hellenic Institute of Holography, Gentet and Shevtsov's system is based on three solid-state continuous wave lasers: a semiconductor red laser at 639 nm giving 25 mW, a diode-pumped solid-state laser at 532 nm giving 120 mW and another diode-pumped solid-state (Cobolt) laser at 473 nm giving 70 mW. The overall dimensions of the camera system, which allows a hologram format of up to 30 cm × 40 cm to be recorded, are 30 × 40 × 50 cm with a weight of 12 kg.

In copying oil paintings and indeed other cultural heritage items, the reproduction of spectral information is of particular importance. As we mentioned in Chapter 5, three-colour analogue holograms produce a good, but certainly not perfect, spectral representation. Future work will hopefully extend holographic copying to four or five wavelengths.

* With the available lasers at the time, wood provided better interferometric stability.

14.3.2 Digital Display Holograms for Advertising and Product Promotion

14.3.2.1 Introduction

Large-format full-colour 3D reflective digital holographic displays can be extremely impressive. Over the years, this application has had a number of false starts for a variety of reasons—mostly connected with colour issues, reliability, speed of fabrication, illumination and price point. From time to time, great interest has been shown by advertising agencies, printing companies and by organisations wishing to promote their products. A number of studies by reputable market research companies in the last 10 years have estimated potential yearly returns for this sector approaching the billion dollar level.

Many of the issues that have frustrated the penetration of holographic displays in the advertising display sector are now being resolved. Advances in laser technology are making the process of writing the holograms easier. Better materials, such as a new photopolymer material from Bayer, allow higher quality and cheaper images to be produced. Advances in illumination technology make the final hologram brighter, deeper and easier to light. As such, the recent proliferation of 3D films and the increasing popularity of 3D televisions can reasonably be expected to drive a reappraisal and renaissance of this application.

14.3.2.2 Key Organisations

The main companies currently active in the area of digital holographic printing are Zebra Imaging Inc., Austin, TX, and Geola Digital UAB, Vilnius, Lithuania. Zebra was the first group to produce digital full-colour holograms using continuous wave lasers, hogel by hogel. Geola was the first group to do this using RGB-pulsed lasers. Until recently, XYZ Imaging Inc., Montreal, Canada originally a spin-off from Geola and subsequently operating under the name of RabbitHoles Media Inc.,* also operated a digital holographic printing facility in Canada. Ultimate Holography in France, the Dutch Holographic Laboratory in Holland as well as Ceres Imaging Ltd., Spatial Imaging Ltd. and View Holographics Ltd. in the United Kingdom have also developed equipment for the digital printing of colour reflection holograms.

14.3.2.3 Zebra Imaging

Zebra was formed in 1996 by Michael Klug and fellow graduate students from the Massachusetts Institute of Technology's (MIT) Media Lab [30]. In 1999, Zebra produced the world's largest single-colour hologram, which is shown in Figure 14.22. It was developed in collaboration with the Ford Motor Company and was exhibited at the 1999 North American International Auto Show in Detroit. The hologram showed Ford's concept vehicle, the *P2000 Prodigy*, at a 50% scale. The 3D image bisected the 40 ft.² panel, which was composed of individually printed smaller tiles. The exterior of the *P2000 Prodigy* hologram was translucent red, allowing the viewer to see inside the vehicle's futuristic hydrogen fuel cell power train. The hologram contained 900,000 individual exposures taken directly from computer design data supplied by the Ford development team. Ford's Advanced Design Studio was, at the time, working on "replacement reality" techniques—and full parallax large computer-generated colour holograms fit very well into this description. Four terabytes of data storage was needed for the *Prodigy* hologram.

In December 2001, three large Zebra holograms—*Austin Dimensions; Music, Nature and Technology*—were installed at the Austin-Bergstrom International Airport in Texas. The holograms were presented to the airport by Zebra Imaging, Samsung Austin Semiconductor and Frog Design. Measuring 30 ft. long and 4 ft. high, the three panels comprised the world's largest holographic display at the time of its completion. A photograph of the holograms is shown in Figure 14.23.

14.3.2.4 XYZ, Geola and Sfera-S

In 2001, Geola teamed up with Québécois entrepreneur, Eric Bosco, to launch the company, XYZ Imaging (Imagerie XYZ Inc.) in Montreal. This was the first company to target commercial mass

* At the time of going to press, we understand that XYZ Imaging Inc. were in the process of restarting their activities.

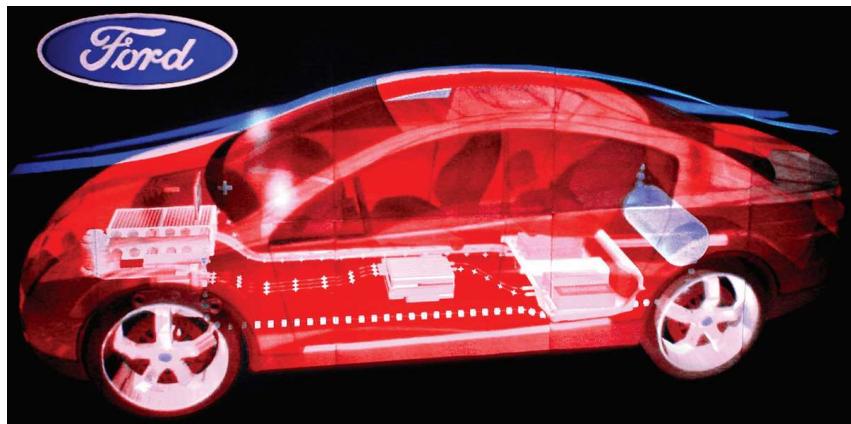


FIGURE 14.22 The 40 ft² full-colour digital reflection hologram of Ford's *P2000 Prodigy* produced by Zebra Imaging in 1999.

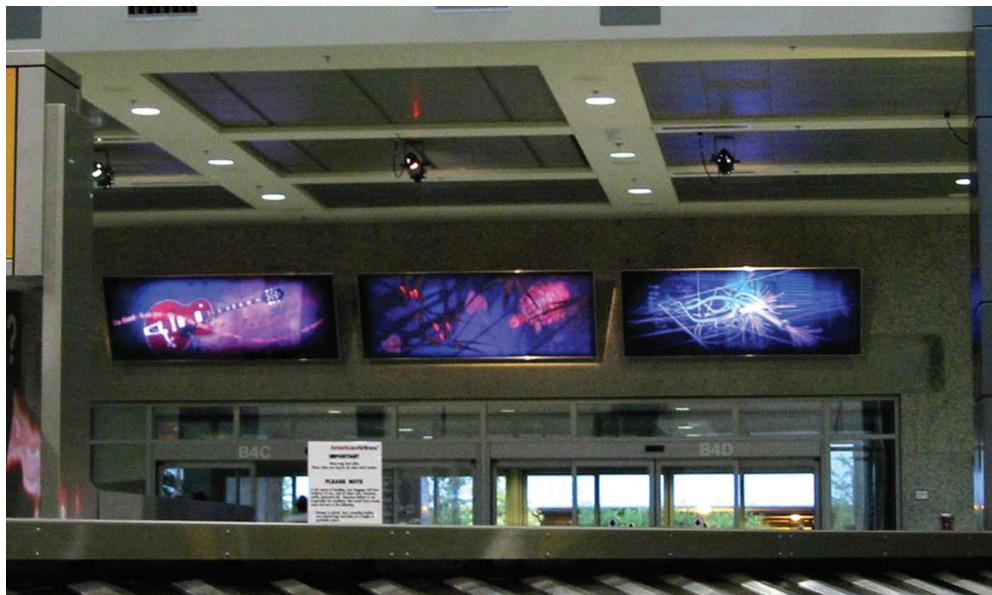


FIGURE 14.23 Austin Dimensions Zebra airport holograms (2001).

production of full-colour 3D holographic printers rather than simply supplying holograms to clients. At the time, the Lithuanian operation of Geola numbered around 30 people. A new company, comprising around 10 people, was created in the United Kingdom to manufacture the RGB lasers for the printers. This was Geola Technologies Ltd., which was based on-campus at the University of Sussex. XYZ itself hired an additional 25 people. The Russian company Sfera-S AO, formed by ex-Micron director at Slavich AO, Yuri Sazonov, was also a critical partner to the XYZ–Geola group. Sazonov worked with Geola and XYZ to produce a vital silver halide film compatible with RGB-pulsed lasers. Zebra used DuPont's panchromatic photopolymer material due to its special contacts, but this material was not available on the general market, and it also had uncertain compatibility with Geola's lasers.

From 2001 to 2006, XYZ, Geola and Sfera-S worked together as a team of around 60 people towards the goal of establishing digital holographic printing as a viable industry for advertising display applications. The initial plan was to develop a true commercial digital holographic printer based on an RGB pulsed laser. Initially, an in-house service bureau in Montreal was to be set up to test the printer under real operational conditions and also to identify any problems with the holograms. However, investor concern over the project timescale led to the service bureau being farmed out to a third-party printing company, and sales of the printers to end-users being brought forward. Despite this, the printer, original

equipment manufacturer (OEM) laser and film development were rather successful, and by 2003 to 2004, excellent quality holograms on a new commercial film were being produced regularly on the new printers. XYZ started to win prestigious awards within the advertising industry for its holograms and, under pressure for sales revenue and because of the optimism created by the awards, commenced commercial sales of its printers. Unfortunately, in hindsight, the company rushed into this process. In addition, its first sales were international and required numerous international service visits. The lead investors, seeing these problems, put the company up for sale, and despite offers from its founding shareholders, new management was brought in and took the decision to sell to a third party. At this point, the synergy of the greater group was broken and the new company, which was again later sold, became an independent 3D printing company known under the name RabbitHoles Media.



FIGURE 14.24 President and Founder of XYZ Imaging, Eric Bosco, in front of one of his company's first digital full-colour holograms. Note that the display has been tiled from smaller panels.



FIGURE 14.25 Digital display hologram made by XYZ Imaging advertising Puma shoes in downtown Montreal (2005).



FIGURE 14.26 Another Puma digital display hologram made by XYZ (2005).



FIGURE 14.27 Digital display hologram made by XYZ, mounted on a custom display unit manufactured by a UK Geola subsidiary, Power Imaging Ltd. (2006).



FIGURE 14.28 Avatar digital hologram (16 in. × 16 in.) by RabbitHoles Media (2011). The company produced limited edition series in three formats: 32 in. × 32 in. holograms (edition of 10), 16 in. × 16 in. holograms (edition of 50) and 8 in. × 8 in. holograms (edition of 200).

During its heyday in the mid-2000s, XYZ produced many interesting holograms. Some of these are shown in Figures 14.24 through 14.27. Under the name of RabbitHoles, XYZ developed close connections with the movie industry in Hollywood and produced many excellent promotional display holograms for movies [31]. Most notable was their hologram of the 3D movie *Avatar* (Figure 14.28), which won the RealD award for Innovation in 3D. The company also produced holograms for trade show promotion. Notable clients included Bombardier, Audi, Toyota, Puma and Zeiss Meditec.

14.3.3 Digital Display Holograms for Mapping and Architectural Design

14.3.3.1 Horizontally Mounted Overhead-Illuminated Holograms

Horizontally mounted digital full-colour, full-parallax reflection holograms with overhead illumination are useful for a variety of applications in mapping and architectural design. Large full-parallax holograms may be laid out on a table and viewed by a group of people. 3D terrain can project up out of the hologram and viewers can easily perceive the 3D structure of mountains and valleys. Such holograms can potentially be rolled up into tubes and taken out for display when required. As such, they are of potential interest to military organisations.

At the time of writing, the US Army had purchased more than 10,000 horizontally mounted overhead-illuminated (HMOI) holographic displays for military mapping applications from Zebra Imaging. Zebra has also now started producing holographic printers for the US Army. The company has developed a new generation of high-speed colour and enhanced monochromatic digital hologram printers, which it calls *Imagers*. These allow the rapid production of digital holograms and, in particular, HMOI holographic displays. Like the printers made by Geola and XYZ, a wide variety of data sources can be used for the creation of the 3D images including computer-aided design/engineering/manufacturing, aerial photos, radar and laser scans such as synthetic aperture radar (SAR) and light detection and ranging.

The Hellenic Institute of Holography has also recorded map holograms for geographical services of various national armies, including the Greek Army. The institute recently cooperated with the NATO Fast Deployment Unit for the Balkan areas based in the city of Thessaloniki (NDC-GR) in the making of one trial map. In this case, geographical data, which were freely available on the web, were used (colour satellite terrain picture of an area near the Evros River in Northern Greece). Figure 14.29 shows a topographic reflection hologram, 50 cm × 41 cm, created by General Command of Mapping, Ankara, Turkey and printed by the Geola organisation [32].



FIGURE 14.29 (a) Normal relief map, and (b) topographic holomap or Digital Carto–hologram of the same area as in (a). Photos from “Ihsan Seref Dura Exhibition Room” at General Command of Mapping, Ankara, Turkey. (Reproduced with permission from H. Dalkiran, GCM, printed at Geola 2010.)

HMOI holographic displays can also be used for architectural and urban planning applications (Figure 14.30). Urban planners in the United Kingdom have long sought detailed physical models from developers for large residential or commercial developments. Such models allow an effective visualisation of a major project but can be very costly to make and extremely costly to modify when the planners request design changes. More recently, 3D virtual models have tended to replace such physical models. Such models are cheaper, can be modified easily and portray truly photorealistic detail, but usually, for convenience, the planner will see just a programmed “fly-through”. Writing such computer models to an HMOI display produces the best of both worlds. The HMOI hologram is cheap, easily transportable (unlike the model) and can be laid out quickly for viewing. Once lit, the display acts and feels just like a physical model except that the rendering is usually completely photorealistic. By replacing one display with another (an operation of several seconds), an alternative model can be displayed. In addition, design changes can easily be implemented in the computer model and another display printed—rather easier than changing the model!

14.3.4 Digital Holographic Colour Portraits

Despite early enthusiasm, pulsed-laser monochrome portraits never became terribly popular. Their single-colour waxy appearances seemed, on the one hand, strangely devoid of life and, on the other hand, appeared just too realistic. Digital holograms, however, can now offer a far more palatable solution. In the simplest variant, a series of photographs are taken using a holocam device. As we described in Chapter 10, this is just an automated camera on a horizontal rail. As the camera moves along the rail (typically from one to several metres long), digital photographs (typically several hundred to several thousand) are taken of a subject from different angles (Figure 14.31). The recorded images are then processed and a full-colour horizontal parallax-only (HPO), direct-write digital holography (DWDH)* portrait is printed. An example of such a hologram is the 30 cm × 40 cm 2006 colour holographic portrait of Nick Phillips shown in Figure 14.32. There are several advantages inherent to this technique. First and

* Full-colour reflection portraits can also be written using MWDH.



FIGURE 14.30 Full-parallax HMOI hologram displays by Zebra Imaging. (a) Double helix model, (b) military map and (c) city map. (Photos courtesy of M. Klug, Zebra Imaging).



FIGURE 14.31 Large-format 2005 studio holocam system manufactured by Geola.

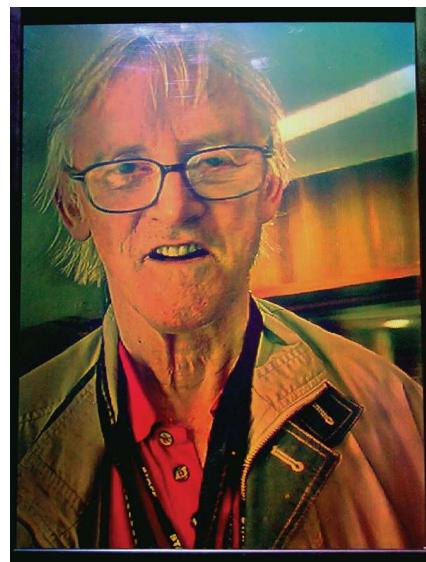


FIGURE 14.32 Full-colour digital hologram portrait (30 cm × 40 cm) of Nick Phillips shot at the 2006 *International Symposium on Display Holography* held in Wales and produced by Geola.

foremost, the hologram is a true-colour portrait, but almost as important as this is how the image data are derived. The subject is lit, not by highly coherent monochromatic laser beams, but by normal wide-band white illumination. This bypasses the old problems associated with the skin's reflective properties at certain laser frequencies. Because the image data is collected in digital form, the data may be easily manipulated and retouched. Backgrounds may be inserted and other features such as text inserted into the hologram. Finally, limited animation may be included, which makes the hologram feel more alive.

Full-parallax digital holographic portraits are also possible. In Chapter 10, we reviewed how structured-light techniques have led to the ability to acquire complete 3D computer image data from a subject in real-time and in Appendix 9 we review a revolutionary new technique which allows a single device such

as iPad from Apple Computers to acquire complete 3D data-sets. Such data sets allow the printing of full-parallax colour portraits using the DWDH technique.

14.3.5 Digital Art Holograms

Artists have been attracted to the new digital full-colour holograms. Paula Dawson in Australia, who has always been interested in large-format holograms, has recently created several large digital colour holograms which were printed by Geola. The *Luminous Presence* hologram ($0.95\text{ m} \times 1.5\text{ m}$), shown in Figure 14.33, was created in 2007. Martin Richardson at DeMontfort University in the United Kingdom has also created art portraits with several composite image sets. One such hologram is *Psychedelic Amy*, which is shown in Figure 14.34. At first sight, this is a digitally retouched portrait of a young woman, but a ghost-like image can be seen to reflect in the subject's eyes as one moves from left to right, giving the impression of sixties-like psychedelic imagery. Many other artists have started to work with this new and exciting medium.

14.3.6 Smaller Full-Colour Holograms

Small photopolymer holograms tend to be the easiest type of hologram to mass replicate. They also exhibit a high diffractive efficiency, and if image depth is small, they can be less sensitive to lighting conditions. As such, they are sometimes used on book or magazine covers. In 1999, a DAI Nippon



FIGURE 14.33 Full-colour digital hologram entitled *Luminous Presence* by Paula Dawson (printed by Geola, 2007).



FIGURE 14.34 Full-colour digital hologram entitled *Psychedelic Amy* by Martin Richardson (printed by Geola, 2010).

colour hologram was used on the cover of the Eighth Edition of *Holography Marketplace*. The mass-produced hologram shown in Figure 14.35 was recorded in DuPont's panchromatic photopolymer material. Although the DuPont material is not yet available to general customers, Ultimate Holography North America (UHN) in Canada has recorded similar colour holograms on the new panchromatic photopolymer material from Bayer—an example is shown in Figure 14.36.



FIGURE 14.35 Small, full-colour analogue reflection hologram by Dai Nippon (DNP) recorded using the DuPont panchromatic photopolymer.



FIGURE 14.36 Small, full-colour analogue reflection hologram by UHN recorded using the recent panchromatic photopolymer material from Bayer.



FIGURE 14.37 Model 840 Classica Brass Fire from Valor incorporating an analogue reflection hologram of the fire bed.

14.3.7 Holographic Fuel-Effect Electric Fires

To date, monochromatic holograms have only found display applications of significant sales volume in the document security field. One exception to this rule, however, is an application based on the use of a monochromatic reflection hologram in the mass-produced Valor *Electric Fireplace*. Valor has recently introduced a new type of household electric fireplace (Figure 14.37), in which a hologram is used to create a realistic-looking fuel bed of burning coal. Traditionally, fireplaces mimicking burning coal have made use of a bed of cut-glass “pebbles” illuminated by red light. However, such technology requires a bulky device that all too often simply does not fit into modern homes which frequently have no chimney. The great advantage of using a hologram to create the illusion of a fire is that it gives the illusion that the fireplace is much deeper than it really is. According to Valor, the design “changes the way you view fuel effect fires forever, offering you an unrivalled degree of realism”.

Valor’s fire bed hologram is relatively small—20 cm × 30 cm. The master reflection holograms were made by Colour Holographic Ltd. in the United Kingdom. Mass production of the holograms was then carried out by DuPont Holographics, Inc. in Logan, Utah using a monochromatic DuPont photopolymer film. The laminated transparent reflection hologram is mounted at an angle in the fireplace and is illuminated via a mirror by a 50 W, 12 V halogen lamp mounted at the rear of the device. Behind the hologram is a diffuser through which a flame effect is generated using material strips that move in the flowing air generated by a silent fan. These material strips are illuminated by zeroth-order light, which the hologram does not diffract. The diffuser makes the flame effect more realistic by making the moving material strips appear slightly diffuse through the hologram plate. Overall, the result is rather pleasing and the untrained eye does not immediately suspect that the fire is a hologram.

The reason this application works with a monochromatic hologram is that burning coals are often very red. As such, a red monochromatic hologram works satisfactorily. The company is nonetheless now working on an improved version of the fireplace, which will use a three-colour hologram.

14.4 Future Applications

14.4.1 Holographic Windows and Super-Realistic 3D Static Displays

One of the most intriguing applications of holography is the use of large-format digital holograms to mimic windows. This is perhaps the ultimate test of what is possible to extract from static holography.

The basic idea is to create glass holographic panels that, on appropriate illumination, reveal landscape-type images with virtual volumes of up to many square kilometres. At first sight, this might seem impossibly ambitious. After all, the standard credit card hologram might have, under ideal illumination, an image volume of 1 cm^3 , a modern panchromatic hologram of a typical museum artefact might possess an image volume of 0.3 m^3 , and even a large modern digital hologram might only have an image volume of say, 20 m^3 . Despite these figures, however, calculations clearly show that much larger image volumes should be possible.

Put simply, a holographic window is a large-format panchromatic digital reflection hologram that emulates the view outside a typical window. The characteristics of the hologram are chosen such that the intrinsic image blurring properties of the hologram, together with its illumination source, fall below the level of human discrimination. As such, an observer standing in front of this hologram interprets it, not as a hologram, but as a window. If properly implemented, such holograms have the ability to transform basements or small windowless rooms into top-floor penthouses with spectacular views over city centres or mountain ranges. Equally well, architects and interior designers could use this technology in commercial projects.

14.4.1.1 Creation of Space—High Virtual Volume Displays

Closely related to the concept of the holographic window is holography's ability to create space. A super-realistic hologram can effectively create virtual space behind its glass surface. Take for example, a small corridor in a typical London office suite. Often, such corridors are windowless and dark. Developers spend many millions of pounds opening up such small dark corridors with glass walls looking onto a central light well. Such developments produce nicer spaces in which to work, as the eye can see further and the darkness is replaced by light; but the central light well takes space and is extremely expensive. Holography can potentially achieve the same effect with glass holographic panels. Here, the extra space is all virtual and costs nothing. All that needs to be done to create the illusion of the brightly lit central light well is the installation of the panels and the illumination sources.

Holographic windows and super-realistic holographic displays will undoubtedly find a future market in architecture. Such displays may also be expected to find applications in museums. Imagine for example, going to the British Museum and seeing the pyramids in full-scale through a giant window display. Because such displays are intrinsically digital, one is not constrained to modern landscapes of course. Therefore, real-scale super-realistic images may be created of historical events. Even limited animation of the image may be included in such displays. The larger the display, the easier such animation is to encode and the longer it can be.

14.4.1.2 Constraints on Digital Holograms

Clearly, holographic windows or super-realistic high virtual volume (HVV) displays require digital full-parallax writing techniques such as DWDH. Analogue holograms could not, in any form or manner, be expected to generate such HVVs. Today's digital printers have not yet targeted HVV displays; these printers are capable only of producing digital holograms of rather limited depth. The actual depth possible depends on the field of view that the printer is required to create in the printed hologram. When the field of view is maximised, the capacity of the hologram to reproduce large depths decreases. Up until now, this has not been important as typical illumination sources have introduced greater blurring into the hologram than any restriction due to the printer's optical construction. In the following sections, we list the basic requirements for printing and displaying HVV and holographic window displays.

14.4.1.2.1 Optimal Hogel Size

The hogel size of a digital HVV display needs to be chosen in an optimal fashion. As we have already mentioned in Chapter 11, digital diffractive blurring becomes larger than the average human eye perception level at hogel diameters smaller than approximately 0.5 mm. This means that the virtual volume of a HVV display will decrease dramatically if the hogel size is reduced much beyond this point. Nevertheless, a hogel size of 0.5 mm is tolerable for a large-format display. Generally, the eye will

interpret this as a fine mesh covering the window when the display is viewed at a distance of less than 0.5 m.* At an observation distance of more than 0.5 m, the display will appear effectively non-pixelated.

14.4.1.2.2 Field of View

The field of view of a digital DWDH hologram is defined by the numerical aperture of the primary-colour writing objectives within the digital printer. A (horizontal) field of view for a DWDH hologram of between 100° and 120° is typical nowadays. Although HVV displays can be created with a field of view of only 120°, a holographic window ideally requires a field of view that is nearer to 140°. Values greater than these are usually not necessary as windows start to reflect the interior view as the angle becomes much flatter. In addition, an installation can be designed in certain circumstances such that panels may be mounted so that glancing angle views are physically obstructed. This reduces the requirement on field of view. One must also remember that field of view and hologram brightness are conjugate variables—the greater one is, the smaller the other is. As such, the vertical field of view may often be chosen to be rather smaller, thus increasing the effective brightness of the display.

14.4.1.2.3 Pixel Dimensions of Printer SLM

The three primary-colour spatial light modulators (SLMs) in a DWDH printer, together with the numerical aperture of the optical objectives, dictate the maximum possible angular resolution of the hologram at replay. The required horizontal numerical dimension[†] of the SLM is related to the field of view by the formula (Figure 14.38)

$$N \sim \cot \frac{\theta_{\text{Eye}}}{2} \times \tan \frac{\Phi}{2} \sim 2000 \times \tan \frac{\Phi}{2} \quad (14.1)$$

This means that for a horizontal field of view, $\Phi = 120^\circ$, one needs an SLM with a horizontal pixel dimension of 3464. Assuming a slightly smaller vertical field of view of $\Phi = 100^\circ$, the vertical dimension must be 2384. This is nearly double the 1080p HDTV standard of 1900 × 1200. Accordingly, today, even for a 100° × 120° hologram, one must use four 1080p SLM panels per primary colour. Combining the optical images of these panels is of course possible, but it considerably complicates the design of the printer. For a true 140° × 140° field of view, an SLM of dimensions around 5500 × 5500 is required. No doubt, such products will become available within the not-so-distant future and, at this point, digital printer design for HVV displays will become much simpler. Until then, one has to either use multiple panels or accept a smaller field of view if one is not to compromise virtual image volume. For instance, a 1080p panel equates to a maximum field of view of 87° × 62° (H × V), if one is not to incur blurring due to SLM insufficiency. If one does build a digital printer using a single 1080p panel (per primary colour channel) with an optical system having a higher numerical aperture, then one introduces a very significant blurring. The maximum image depth is given by the formula

$$x_{\text{CRIT}} = \frac{h}{\frac{1}{(N-1)} \cot \frac{\theta_{\text{Eye}}}{2} \tan \frac{\Phi}{2} - 1} \sim \frac{h}{\frac{2000}{N} \tan \frac{\Phi}{2} - 1} \quad (14.2)$$

If we choose a horizontal field of view of even 120°, then, at an observation distance of 0.5 m, we can expect to start to observe blurring due to the insufficiency of the printer SLM at an image depth of only 40 cm! The equations are unfortunately highly non-linear, and there is no way around this. A digital printer capable of printing HVV displays requires sufficient SLM resolution. And for the lowest possible standard of 100° × 120°, this equates to requiring four 1080p panels per primary colour channel.

* Of course DWDH itself has limitations in HVV display. For example very fine detail cannot be encoded “close-up”. For this computer generated holography (CGH) techniques must be used.

[†] Also called the horizontal pixel count.

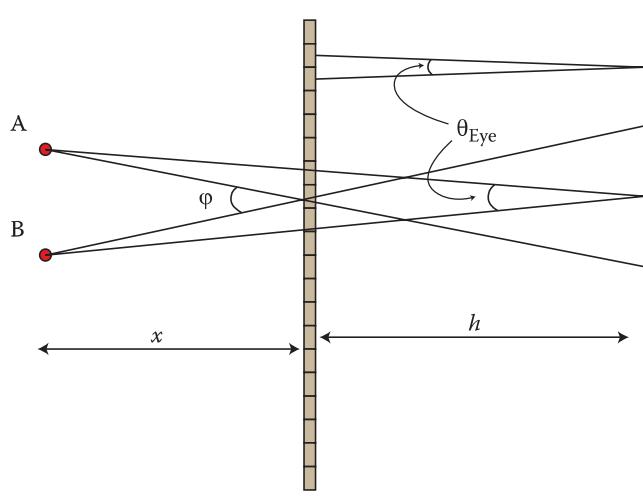


FIGURE 14.38 Diagram showing a side view (or overhead view) of a DWDH hologram. The two points (A and B) are just resolvable when θ_{Eye} is equal to the angular resolution of the human eye. However, the digital hologram can only replay the required rays so that A and B are actually resolved if (a) the SLM which records these rays has sufficient resolution and (b) the hogel size is below a critical value for a given observation distance.

14.4.1.2.4 Aberrations of Printer Optical System

A digital printer capable of printing HVV or window-type displays must have a very high numerical aperture optical system with extremely low aberration. Basically, the optical system must produce an image of the SLM at the given primary wavelength that is capable of resolving the SLM pixels. Holographic optical elements promise to provide a better solution to attaining this goal than current compound objectives.

14.4.1.2.5 Printer Variable Reference Beam System

To avoid introducing aberrations into the final hologram, it is important that the printer possesses a variable reference beam system that is able to properly synthesise a macroscopic point source. Current designs tend to use compound optical systems that work well for normal holographic displays but which may induce too much aberration for HVV displays. This is a technological issue only.

14.4.1.2.6 Constraints on Digital Data

Digital data must be available for each SLM pixel at each hogel coordinate. If part or all of the data originates from a Holocam device, then the Holocam must have the capability of generating the required image resolution. In the case of a final hologram field of view of $140^\circ \times 140^\circ$, a full-parallax Holocam might realistically need to generate approximately 25 million high-resolution images. If the image data originates from a computer model, then a similar number of render views will be required.

An HVV display differs fundamentally from a holographic window display, which has no foreground data. Such window holograms might portray distant landscapes. An example from the real world would be the view out of the window from an extremely high cliff-face. Only very distant features are visible. To produce this type of hologram, only one high-resolution photograph is required.

14.4.1.2.7 Constraints on Illumination Source

The illumination source is vital for all large-depth holograms. In Chapter 11, we reviewed the important role of source-size and chromatic blurring on the hologram. Most full-parallax reflection holograms nowadays are depth-limited primarily by the illumination source. In fact, it is common to use a broadband source and rely on the Bragg grating to filter the light; this gives the hologram a finite depth. Generally, the size of the source is, however, what finally determines the depth unless the hologram is particularly thick. Thick holograms become dim if lit by a broadband source; therefore, one usually restricts the thickness and tries to increase the permittivity modulation to optimise diffractive efficiency.

As we have seen in Chapter 13, new narrow-band semiconductor illumination sources of much better spatial coherence are now becoming available. This breaks the old equation. One no longer needs to rely on the Bragg grating to assure hologram depth. Of course, a certain frequency selection is still needed so that cross-talk between the three primary illuminating colours does not become an issue. However, chromatic blurring can now be determined solely by the bandwidth of the illumination source. As we have seen in Chapter 11, for chromatic blurring in the reflection hologram to fall below the level of human discrimination, the source needs to have a bandwidth of less than 1 to 2 nm. Speckle noise becomes a problem at bandwidths below 1 nm. Therefore, the ideal bandwidth for an HVV display source is 1 to 2 nm. Source-size blurring falls below the human perception level when the size of the source decreases to less than 1 mm per metre of distance from the hologram. To control image aberration, the wavelength of each illumination source must correspond exactly to a primary-colour laser recording wavelength.

The illumination constraints of HVV or holographic window displays are therefore identical to laser TV. This will undoubtedly mean that the price point of the best illumination sources will fall in due course.

14.4.1.2.8 Constraints on Photographic Material

The photosensitive material is an integral part of any HVV display. Because the angle of view is usually very large for these displays, the efficiency of the hologram must also be high. In addition, high image fidelity can only be attained if high spatial frequencies are recordable. As such, an extremely high-resolution material is required.

14.4.1.2.9 Summary

HVV and holographic window displays primarily require improvements in the current digital holographic printing technology. Currently, no printer has the required resolution. However, the modifications necessary are certainly feasible, even with today's technology. The illumination and materials technologies are also here. This is a real application of ultra-realistic imaging that appears to be well within reach but remains unrealised today.

14.4.2 Updateable 3D Holographic Displays

An updateable holographic display is a 3D display system in which the image data can be erased and reloaded. Following Geola's demonstration [33–36] in 1999 that pulsed lasers could be used effectively in a DWDH digital holographic printer to significantly increase the hogel-write rate and eliminate the interferometric stability problems of digital printers, Zebra filed a US patent in 2003 [37], which described an updateable holographic display using a pulsed laser. At the most basic level, all that was required to transform a pulsed laser DWDH printer into an updateable display system was a suitable material that could temporarily store a hologram before being erased. An experimental device based on a 4 in. × 4 in. photorefractive polymer screen* was then demonstrated in 2010 by workers at the University of Arizona in the United States [38]. Using a 50 Hz single-colour pulsed laser, the small display, capable of a 0.5% diffractive efficiency, could be updated every 2 s using a hogel size of 1 mm². Both single and full-parallax images were used. To attain full colour, the group used a transmission geometry with angle multiplexing.

The updateable holographic display is still in a very early stage of development; as such, it is still currently far from what one might describe as constituting an ultra-realistic display technology. Nevertheless, it is very possible that progress will be made here. Two important areas are the screen technology and the hogel write-rate. Clearly, the technology faces some significant challenges; a typical-size commercial display screen is unlikely to be smaller than 20 cm × 30 cm, and the required hogel size is unlikely to be greater than 0.5 mm. This leads to 240,000 hogels in the best of cases, which one assumes will need to be updated within some seconds to be useful. This in turn leads to a write-rate of approximately 100,000 hogels per second. Using optical multiplexing of a factor of 100, this leads to a required basic write-rate of 1 kHz. Pulsed lasers are certainly available at this repetition rate, and even though current twisted nematic liquid crystal displays cannot operate so fast, there are clearly ways around this. Perhaps more

* Produced by Nitto Denko Technical.

important, however, is the data processing required to create the high-resolution image data. Arrays of computers would be needed today to cope with the full-parallax data processing required to update a high-resolution screen of useful size. One answer could be to sacrifice the vertical parallax and to print only HPO holograms. Here, the data-processing issues are trivial, but there are far fewer applications for such restricted parallax displays. Given time, however, the cost of the required computational resources required for full-parallax calculation should fall dramatically.

Updateable holographic displays, unfortunately, suffer from the problem that current devices are based on fast macroscopic movement of an optical head over a suitable photosensitive material. This is a fundamental limitation; unless it can be addressed and a fully solid-state system is identified, the market for updateable displays is likely to be relatively specific. Progress in a number of areas will, in any case, be required to bring such displays to the level of a commercial product. Their first application is likely to be in the military domain, where somewhat larger updateable displays could be useful in mapping and tactical visualisation.

14.4.3 Real-Time 3D Display Technologies Based on Holography

14.4.3.1 Simple Autostereoscopic Systems Based on Holographic Screens

Glasses-free autostereoscopic screens for 3D visualisation systems may be produced using both reflection and transmission holograms [39,40]. Such systems offer some significant advantages over the better-known lenticular or barrier solutions: generally, these types of screens sacrifice two-dimensional image resolution to attain 3D, which is not the case for the holographic screen.

Generally, real-time holographic autostereoscopic visualisation systems work by projecting both a left-eye and a right-eye image, using standard digital projector technology, onto a special holographic screen. The screen then directs the respective images to each eye. By using ultrafine-grain holographic materials to record special gratings, high-resolution screens may be constructed. The difficulty with all standard simple glasses-free autostereoscopic visualisation systems (including those with holographic screens) is that the observer must hold his or her head in a unique horizontal position; this is the unique location where the two images fall into the correct eyes.

One way around this *observer location* problem is to use head tracking. Here, either the entire system, including the screen, continuously moves, tracking the observer's head or, alternatively, optics within the device move in order to move the left and right "eye-boxes". Both solutions are workable in some cases, although moving the optics has quite severe limitations. One of the authors (HB) has recently been involved in developing such a system for 3D visualisation by surgeons of medical endoscopy images with the US company Absolute Imaging Inc. Both transmissive and reflective systems are being developed. Analogue holographic screens have been recorded in cooperation with the Centre of Modern Optics in North Wales. Digitally written screens have been supplied by the Geola organisation. Figure 14.39 shows a prototype autostereoscopic system based on a transmission holographic screen.

14.4.3.2 Multiprojector Autostereoscopic Systems Based on Holographic Screens

Another way around the observer location problem is to use a holographic screen in conjunction with many projectors. Sang et al. [41] have recently produced such a system using a transmission holographic screen of $1.8\text{ m} \times 1.3\text{ m}$ using 64 digital projectors (Figure 14.40). The system is capable of displaying real-time full-colour HPO 3D images over a total horizontal viewing angle of 45° and can be viewed by many observers at once. Depths of up to 1 m are possible. The group have also developed an array of 64 digital cameras (640×480 pixels) so that real-time camera data may be displayed on the screen. A similar technology is being developed in Hungary [42–45]. These types of systems show some promise as extremely small and cheap pico-projectors are becoming available. This will potentially allow the incorporation of many more projectors, thereby improving depth and image resolution. Nevertheless, the fundamental architecture of such a system imposes certain limits and it is therefore difficult to see how useful full-parallax compact systems might be produced. The most likely scenario is that these systems will be adopted first for applications such as military engagement simulation in which large and bulky real-time HPO 3D glasses-free display systems could reasonably be employed for multiperson use.



FIGURE 14.39 *Absolute* medical autostereoscopic 3D display system being demonstrated to Dalibor Vukičević of the University of Strasbourg by *Absolute*'s president, William Pinkerton.

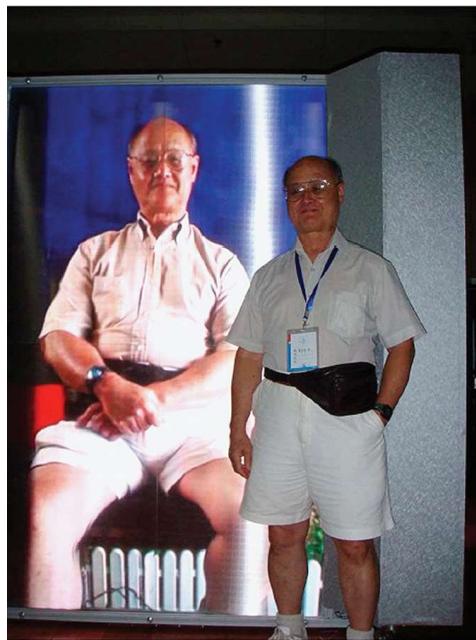


FIGURE 14.40 Professor Tung H. Jeong standing in front of the multiprojector autostereoscopic display system of Sang et al. [41]. The system is based on a large holographic screen, showing here the 3D portrait of Professor Jeong.

14.4.4 Future Real-Time True Holographic Displays

The ultimate application of holography is perhaps the true real-time display of full-parallax 3D images. Potentially, one can easily imagine a type of high-resolution two-dimensional transmission phase-modulated holographic screen (essentially a high-resolution SLM) that could be electrically updated in real-time. RGB laser illumination of such a screen could then be used to generate full-colour real-time images using temporal multiplexing to avoid chromatic cross-talk. Presumably, thin reflective screens could also be used, thereby producing a model for true 3D TV.

However, what appears in various science fiction films to be ostensibly simple is, in fact, rather difficult to achieve. The best way to understand why this is so is to consider the information rate required to write the 3D images. As we have seen in Chapter 9, even a discretised DWDH full-parallax hologram contains a lot of information. Supposing we are able to limit the optical information required to that pertaining to 1 mm^2 hogels,* then for a display with a reasonable field of view and a reasonable angular resolution, we can calculate that we will require approximately 300 MB of uncompressed memory storage per square centimetre of display. Because for real-time operation, the screen must be updated at a rate of at least 30 Hz, one can see that for a reasonably sized screen of say $40\text{ cm} \times 50\text{ cm}$, an information rate of nearly 20 TB/s will be required. In the simplest case of writing the real-time optical interference pattern digitally to a single high-resolution screen, the actual raw digital information required will be rather higher. For example, if one requires a field of view of 100° in both the vertical and horizontal directions, then any transmission grating capable of replaying this will require, using the results of Chapter 11, a maximum pixel spacing of $\lambda/2\sin(50^\circ)$ or roughly $0.35\text{ }\mu\text{m}$. For a $40\text{ cm} \times 50\text{ cm}$ display, this then equates to 2×10^{12} pixels. Using temporal multiplexing for three-colour display, this then gives an information rate of approximately 200 TB/s, but even the lower rate of 20 TB/s, which corresponds to producing a real-time hologram from a reduced data set, is extremely high for today's technology and would require massive parallelisation strategies for the image data calculations.

Another major problem is the lack of high-resolution electrically addressable displays. Twisted nematic liquid crystal displays are available today in 1080p panels down to pixel sizes of somewhat less than $10\text{ }\mu\text{m}$. Recently, Boulder Nonlinear Systems, Lafayette, CO, released a $2\text{ cm} \times 2\text{ cm}$ one-dimensional panel (composed of 12,288 liquid crystal columns rather than square pixels) with a pixel spacing of $1.2\text{ }\mu\text{m}$ and a gap between pixels of $0.6\text{ }\mu\text{m}$. The panel is intended for solid-state beam-steering applications and can deflect a beam over a range of around $\pm 5^\circ$. Although extremely encouraging, this is still some way from the submicron resolution required for holographic displays with a good viewing angle. Perhaps more important, however, is the fact that today's micropanels would need to be scaled up in area by approximately 1000 times to be useful as a simple holographic screen.

The lack of large ultrahigh-resolution electrically addressable spatial light modulators has led research groups in this field to adopt demagnification and stacking strategies. Several groups have been active in the field for some years. The first group to start serious work was Stephen Benton's group at MIT in the United States [46,47]. Benton used the HPO approximation to simplify these initial experiments, employing computational calculation of the fringe patterns, which were then written in sequence to a one-dimensional tellurium dioxide acousto-optic modulator. A demagnified image of this modulator was raster scanned to produce a composite HPO holographic image of $75\text{ mm} \times 150\text{ mm} \times 160\text{ mm}$ with an angle of view of 30° . The work was subsequently extended to full-colour. The most recent MIT work uses a new light modulator technology that uses surface acoustic waves in a slab of lithium niobate [48].

Work at QinetiQ in the United Kingdom [49] has focussed on full-parallax displays using optically addressable spatial light modulators—this type of SLM possesses a significantly better spatial resolution than current electrically addressable spatial light modulators. Computational calculation of the fringe patterns of a full-parallax hologram are performed using clusters of computers and the resulting data displayed on a 1024×1024 ferroelectric LCD running at several kilohertz. A demagnified image of this LCD is then (optically) tiled onto an optically addressable spatial light modulator using time multiplexing. A number of such modules can be stacked together to produce a larger full-colour, full-parallax display. Tiling from electrical to optical SLMs can currently produce an effective linear increase in SLM resolution of up to 10 times while maintaining video speeds. Although extremely impressive, these systems are still rather small and have a limited field of view. QinetiQ have so far demonstrated an effective 24 billion pixel display (Figure 14.41)—which really does prove that the technology can work—but the pixel count for a typical-size commercial screen will need to increase by nearly two orders of magnitude before such displays are optically equivalent to today's static full-colour digital holograms.

Stephen Benton was a firm believer that, given enough time, the computational task of calculating the fringe patterns required for full-colour, wide-angle real-time holographic display would cease to

* In fact, the quantisation of the hologram into hogels forms the basis of diffraction-specific algorithms for calculating the fringe structure.



FIGURE 14.41 Real-time digital, full-parallax holographic display (3×8 billion pixels) produced by QinetiQ.

become an issue. In 2004, the QinetiQ system used a cluster of 102 Pentium III CPUs each with 1 GB of memory to calculate the required real-time fringe patterns. Since then, as one might expect, the cost and physical size of computing power have decreased substantially. Indeed, it would seem that despite the seemingly enormous task of calculating real-time fringe patterns, if current trends continue, Benton will almost certainly be proved correct. When this stage is reached, ultra-realistic holographic imaging will truly have come of age.

REFERENCES

1. L. Solymar and D. J. Cooke, *Volume Holography and Volume Gratings*, Academic Press, New York (1981).
2. H. J. Coufal, D. Psaltis and G. T. Sincerbox (eds), *Holographic Data Storage*, Springer, New York (2000).
3. Ting-Chung Poon (ed.) *Digital Holography and Three-Dimensional Display—Principles and Applications*, Springer, New York (2010).
4. K. Curtis, L. Dhar, A. J. Hill, W. L. Wilson and M. R. Ayres, *Holographic Data Storage : From Theory to Practical Systems*, Wiley-Interscience, John Wiley & Sons Inc. (2010).
5. S. Benton and V. M. Bove Jr., *Holographic Imaging*, Wiley-Interscience, John Wiley & Sons Inc. (2008).
6. J. Robillard and H. J. Caulfield, *Industrial Applications of Holography*, Oxford University Press, New York (1990).
7. F. Havermeyer, L. Ho and C. Moser, “Compact single mode tunable laser using a digital micro-mirror device,” *Opt. Expr.*, **19**, 14642–14652 (2011).
8. P. Leisher, K. Price, S. Karlsen, D. Balsey, D. Newman, R. Martinsen and S. Patterson, “High-performance wavelength-locked diode lasers,” in *High-Power Diode Laser Technology and Applications VII*, M. A. Zediker, ed., Proc. SPIE **7198**, 7198–38 (2009).
9. S. L. Rudder, J. C. Connolly, G. J. Steckman, “Hybrid ECL/DBR wavelength & spectrum stabilized lasers demonstrate high power & narrow spectral linewidth,” in *Laser Beam Control and Applications*, A. V. Kudryashov, A. H. Paxton, V. S. Ilchenko, A. Giesen, D. Nickel, S. J. Davis, M. C. Heaven and J. T. Schriempf, eds., Proc. SPIE **6101**, 61010I 1–8 (2006).
10. S. J. Barden, J. A. Arns and W. S. C. Willis, “VPH gratings and their potential for astronomical applications,” in *Optical Astronomical Instrumentation*, S. D’Odorico, ed., Proc. SPIE **3355**, 866–876 (1998).
11. J. G. Robertson, K. Taylor, I. K. Baldry, P. R. Gillingham and S. C. Barden, “ATLAS: A Cassegrain spectrograph based on VPH gratings,” in *Optical and IR Telescope Instrumentation and Detectors*, M. Iye and A. F. Moorwood, eds., Proc. SPIE **4008**, 194–202, (2000).
12. P. K. Rambo, J. Schwarz, I. C. Smith, C. S. Ashley, E. D. Branson, D. R. Dunphy, A. W. Cook, S. T. Reed and W.A. Johnson, “Development of an efficient large-aperture high damage-threshold sol-gel diffraction grating,” *SANDIA REPORT*, SAND2004-5496 (2004).

13. H. Akbari and H. I. Bjelkhagen, "Pulsed holography for particle detection in bubble chambers," *Opt. Laser Technol.* **19**, 249–255 (1987).
14. J. Duplissy, M. B. Enghoff, K. L. Aplin, F. Arnold, H. Aufmhoff, M. Avngaard, U. Baltensperger, T. Bondo, R. Bingham, K. Carslaw, J. Curtius, A. David, B. Fastrup, S. Gagné, F. Hahn, R. G. Harrison, B. Kellett, J. Kirkby, M. Kulmala, L. Laakso, A. Laaksonen, E. Lillestol, M. Lockwood, J. Mäkelä, V. Makhmutov, N. D. Marsh, T. Nieminen, A. Onnela, E. Pedersen, J. O. P. Pedersen, J. Polny, U. Reich, J. H. Seinfeld, M. Sipilä, Y. Stozhkov, F. Stratmann, H. Svensmark, J. Svensmark, R. Veenhof, B. Verheggen, Y. Viisanen, P. E. Wagner, G. Wehrle, E. Weingartner, H. Wex, M. Wilhelmsson and P. M. Winkler, "Results from the CERN pilot CLOUD experiment," *Atmos. Chem. Phys.* **10**, 1635–1647 (2010).
15. H. Bingham, J. Lys, L. Verluyten, S. Willocq, J. Moreels, K. Geissler, G. Harigel, D. R. O. Morrison, F. Bellinger, H. I. Bjelkhagen, H. Carter, J. Ellermeier, J. Foglesong, J. Hawkins, J. Kilmer, T. Kovarik, W. Smart, J. Urbin, L. Voyvodic, E. Wesly, W. Williams, R. Cence, M. Peters, R. Burnstein, R. Naon, P. Nailor, M. Aderholz, G. Corrigan, R. Plano, R. L. Sekulin, S. Sewell, B. Brucker, H. Akbari, R. Milburn, D. Passmore and J. Schneps, "E-632 Collaboration: Holography of particle tracks in the Fermilab 15-Foot Bubble Chamber," *Nucl. Instr. and Meth.* **A297**, 364–389 (1990).
16. X. Sang, C. Yu, M. Yu and D. Hsu, "Applications of digital holography to measurements and optical characterization," *Opt. Eng.* **50**, 091311-1-8 (2011).
17. A. Osanlou, "Holographic digital microscopy in on-line process control," *Opt. Eng.* **50**, 091312-1-4 (2011).
18. J. Watson, "Submersible digital holographic cameras and their application to marine science," *Opt. Eng.* **50**, 091313-1-5 (2011).
19. L. Xu, J. Miao and A. Asundi, "Properties of digital holography based on in-line holography," *Opt. Eng.* **39**, 3214–3219 (2000).
20. D. Carl, B. Kemper, G. Wernicke and G. Von Bally, "Parameter-optimized digital holographic microscope for high-resolution living cell analysis," *Appl. Opt.* **43**, 6536–6544 (2004).
21. P. Jourdain, N. Pavillon, C. Moratal, D. Boss, B. Rappaz, C. Depeursinge, P. Marquet and P. J. Magistretti, "Determination of transmembrane water fluxes in neurons elicited by glutamate ionotropic receptors and by the cotransporters KCC2 and NKCC1: A digital holographic microscopy study," *Journal of Neuroscience* **31**, (No. 33)11846–11854 (2011).
22. H. Bjelkhagen, D. Boggett, P. Crosby, P. Henderson, S. Lilington, E. Mirlis, B. Napier, A. Osanlou, S. Rider, A. Robertson and A. Shore, "Full colour holographic endoscope (Holoendoscope) project" UK-DTI TP/3/IMG/6/I/15727 (2006).
23. H. I. Bjelkhagen, J. Chang and K. Moneke, "High-resolution contact Denisyuk holography," *Appl. Opt.* **31**, 1041–1047 (1992).
24. M. Friedman, H. I. Bjelkhagen and M. Epstein, "Endoholography in medicine," *J. Laser Appl.* **1**, 40–44 (1988).
25. H. I. Bjelkhagen, "Color holography for museums: bringing the artifacts back to the people," in *Practical Holography XXV: Materials and Applications*, H. I. Bjelkhagen, ed., Proc. SPIE **7957**, 0B-1 – 7 (2011).
26. H. I. Bjelkhagen and J. Cook, "Colour holography of the oldest known work of art from Wales," *The British Museum Technical Research Bulletin* **4**, 87–94 (2010).
27. A. Lembessis, "Realistic colour 3D holography: is it here?" *HOLO-PACK HOLO-PRINT 2011*, Las Vegas (2011).
28. H. I. Bjelkhagen and D. Vukičević, "Color holography: a new technique for reproduction of paintings," in *Practical Holography XVI and Holographic Materials VIII*, S. A. Benton, S. H. Stevenson and T. J. Trout, eds., Proc. SPIE **4659**, 83–90 (2002).
29. Y. Gentet and M. K. Shevtsov, "Mobile holographic camera for recording color holograms," *J. Opt. Technol.* **76**, 399–401 (2009).
30. M. Klug, "Display applications of large scale digital holography," in *Holography: A Tribute to Yuri Denisyuk and Emmett Leith*, H. J. Caulfield, ed., Proc. SPIE **4737**, 142–149 (2002).
31. V. Bates, "Down the Rabbit hole – Avatar Neytiri with Sprite," *HOLO-PACK HOLO-PRINT 2010*, Malaysia (2010).
32. H. P. Dalkiran, "Hologram – The future of cartographic publishing (Holocartography)," *HOLO-PACK HOLO-PRINT 2010*, Malaysia (2010).
33. D. Brotherton-Ratcliffe, F. M. Vergnes, A. Rodin and M. Grichine, *Method and apparatus to print holograms*, Lithuanian Patent, LT4842, (1999).

34. D. Brotherton-Ratcliffe, F. M. Vergnes, A. Rodin and M. Grichine, *Holographic printer*, US Patent 7,800,803 (filed 1999, granted 2010).
35. D. Brotherton-Ratcliffe and A. Rodin, *Holographic printer*, US Patent 7,161,722 (filed 2002, granted 2007).
36. A. Rodin, F. M. Vergnes and D. Brotherton-Ratcliffe, *Pulsed multiple colour laser*, EU Patent, EPO 1236073 (2001).
37. M. A. Klug, C. Newswanger, Q. Huang and E. Holzbach, *Active digital hologram display*, US Patent 6,859,293 (filed 2003, granted 2005).
38. P.-A. Blanche, A. Bablumian, R. Voorakaranam, C. Christenson, W. Lin, T. Gu, D. Flores, P. Wang, W.-Y. Hsieh, M. Kathaperumal, B. Rachwal, O. Siddiqui, J. Thomas, R. A. Norwood, M. Yamamoto and N. Peyghambarian, “Holographic three-dimensional telepresence using large-area photorefractive polymer,” *Nature* **468**, 80-83, (4 Nov. 2010).
39. C. Newswanger, *Real time autostereoscopic displays using holographic diffusers*, US Patent 4,799,739 (filed 1987, granted 1989).
40. D. Brotherton-Ratcliffe, H. Bjelkhagen and J. Fischbach, *Holography apparatus and system*, US Patent Application, Patent Serial Number 11/870,442 (2007).
41. X. Sang, F. C. Fan, C. C. Jiang, S. Choi, W. Dou, C. Yu and D. Xu, “Demonstration of a large-size real-time full-color three-dimensional display,” *Opt. Lett.* **34**, 3803–3805 (2009).
42. T. Balogh, *Method and apparatus for displaying three dimensional images*, US Patent 6,201,565 (filed 1998, granted 2001).
43. T. Balogh, *Method and apparatus for generating 3D images*, PCT Application WO2005/117458A2 (2004).
44. T. Balogh, *Method and apparatus for Displaying 3D images*, US Patent 6,999,071 (filed 2000, granted 2006).
45. T. Balogh, *Method and apparatus for producing 3D pictures*, PCT Application, WO9423541A1 (1993).
46. M. Luente, S. A. Benton and P. St.-Hilaire, “Electronic holography: the newest,” in *Int'l Symposium on 3D Imaging and Holography*, Osaka, Japan, November (1994).
47. P. St-Hilaire, S. A. Benton, M. Luente, J. Underkoffer and H. Yoshikawa, “Realtime holographic display: improvements using multichannel acousto-optic modulator and holographic optical elements,” in *Practical Holography V*, S.A. Benton, ed., Proc. SPIE **1461**, 254–261 (1991).
48. D. E. Smalley, *High-Resolution Spatial Light Modulation for Holographic Video*, Thesis submitted for Master of Science in Media Technology, MIT (2008).
49. C. Slinger, C. Cameron, S. Coomber, R. Miller, D. Payne, A. Smith, M. Smith, M. Stanley and P. Watson, “Recent developments in computer-generated holography: toward a practical electroholography system for interactive 3D visualisation,” in *Practical Holography XVIII: Materials and Applications*, T. H. Jeong and H. I. Bjelkhagen, eds., Proc. SPIE **5290**, 27–41 (2004).

15

Acronyms

A/D	Analogue to digital
AgX	Silver halide
ANSI	American National Standards Institute
AR	Antireflective (coating)
CAD	Computer-aided design
CAE	Computer-aided engineering
CAM	Computer-aided manufacturing
CCD	Charge-coupled device
CEU	Control electronics unit
CFL	Compact fluorescent lamp
CFSL	Continuous full-spectrum lighting
CGH	Computer-generated hologram
CIE	Commission Internationale de l'Eclairage
CLP	Centre for Laser Photonics (located in OpTIC, Wales, UK)
CMO	Centre for Modern Optics (located in OpTIC, Wales, UK)
CMOS	Complementary metal oxide semiconductor
CRI	Colour rendering index
CT	Computed tomography
CW	Continuous wave (laser)
CW	Coupled wave (theory)
CWC2	Cook and Ward developer
CWT	Coupled wave theory
D/A	Digital to analogue
DCG	Dichromated gelatin
DE	Diffraction efficiency
DEA	Dielectric elastomer actuator
DHM	Digital holographic microscopy
DHR	Digital holographic recording
DI-HO	Digital input – holographic output
DMX	DMX512-A, the ESTA standard for controlling lighting equipment and related accessories
DPSS	Diode-pumped solid-state
DPSSL	Diode-pumped solid-state laser
DWDH	Direct-write digital holography
ECDL	External cavity diode laser
EDTA	Ethylenediaminetetraacetic acid
EM	Electro-magnetic
ERS	Ellipsoidal reflector spotlight
ESTA	Entertainment Services Technology Association
FFT	Fast Fourier transform

FOV	Field of view
F/P	Fabry-Perot (interferometer)
FWHM	Full-width half-maximum
GCM	General Command of Mapping (Turkish military organisation)
GEOLA	General Optics Laboratory (Anglo-Lithuanian group of Companies from 1995 now comprising Geola Digital UAB in Lithuania and Geola Technologies Ltd. in the UK)
GUI	Graphical user interface
H ₁	Master hologram to be used for copying
H ₂	Transmission or reflection copy of a master hologram (H ₁)
H ₃	Hologram copied from an H ₂ (e.g., an embossed hologram)
HDTV	High-definition television display
HeNe	Helium–neon laser
HeCd	Helium–cadmium laser
HF	High-frequency (electromagnetic radiation)
HiH	Hellenic Institute of Holography
HIRF	High-intensity reciprocity failure
HiPER	High power laser energy research facility
HMOI	Horizontally mounted, overhead-illuminated (type of hologram)
HOE	Holographic optical element
Hogel	Holographic element (the component hologram of a DWDH hologram)
Holocam	Holographic (image data) camera (camera on a motorised rail for 3D image acquisition)
Holopixel	Holographic pixel (same as hogel)
HPIV	Holographic particle image velocimetry
HPO	Horizontal parallax-only (type of hologram without vertical parallax)
HRLF	Holographic reciprocity law failure
HUD	Heads-up display
HVV	High virtual volume (type of holographic display)
HWHM	Half-width half-maximum
ICNIRP	International Commission on Non-Ionizing Radiation Protection
IEC	International Electrotechnical Commission
IMS	Industrial methylated spirits
IPS	Isopropyl alcohol
IR	Infrared
I-to-S	Transformation from image data to SLM data in a digital DWDH or MWDH printer
ISIS	Interference security image structures
Laser	Light amplification by stimulated emission of radiation
LCD	Liquid crystal display
LCOS	Liquid crystal on silicon (reflective SLM)
LED	Light-emitting diode

LIDAR	Light detection and ranging
LIRF	Low-intensity reciprocity failure
LSR	Laser speckle reducer
MAX	Abbreviation for 3D StudioMax™ (3D modelling program)
MaxSCRIPT™	Max Scripting language (for 3D modelling)
MB	Methylene blue (dye)
MCU	Main camera unit
MMA	Methyl methacrylate
MOPA	Master oscillator power amplifier
MPGH	Multiple-photo generated holography
MRI	Magnetic resonance imaging
MTF	Modulation transfer function
MWDH	Master-write digital holography
N-CWT	N-Coupled wave theory
N-PSM	N-Parallel Stacked Mirror (theory)
NIKFI	Cinema and Photographic Research Institute in Moscow
NIP	National ignition facility
OEM	Original equipment manufacturer
OPO	Optical parametric oscillator
OTF	Optical transfer function
OVD	Optical variable device
PCB	Printed circuit board
PBU	Phillips Bjelkhagen ultimate bleach
PC	Personal Computer
PET	Polyethylenterephthalate
PMC	Printer and monitoring controller (software package used by Geola UAB in its 2001 DWDH printing systems)
PMMA	Polymethyl methacrylate
PSM	Parallel stacked mirror (theoretical model of a colour holographic grating)
PTP	Peak-to-peak
PVA	Polyvinyl alcohol
R6G	Rhodamine 6G (dye)
RCA	Royal College of Art (London, UK)
RCW	Rigorous coupled wave (theory)
RCWT	Rigorous coupled wave theory
RFM	Ripping, filing and monitoring software package (used by Geola UAB in its 2001 DWDH printing systems)
RGB	Red, green, blue
RMS	Root mean square
SAR	Synthetic aperture radar
SBS	Stimulated Brillouin scattering
SCR	Silicon controlled rectifier
SHG	Second harmonic generation
SHSG	Silver halide–sensitised gelatin
SLM	Single longitudinal mode
SLM	Spatial light modulator

SM-6	Popular developer for pulsed AgX materials (Salim's mistake)
SRS	Stimulated Raman scattering
SSDL	Solid-state dye laser
TEA	Triethanolamine
TEM ₀₀	Transverse electromagnetic mode 00
TFT	Thin-film transistor
THG	Third harmonic generation
TIR	Total internal reflection
TMG	1,1,3,3-Tetramethylguanidine (chemical promoter)
TTL	Transistor-transistor logic
UV	Ultraviolet
VCSEL	Vertical surface cavity emitting laser

Laser Crystals

Nd:YAG	Neodymium-doped yttrium aluminium garnate
Nd:YALO	Neodymium-doped yttrium aluminium oxide
Nd:YAP	Neodymium-doped yttrium aluminium perovskite
Nd:YLF	Neodymium-doped yttrium lithium fluoride
Nd:YVO ₄	Neodymium-doped yttrium orthovanadate
Yb:YAG	Ytterbium yttrium aluminium garnate

Non-linear Crystals

BBO	Beta barium borate
DKDP	Deuterated potassium dihydrogen phosphate
KDP	Potassium dihydrogen phosphate
KNbO ₃	Potassium niobate
KTP	Potassium titanyl phosphate
LBO	Lithium triborate
PPKTP	Periodically poled KTP

Passive Q-Switches

Co:MALO	Cobalt-doped MALO—Co:MgAl ₂ O ₄
Cr:YAG	Chromium-doped yttrium aluminium garnate
V:YAG	Vanadium-doped yttrium aluminium garnate

Semiconductor Materials

AlGaN	Aluminium gallium nitride
AlGaP	Aluminium gallium phosphide
GaAs	Gallium arsenide
GaAsP	Gallium arsenide phosphide
AlGaAsP	Gallium arsenide phosphide
GaInN	Gallium indium nitride
GaN	Gallium nitride
GaP	Gallium phosphide
SiC	Silicon carbide
ZnSe	Zinc selenide

Chemicals

Acetic acid	CH ₃ COOH
Amidol	(see diaminophenol dihydrochloride)

Ammonium dichromate	$(\text{NH}_4)_2\text{Cr}_2\text{O}_7$
Ammonium rhodanide	(see Ammonium thiocyanate)
Ammonium thiocyanate	NH_4SCN
Ammonium thiosulphate	$(\text{NH}_4)_2\text{S}_2\text{O}_3$
L-Ascorbic acid	$\text{CH}_2\text{OHCHOH}(\text{CHCOH}:\text{COHCOO})$
p-Benzoquinone	$\text{C}_6\text{H}_4\text{O}_2$
Benzotriazole	$\text{C}_6\text{H}_4\text{NHN}_2$
Borax	(see Sodium tetraborate, decahydrate)
Bromine	Br
Calgon	(see Sodium hexametaphosphate)
Carbon tetrachloride	CCl_4
Catechol	$\text{C}_6\text{H}_4\text{-}1,2\text{-(OH)}_2$
Chlorohydroquinone	$\text{ClC}_6\text{H}_3\text{-}1,4\text{-(OH)}_2$
Chrom (III) acetate hydroxide	$(\text{CH}_3\text{CO}_2)_7\text{Cr}_3\text{OH}_2$
Chromium III chloride	$\text{CrCl}_3 \cdot 6\text{H}_2\text{O}$
Chromium III nitrate	$\text{Cr}(\text{NO})_3 \cdot 9\text{H}_2\text{O}$
Chromium III potassium sulphate	$\text{CrK}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$
Citric acid, monohydrate	$\text{HOC}(\text{COOH})(\text{CH}_2\text{COOH})_2 \cdot \text{H}_2\text{O}$
Cupric bromide	CuBr_2
Cupric sulphate	CuSO_4
Cupric sulphate, pentahydrate	$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$
Decahydronaphthalene	$\text{C}_{10}\text{H}_{18}$
Decalin	(see Decahydronaphthalene)
2,4-Diaminophenol dihydrochloride	$(\text{NH}_2)_2\text{C}_6\text{H}_3\text{OH} \cdot 2\text{HCl}$
Dimethyl phalate	$\text{C}_8\text{H}_4(\text{COOH CH}_3)_2$
EDTA	(see Ethylenediaminetetraacetic acid)
Elon	(see p-Methylaminophenol sulphate)
Ethanol	(see Ethyl alcohol)
Ethylenediaminetetraacetic acid	$(\text{HOOCCH}_2)_2\text{NCH}_2\text{CH}_2\text{N}(\text{CH}_2\text{COOH})_2$
Ethyl methyl ketone	$\text{CH}_3\text{:C}(\text{CH}_3)\text{COO CH}_3$
Ferric chloride	$\text{FeCl}_3 \cdot 4\text{H}_2\text{O}$
Ferric nitrate, nonahydrate	$\text{Fe}(\text{NO}_3)_3 \cdot 9\text{H}_2\text{O}$
Ferric sulphate	$\text{Fe}_2(\text{SO}_4)_3$
Formaldehyde	HCHO
Formalin	(see Formaldehyde)
Glutaraldehyde	$\text{CH}_2(\text{CH}_2\text{CHO})_2$
Glycerin	(see Glycerol)
Glycerol	$\text{CH}_2\text{OHCHOHCH}_2\text{OH}$
Hydroquinone	$\text{C}_6\text{H}_4\text{-}1,4\text{-(OH)}_2$
Hypo	(see Sodium thiosulphate, pentahydrate)
Iodine	I_2
Isopropyl alcohol	$(\text{CH}_3)_2\text{CHOH}$
Kodalk	(see Sodium metaborate)
Methanol	(see Methyl alcohol)
Methyl alcohol	CH_3OH
p-Methylaminophenol sulphate	$(\text{HOC}_6\text{H}_4\text{NHCH}_3)_2 \cdot \text{H}_2\text{SO}_4$
4-Methylaminosulphate	$\text{CH}_3\text{NH}(\text{C}_6\text{H}_4)_2\text{H}_2\text{SO}_4$
Methylphenidone	$\text{C}_{10}\text{H}_{12}\text{N}_2\text{O}$

Methyl methacrylate	$\text{CH}_2:\text{C}(\text{CH}_3)\text{COO CH}_3$
Metol	(see <i>p</i> -Methylaminophenol sulphate)
Paraformaldehyde	$(\text{CH}_2\text{O})_x$
Phenidone (A)	(see 1-Phenyl-3-pyrazolidone)
Phenidone (B)	(see Methylphenidone)
Phenosafranine	$\text{C}_{18}\text{H}_{15}\text{ClN}_4$
<i>p</i> -Phenylenediamine	$\text{C}_6\text{H}_4(\text{NH}_2)_2$
1-Phenyl-3-pyrazolidone	$\text{C}_6\text{H}_5-\text{C}_3\text{H}_5\text{N}_2\text{O}$
Potassium baborate	(see Potassium tetraborate)
Potassium bisulphite	(see Potassium hydrogen sulphate)
Potassium bromide	KBr
Potassium carbonate	K_2CO_3
Potassium citrate	$\text{K}_3\text{C}_6\text{H}_5\text{O}_7 \cdot \text{H}_2\text{O}$
Potassium dichromate	$\text{K}_2\text{Cr}_2\text{O}_7$
Potassium dihydrogen orthosulphate	$\text{K}_2\text{H}_2\text{PO}_4$
Potassium ferricyanide	$\text{K}_3\text{Fe}(\text{CN})_6$
Potassium hydrogen sulphate	KHSO ₄
Potassium hydroxide	KOH
Potassium iodide	KI
Potassium metabisulphite	$\text{K}_2\text{S}_2\text{O}_5$
Potassium nitrate	KNO_3
Potassium permanganate	KMnO ₄
Potassium persulphate	$\text{K}_2\text{S}_2\text{O}_8$
Potassium phosphate, dibasic	K_2HPO_4
Potassium pyrosulphite	(see Potassium metabisulphite)
Potassium rhodanide	(see Potassium thiocyanate)
Potassium thiocyanate	KSCN
Propanol	(see Propyl alcohol)
Propyl alcohol	$\text{CH}_3\text{CH}_2\text{CH}_2\text{OH}$
Pyrocatechol	(see Catechol)
Pyrogallol	$\text{C}_6\text{H}_3-1,2,3-(\text{OH})_3$
Quinol	(see Hydroquinone)
Quinone	(see <i>p</i> -Benzoquinone)
Rhodamine 6G	$\text{C}_{28}\text{H}_{31}\text{N}_2\text{O}_3\text{Cl}$
Rhodamine B	$\text{C}_{28}\text{H}_{31}\text{ClN}_2\text{O}_3$
Silver nitrate	AgNO_3
Sodium acetate	CH_3COONa
Sodium bisulphite, monohydrate	$\text{NaHSO}_4 \cdot \text{H}_2\text{O}$
Sodium borohydride	NaBH ₄
Sodium carbonate, anhydrous	Na_2CO_3
Sodium hexametaphosphate	$(\text{NaPO}_3)_6$
Sodium hydrogen carbonate	NaHCO_3
Sodium hydrogen sulphite	(see Sodium bisulphite)
Sodium hydroxide	NaOH
Sodium metabisulphite	$\text{Na}_2\text{S}_2\text{O}_5$
Sodium metaborate tetrahydrate	$\text{NaBO}_2 \cdot 4\text{H}_2\text{O}$
Sodium metasilicate	$\text{Na}_2\text{SiO}_3 \cdot 5\text{H}_2\text{O}$
Sodium nitrite	NaNO ₂

Sodium sulphate	Na_2SO_4
Sodium sulphite, anhydrous	Na_2SO_3
Sodium tetraborate, decahydrate	$\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$
Sodium thiocyanate	NaSCN
Sodium thiosulphate, anhydrous	$\text{Na}_2\text{S}_2\text{O}_3$
Sodium thiosulphate, crystal, pentahydrate	$\text{Na}_2\text{S}_2\text{O}_3 \cdot 5\text{H}_2\text{O}$
Sulphuric acid	H_2SO_4
Tetraethylammonium bromide	$(\text{C}_2\text{H}_5)_4\text{NBr}$
Thiocarbamide	(see Thiourea)
Thiourea	H_2NCSNH_2
Triethanolamine	$(\text{HOCH}_2\text{CH}_2)_3\text{N}$
Urea	NH_2CONH_2
Vitamin C	(see Ascorbic acid)
<i>o</i> -Xylene	$\text{C}_6\text{H}_4(\text{CH}_3)_2$

Appendix 1: Historical Origins of Display Holography: Spreading Awareness

A1.1 Hologram Exhibitions

It has always been important to make people aware of holography, and one of the best ways to do this has been through the organisation of hologram exhibitions. The very first hologram exhibition took place in 1968 at Cranbrook Academy in Michigan. Margaret Benyon exhibited her first holograms in 1969 at the Nottingham University Art Gallery in England. In 1975, a large exhibition, *Holography '75: The First Decade*, was held at the International Center of Photography in New York. It was organised by Joseph (Jody) Burns, Jr. and Rosemary (Posy) Jackson and represented the work of artists and scientists from the United States and six other countries. Artists and scientists met for the first time to acknowledge each other's presence in the field—this resulted in the first use of the word *holographer*.

A sense of community grew. In March 1976, Burns and Jackson arranged the first exhibition outside the US at the Cultural Centre in Stockholm, Sweden: *Holografi: Det 3-Dimensionella Mediet*. The exhibition was extremely successful, with 60,000 visitors during the 2 weeks (12–28 March) over which it was held, breaking all previous attendance records for the most people per day, per week and per event. The second exhibition in New York, *Through the Looking Glass*, opened in December 1976, this time at the newly opened Museum of Holography. In 1977, the exhibition moved to Toronto in Canada and later to other places.

In 1977, the first French hologram exhibition, *Sculptures de Lumière*, took place in Strasbourg. Nick Phillips of Loughborough University was part of the HOLOCO team (with Anton Furst and John Wolff) and was involved in recording large-format display holograms. HOLOCO was supported by the rock group, the WHO. Phillip's holograms were on display at two large exhibitions, both arranged at the Royal Academy of Arts in London: *Light Fantastic I* in 1977 and *Light Fantastic II* in 1978. The first exhibition was very successful, with 96,000 people in 28 days and created a lot of interest in holography. The second exhibition attracted 250,000 visitors.

Many successful hologram exhibitions were arranged in many countries in the late 1970s and early 1980s. *Alice in the Light World* was one such exhibition held in August 1978 at the Isetan Department Store in Tokyo, Japan. Many of Nicholson's portraits from Hawaii were part of this show. In 1979, the number of exhibitions increased; shows occurred, for example, in Eindhoven, Rotterdam, Berlin, Milan and Liverpool in Europe. In 1980, there were many international exhibitions and also one of interest in London—*Light Years Ahead*—at the Photographer's Gallery; this show broke all previous attendance records.

Another UK exhibition was *Light Dimensions*, which took place in Bath in 1983 at the National Centre of Photography. This was the world's biggest hologram exhibition thus far and was organised by Eve Ritscher Associates in London. Her Royal Highness Princess Margaret, Countess Snowdon, opened the exhibition on 21 June 1983. The exhibition moved to The Science Museum in London in December that year and continued until the end of April 1984.

The Art and Science of the Soviet Union was a 1985 exhibition of holograms from the former USSR organised at the Trocadero in London. Mostly, it contained monochrome Denisuk holograms of artefacts from the Hermitage and other museums (Figure A1.1). A total of 132 holograms—many, very large rainbow plates—were on display at the exhibition *Découvrez l'Holographie*, which was organised by Ap-Holographie in France, and supported by Philip Morris, France and Paris Match. It opened in May 1985 and took place at the Palais de la Découverte in Paris.

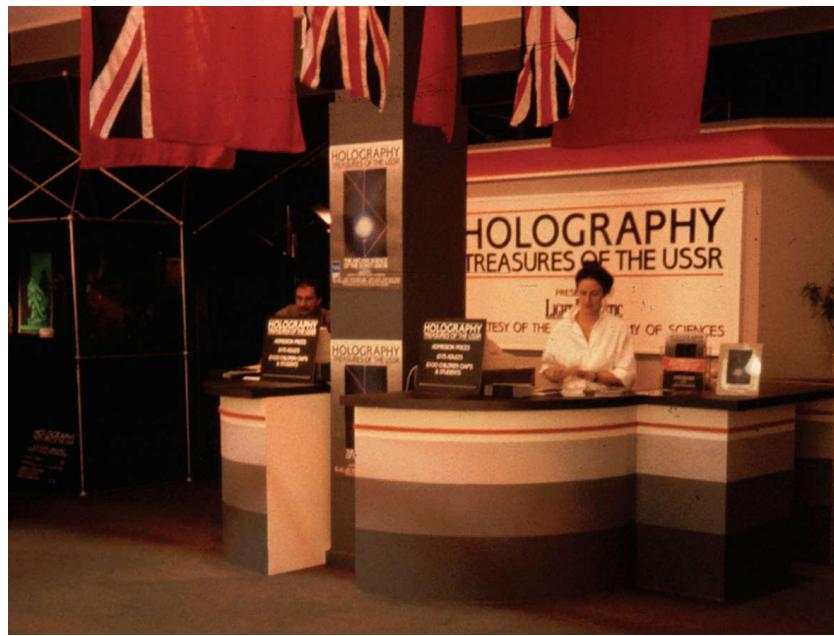


FIGURE A1.1 Trocadero exhibition entrance.

Teit Ritzau in Copenhagen, Denmark, arranged a permanent hologram exhibition *Holography or Reality* at the *Holographic World* in the Tivoli Gardens, H. C. Andersen's Castle, in 1986. It had many large Denisyuk holograms from the Cinema and Photographic Research Institute (NIKFI) in Moscow. The holograms were integrated in specially designed rooms. It continued for many years.

The travelling exhibition, *Images in Time and Space*, was referred to as “the most significant international holography exhibition of the decade.” It was organised by the Associates of Science and Technology Inc. in Ottawa, Canada. It opened at Montreal’s EXPOTEC in 1987. With more than 170 historical, scientific and artistic works, it drew large crowds, estimated at 250,000, in its first three month period. In 1988, it moved to San Jose, California where the exhibition ran into financial problems. This marked the end of large international holographic exhibitions.

Smaller exhibitions, however, continued to be organised in association with holographic conferences such as those at Lake Forest and the Millennium Conference on Holography in St Pölten, Austria. After the fall of the Berlin Wall, there were also a number of significant exhibitions of holography in the countries of Eastern Europe. One such exhibition was organised by one of the authors (DBR) in Bucharest, Romania in 1996. The exhibition, which featured many large-format holograms from Australian Holographics and Geola, was held at the Bucharest Polytechnical University and attracted many visitors; the opening ceremony was attended by the vice president of the house of deputies, Ion Ratiu. Sizeable exhibitions were also held in Hungary, Lithuania and Estonia.

In 2009, Jonathan Ross in London started to arrange touring hologram exhibitions in the United Kingdom, mainly with holograms in his collection. These exhibitions attracted many visitors, most of whom were aware only of credit card holograms at best, and who now had the opportunity to see large art and display holograms for the first time. A large Belarusian exhibition, *HOLOEXPO 2011*, with holograms from around the world, opened in Minsk in October 2011. In addition to various historic holograms, it included many new analogue and large digital colour holograms.

A1.2 Commercial and Educational Entities Involved in Holography

A1.2.1 Companies Producing Holograms

Over the years, holographic companies and individual holographers around the world have produced holograms and hologram products. One of the first companies to offer holograms for sale was *The Holography Company*.

Corporation in Norristown, Pennsylvania. The company's president, Larry Goldberg, claimed that they were "creating a new reality". The holograms were bleached transmission film holograms delivered with a converter filter, peaked at a wavelength of 580 nm, designed to be used with a slide projector for illumination of the hologram with monochromatic light. The company also offered the "solar powered" *Spectrol hologram*, which was a rainbow hologram that could be viewed in sunlight or illuminated with an ordinary tungsten light bulb. The holograms were distributed through the Edmund Scientific company in the United States.

In the early 1970s, companies started to offer custom-made display holograms or limited edition holograms produced by artists and commercial holographers around the world. Only a few of these companies remain today and, if they are still active, they are mainly producing holograms for the document security market. *Hologram Industries* run by Hughes Souparis in France is an example of this.

A1.2.2 Hologram Galleries

In addition to advertising or exhibition holograms, holographic products and holograms were also manufactured for shops and galleries. Outlets were established in many shopping malls and city centres during the 1980s. There were many hologram galleries established in both Europe and the United States. One of the first to open a gallery was Gary Zellerbach with his *HOLOS Gallery* in San Francisco. In 1981, Peter Woodd opened Britain's first permanent gallery of holography named *Light Fantastic*, which was located in Covent Garden in London. However, most of these galleries closed in the early to mid-1990s, including the *Light Wave Gallery* in North Pier, Chicago shown in Figure A1.2.



FIGURE A1.2 Part of the *Light Wave* hologram gallery in Chicago.

A1.2.3 Holographic Museums

Various holographic museums opened during the 1970s, but most of them have closed now. One such museum was Posy Jackson's *Museum of Holography* located in New York, which opened in 1976. The important historic hologram collection of this museum is now preserved as part of the *MIT Museum*. One of the earliest and the longest running museums, which opened in 1977, was Loren Billings' *Gallery 1134: Fine Arts Research and Holographic Center*. Later, it was renamed as the *Holographic Museum in Chicago*. It closed in 2010 when Loren Billings was not able to continue running it because of age and poor health. In 1977, Denisyuk opened the *Museum of Holography* in St. Petersburg; now closed, the holograms are kept in the Vavilov Institute.

In Europe, Matthias Lauk's *Museum für Holographie & Neu Visuelle Medien* in Pulheim near Cologne, Germany, opened in 1979 with a large collection of holograms, including several of the McDonnell Douglas holograms, but it is no longer open. In 1980, Anne Marie Christakis opened the *Musée Française de l'Holographie* in Paris, but again, this museum closed some years ago. Christakis retained the collection of holograms and arranges exhibitions now and then. Another German museum, *Holowood*, was opened in the 1990s by Matthias Frieb in Bamburg and included several of the large-format rainbow holograms from Australian Holographics. Frieb's museum closed when the Bamburg city council undertook extended public works and tourists could no longer easily reach the museum's location.

A1.2.4 Educational Institutes

In the beginning, university physics and electrical engineering departments around the world were the main places where students could learn about holography. Many of the museums also opened educational centres aimed at display and art holography with artist-in-residence programmes. Tung H. Jeong (TJ) at Lake Forest College, located outside Chicago, started in 1972 to offer week-long courses on holography at the college and, later, he arranged the first *International Symposium on Display Holography* with associated workshops in 1982. Jeong's conferences focussed on art and display holography and attracted both scientists and artists from around the world. Jeong was also behind the *School of the Art Institute of Chicago*; he started giving lectures on holography there in 1975. Ten years later, holography became a full-time instructional area. Specialised schools of holography for artists opened as well—for example, Lloyd Cross' school in San Francisco, which opened in 1971. In the beginning, the San Francisco area became a main centre for artists to learn about holography. The *Holography Institute* opened in 1980 and, for many years, was run by Jeffrey Murray.

In 1973, Jody Burns, Jr. opened *The New York School of Holography*. In 1978, it was taken over by Daniel Schweitzer (1946–2001) and Samuel Moree and is now run by Moree as *Holographic Ocean Laboratories*. In 1974, David Hlynksy and Michael Sowdon opened the *Fringe Research* educational centre in Toronto, Canada. In Europe, in 1980, Michael Wenyon started to run workshops at the *Goldsmiths' College of Art* in London; he continued this activity until 1983. London became the United Kingdom's centre for educational art and display holography through the *Royal College of Art* courses and Edwina Orr's private *Richmond Holographic Studios*. Both places offered pulsed holography courses. Dieter Jung's *Academia of Media Arts* in Cologne, Germany also attracted many art holography students over the years. In Japan, the *Holography School of the Tamara Art College* in Kawasaki opened in 1980.

Appendix 2: History of the Geola Organisation

A2.1 The Beginning

The chain of events that led to the creation of Geola can probably be traced back to a trip made by one of us, David Brotherton-Ratcliffe, to Soviet Russia in the late 1980s just before perestroika. David returned from the three-week trip, famished but deeply impressed by what he had seen in Moscow and Leningrad. At the time, he had been working as a postdoctoral fellow at the Flinders University of South Australia on theoretical physics problems connected with controlled nuclear fusion. In 1986, he had met another Flinders postdoctoral fellow, Igor Bray, now director of the Institute of Theoretical Physics at Curtin University in Western Australia. It was with Igor that David made this first trip to Russia.

David left Flinders University in 1989 to form his own company, Australian Holographics Pty. Ltd., with the aim of manufacturing large-format holograms for advertising and display. While searching for laboratory optics and lasers for this company, David met Dmitry Konovalov, a newly arrived Russian postdoc at Flinders. Dmitry, who now worked with Igor, told David that it might be possible to source the items that Australian Holographics needed from the Soviet Union. This was now the time of perestroika and the borders of the Eastern Bloc were starting to open. The two men guessed that if Soviet laser and optical equipment turned out to be useful to Australian Holographics, then perhaps universities and other companies in Australia would also be interested. With this idea in mind, a new company, General Optics Pty. Ltd. was formed and Dmitry started to look for contacts in the Soviet Union.

On 11 March 1990, Lithuania had declared independence from Soviet Russia and it became possible to think seriously about exporting from Russia through Lithuania to the West. By mid-1991, Dmitry had made contact with Lev (Leon) Isacenkov, who was an early postindependence Lithuanian entrepreneur—Leon spoke fluent Russian, Lithuanian and English—and could also be contacted, albeit with great difficulty, by telephone. By the end of 1991, Dmitry had come up with a list of possible photonic products, all with price tags of around a hundred times cheaper than equivalent western products. The one problem was that there was no serious possibility of transporting the goods out of Lithuania. David, remembering the fascination of his earlier trip to Soviet Russia, therefore, decided to drive a lorry himself to Lithuania to pick up the first shipment for General Optics (Figures A2.1, A2.2 and A2.3).

David set off from London in early February 1992. He invited John Fenton, a London music producer to accompany him on the trip for company and also for greater security. The trip took them through France, Belgium, West Germany, East Germany, Poland and finally into Lithuania through the Lazdijai border crossing. The roads at this time were often rather poor and weather conditions in Poland were difficult. The border crossings to both Poland and Lithuania had immense queues of traffic. The Lithuanian queue was easily several miles long; here, David and John understood through somewhat awkward attempts to communicate with other drivers that the wait could last many days. With temperatures going down to -20°C , they decided to take their chances and made their way to the front of the enormous queue. After being stopped by armed Polish guards with submachine guns, they were forced to stay in the lorry for several hours. Eventually John decided to brave it and walked to the soldier's cabin with the idea of explaining that he was a western journalist. Somehow, this worked and 10 minutes later, the lorry was waived through to the Lithuanian side. Here, John distributed copies of western music magazines to the soldiers who became amazingly friendly and welcomed the two Englishmen to Lithuania. As the lorry left the Lazdijai border crossing, the Lithuanian border guards, music magazines in hand, all gathered to wave them off.

Unfortunately, with the bad roads, bad weather and the problems with the border crossings, the lorry was 30 hours late for its scheduled meeting with Leon Isacenkov. They had agreed to all meet up at the border crossing, but after 12 hours of waiting, Leon had returned to Vilnius, disappointed. David and



FIGURE A2.1 David Brotherton-Ratcliffe, in early 1992, seen here driving General Optics' first lorry back from Lithuania.



FIGURE A2.2 Leon Isacenkov (early 1992) in the Vilnius offices of Infortechnika UAB during the first meeting with David Brotherton-Ratcliffe.

John had barely enough petrol left in the lorry to reach Vilnius. In those days, there were no satellite navigation systems and all the road signs in Lithuania were still written in Cyrillic. David had been awake for probably 45 hours by now, as John did not drive. The travellers therefore decided to park the lorry and sleep.

In the morning, David and John awoke to find the lorry surrounded by people who had apparently never seen a western lorry before. This would become a familiar pattern that would follow the two men pretty much wherever they went in the coming weeks. In the centre of Vilnius, with the petrol gauge firmly on empty, David parked the lorry, and once again, the lorry was surrounded, this time by a group of children. By chance, one of the children, a boy who was around 11 years old, spoke a little English; David asked him if he could somehow make a telephone call to the telephone number he had for Leon Isacenkov. The child took David back to his parents' apartment where he dialled Leon's number using the family telephone. The child's parents were both out. Within 30 minutes Leon arrived.



FIGURE A2.3 Large building on Naugarduko Street (seen here in 2003), which has been home to Geola since 1994. Geola occupies around 1000 m² of laboratories and offices over three floors. The entrance is almost exactly in the centre of the picture, just in front of the bus stop.

A2.2 First Meeting in Vilnius

Leon had worked in a well-known electronics institute in Vilnius. After Lithuania's declaration of independence, however, he decided to leave and set up his own business, Infortechnika UAB, assembling and selling computers. Over the next few weeks, David accompanied Leon on numerous visits to factories and universities in Lithuania and in Moscow. At the time, the train connection between Vilnius and Moscow was still uncontrolled and it was easy to go between the two cities. By the end of the trip, an official cooperation had been signed by General Optics and Infortechnika for the export of scientific equipment. David returned to the United Kingdom with a lorry full of holographic plates, large-frame helium-neon lasers, optical tables, holograms and optical components (Figure A2.4). This was the first of what would be many such journeys.

The equipment purchased in Lithuania from Infortechnika proved extremely interesting. General Optics started to make sales, most notably to university departments. Australian Holographics opened a shop called Rainbow Bridge in South Australia, which sold, amongst other things, the Russian holograms from General Optics. Australian Holographics also started to use the Soviet equipment itself in its own laboratories.

As their joint business started to produce results, Leon and David decided that they needed a laboratory in Vilnius to test the equipment being exported. Orders soon became quite complex, with university customers ordering custom-pulsed lasers. These lasers needed verification before shipping. With this in mind, Leon asked his friend and colleague, Viktor Karaganov to head up a small team responsible for this task in Vilnius. Within the year, however, Viktor had decided to take up an offer of employment at Flinders University in Australia (invited by Igor and Dmitry's group) and with his wife, Sveta, left for Australia in 1993. In Australia, Sveta became office manager and accountant for Australian Holographics. Previously, she had been an accountant and lawyer for Infortechnika. Viktor also became

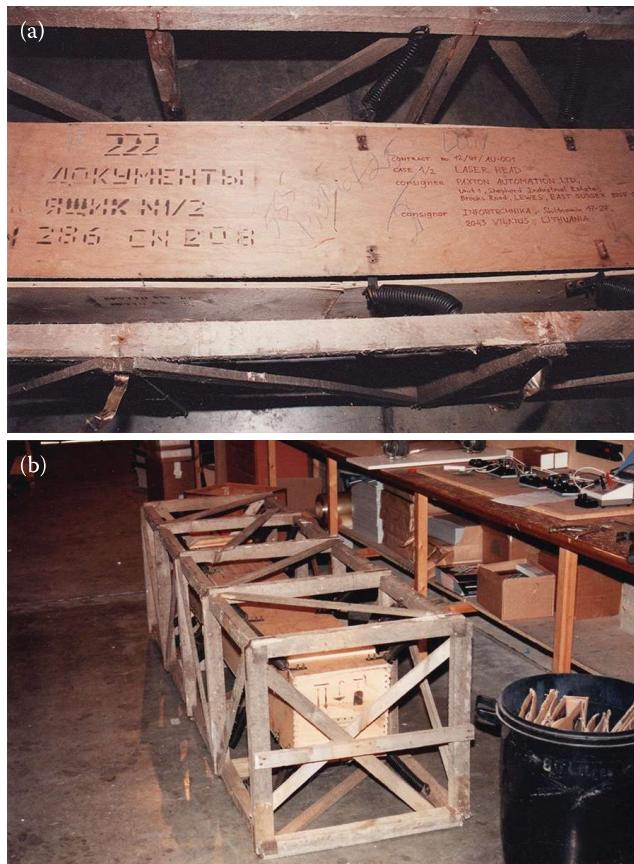


FIGURE A2.4 Large-frame He–Ne lasers brought back from Lithuania on the first lorry in 1992 being unloaded at Paxton Automation Ltd. in the United Kingdom. Note the contract marking visible in (a) on one of the cases: 12/91/AU-001—this was the first of many contracts between General Optics and Infotechnika.

a part-time scientific consultant for the company. In Vilnius, Mikhail Grichine left his job at a large scientific laser laboratory in the Shatura region outside Moscow and moved to Vilnius with his wife Galia to replace Viktor. Mikhail was an expert in Soviet laser technology and saw the possibility not just to export lasers but also to build them.

In fact, as the sales of General Optics got better and better, a general problem appeared with the manufacturers in Russia. Their electricity was being turned off, they were complaining about unserviceable equipment and salaries were not being paid. The vast Soviet scientific infrastructure was in effect collapsing. In certain cases, laboratories and companies managed to rearrange themselves in time and avoid being closed. In other cases, activities ceased. In any case, it was slowly becoming more difficult to source the cheap optical items Dmitry Konovalov had identified in 1991. The decision was therefore taken to follow Mikhail's suggestion and to design and actually produce pulsed lasers in Vilnius.

In parallel to the events in Vilnius and Australia, in 1993, David took part in several artistic exhibitions of holography in Paris. In one such exhibition, he met Florian Vergnes, then a young Parisian entrepreneur just out of university. Florian was fascinated by the large holograms that David displayed at the exhibition. When David showed him examples of Soviet optical components several days later, Florian proposed that they open a joint company together. David accepted and, with several other associates, they formed LMC France Instruments SARL. Within several months, LMC had signed a rental contract and the company started business in St. Denis. For the next few years, David lived mostly in Paris, travelling frequently to Adelaide and Vilnius. In the United Kingdom, Ralph Cullen from UK Optical Supplies Ltd. had become a distributor for General Optics and was helping with UK sales. LMC developed a customer base across mainland Europe. Simon Edhouse had joined Australian Holographics

and was now also sales director for General Optics. Simon started to look after sales in Australasia and David concentrated more on Europe.

A2.3 Incorporation of Geola UAB

By some time in early 1994, problems with the supply end of the import/export business had reached a level where it was decided that more resources had to be directed at producing a stable western-type office in Vilnius. More effort needed to be spent forging relationships with newly emerging Soviet enterprises and in-house manufacturing had to be accelerated. With this in mind, David moved from Paris to Vilnius, leaving Simon Edhouse to run General Optics in Australia and Florian Vergnes to run LMC in France. With Mikhail Grichine and Leon Isacenkov, David identified two new engineers who joined the Vilnius team. These were Gleb Skokov (Figure A2.5) and Aleksej Rodin. Gleb had just finished technical university in optical physics and Aleksej was in the middle of his PhD in quantum electronics. A new company called Geola UAB was formed. It was Mikhail who thought up the name—an acronym for General Optics Laboratory.

Serendipitously, just at this time, the large Soviet electronics institute on Naugarduko Street, where Leon had worked during the communist period, was partially privatised and it became possible, for a minimal price, to buy rooms in this enormous building. Both David and Leon took the opportunity and brought in builders to renovate the spaces. Within months, Geola had a small modern office in the heart of Vilnius (Figure A2.3).

Geola worked closely with General Optics, LMC, but also with Australian Holographics. In Vilnius, they built a pulsed holographics facility based on a 5 J neodymium laser for the creation of large-depth transmission holograms (Figure A2.6). Several of Australian Holographics' large stock holograms were produced in this studio as well as holograms for well-known artists such as H. R. Giger (Figure A2.6).

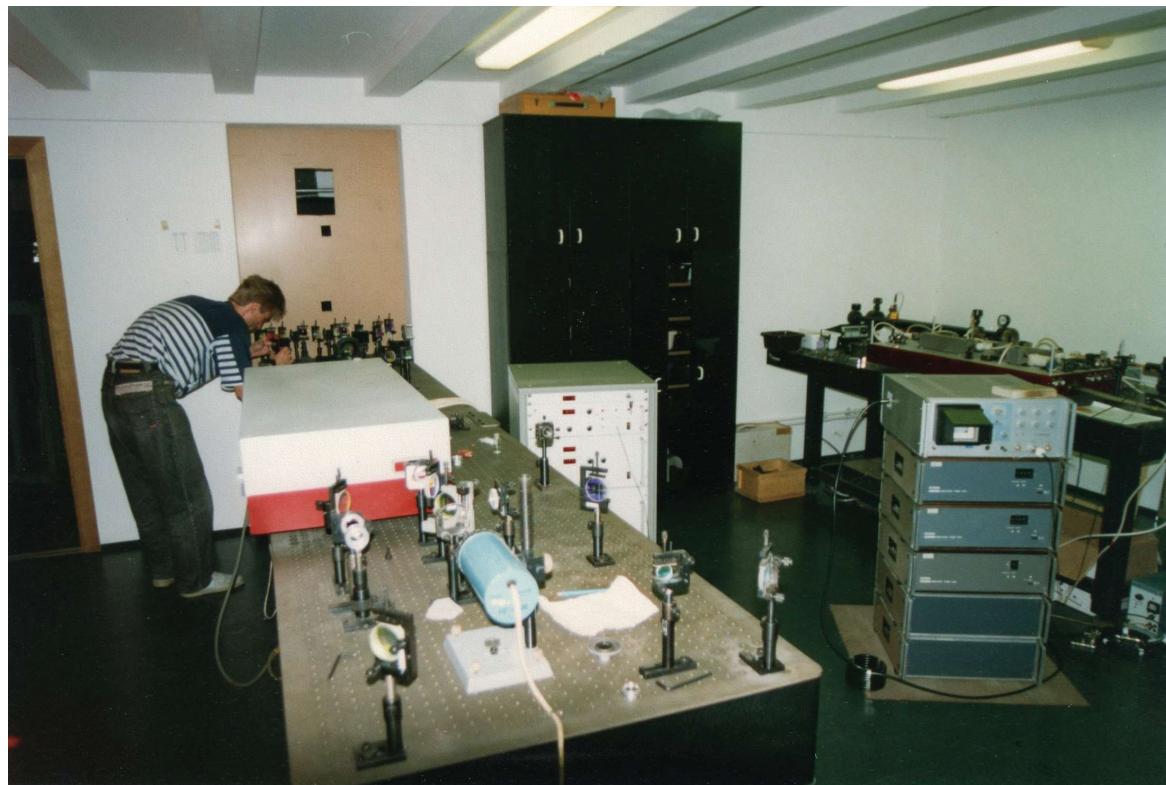


FIGURE A2.5 Gleb Skokov working in 1994 at Geola UAB in the main laser room.

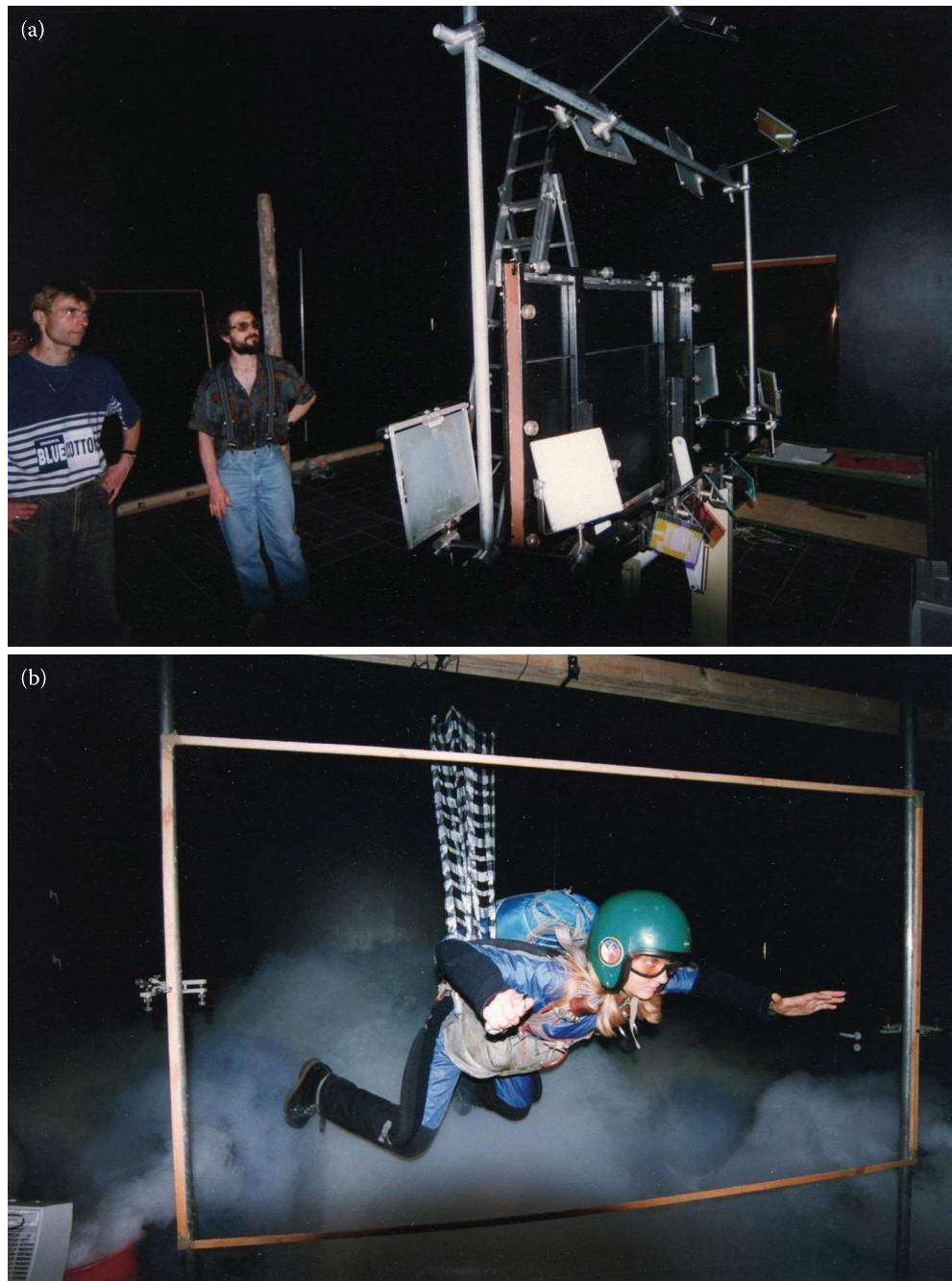


FIGURE A2.6 Pulsed hologram room at Geola UAB in 1994. (a) Gleb Skokov and Mikhail Grichine looking at the large plate holder and laser diffuser system. (b) A large-format rainbow master hologram being shot for Australian Holographics.

Sergei Vorobiov visited and taught Mikhail and David the basics of human pulsed laser portraiture. Aleksej Rodin worked on building a Raman laser to make colour pulsed holograms.

In the winter of 1994, David unfortunately contracted pneumonia—there was often no heating in the average Vilnius apartment then and if you wanted to remain clean, one was obliged to take cold showers. Of course, in the winter, the gas pressure fell, so you could forget about heating the water. Most people showered rarely to cope with this phenomenon, but westerners like David found this habit unacceptable. The inevitable occurred, and David was forced to travel back to Paris, having undergone an old Soviet treatment that would leave him weak for some years to come.

A2.4 First Romanian Exhibition of Large-Format Holography

Some months later, partially recovered from the pneumonia, David was walking near his apartment in the centre of Paris, when he noticed that he was in a street called “*rue de Bucarest*”. Having nothing better to do that weekend and having never visited Romania, he decided to buy an airline ticket to Bucharest. In Bucharest, he tried to identify people who could be interested in holography. During this process, he met several people by chance whom he became close friends with. Over the next few months, David visited Bucharest and Transylvania many times, after which he decided that it would be a good idea to spend some more time there. In mid-1995, he rented a flat in Petru Poni Street next to the north railway station and began organising the first Romanian exhibition of large-format holography with Professor Chisleag from the Polytechnic University of Bucharest.

The Romanian exhibition was a great success, but afterwards, David fell ill again and remained in Romania for more than a year convalescing. During this time, the Australian Holographics business was reorganised; Simon Edhouse formed his own company to market the holograms (this became Australian Holographics Pty. Ltd.), and David retained the manufacturing facility (which became Australian Holographic Studios Pty. Ltd.). Geoff Fox remained chief holographer for some time at AH Studios until his apprentice, Mark Trinne, took over. Florian Vergnes, delegating his duties at LMC, moved to Vilnius and became general manager of Geola UAB. By 1997, Geola had grown to around 10 people and the company had already constructed and delivered its first neodymium-pulsed laser. It had also produced its first holographic portraiture system, and by 1998, it was selling such systems. The company now occupied a space of around 1000 m². Florian visited the Slavich company in Moscow many times and signed an exclusive distribution agreement for the international sale of Russian holographic materials.

In late 1998, David, now completely recovered from his earlier illness, returned to live in Vilnius. He and Mikhail decided to look at making digital holograms with pulsed lasers. Soon, they had a small pulsed laser printer assembled that could print master H₁ type holograms using a Sony LCD display. At this time, Eric Bosco, then working at Pixel Systems Inc. in Montreal, contacted Geola and ordered an 8 J pulsed neodymium laser. Subsequently, it seemed that Eric’s company was interested in a larger collaboration to develop digital pulsed holography. Geola and Pixel started negotiations for an extensive partnership whereby Geola was to undertake a funded scientific research programme with a view to creating a commercial full-colour digital holographic printer. Pixel wanted two types of systems investigated—single-step and two-step printers. Geola had already demonstrated a monochromatic two-step system but lacked appropriate RGB pulsed lasers for a colour system. During the negotiations, David worked out how to build both types of printing system. Mikhail Grichine and Aleksej Rodin also thought they knew how to build the pulsed lasers. In 1999, David and Florian flew to Montreal to sign the multi-million dollar contract between Pixel and Geola. On the aeroplane back, the two men celebrated with a bottle of wine and David, joking, mentioned that the only thing that could go wrong was if the Canadian company was put into liquidation.

Unfortunately, Pixel was indeed inexplicably wound up the next day and Geola was forced to tear up the contract. Nevertheless, Geola had advanced during the negotiations with Pixel to such a point that it was now capable of producing both one- and two-step printers. Within several months, it had demonstrated the technology in its laboratories and David had filed a Lithuanian patent for the first RGB pulsed digital holographic printer.

A2.5 Start of XYZ

Eric Bosco was, however, still on the case and sought to remedy the mess at Pixel by reorganising the company’s digital holography team into a new company, XYZ Imaging Inc. Within a year, Eric had reached an agreement with Geola, Pixel’s main shareholder and with several Canadian venture capital organisations to go ahead with the project. Contracts were signed in Montreal and XYZ and Geola joined forces to develop digital colour pulsed holography. Geola became the major shareholder—a situation that lasted for quite some time—and had two directors on the board of XYZ. Nonetheless, it was the

institutional shareholders who really controlled XYZ, as it was their money. XYZ was to develop one-step printing, but Geola retained rights to its laser and two-step printing technology. The first task for Geola under the new agreements was to construct a digital RGB pulsed laser printer, which was installed in Montreal in 2001 (Figure A2.7). Mikhail Grichine left the company to work at Ekspla UAB, another well-known laser manufacturer in Vilnius, and Stanislovas Zacharovas, who had already been working with Geola for several years as head of sales and marketing, became Geola's director. Aleksej Rodin became chief laser engineer and played a pivotal role in producing a stable pulsed RGB laser. Marcin Lesniewski designed the critical optical writing objectives.

A new UK company, Geola Technologies Ltd., was formed in 2000. Some Canadian investors were nervous about building a manufacturing organisation in an ex-Soviet country like Lithuania. The new UK company would manufacture the critical RGB pulsed lasers required by the XYZ printers. It would also do research for new types of more compact lasers for the future. David moved back to the United Kingdom to become managing director of Geola Technologies Ltd. Aleksej Rodin also moved to the United Kingdom, as did Nataly Vidmer, one of Geola's highly talented precision mechanical engineers. Over the next few years, the team in Vilnius grew to around 30 people whereas the UK team increased to around 10. In Montreal, XYZ hired around 25 people, some of whom came from Geola or were trained in Geola.

Every so often, David and Eric would meet somewhere (Paris, London, Vilnius, Montreal, etc.) to discuss the way forward. Both men were adamant that their respective teams should cooperate well. All too often, sensitivities between the Lithuanian, English and Canadian teams developed which could, without intervention, erupt into disputes. There were also sensitivities on the side of XYZ's institutional investors—some found it difficult to work with a company in an ex-Soviet country and felt intrinsically uncomfortable with this. Eric managed both sides of the equation with great mastery. Although he now had some brilliant engineers in-house—Eugene Kosenko, Jean-Jacques Cotteverte, Vladimir Fedorenchik, Roman Rus and many others—he realised that XYZ and Geola still had much to gain through close cooperation.

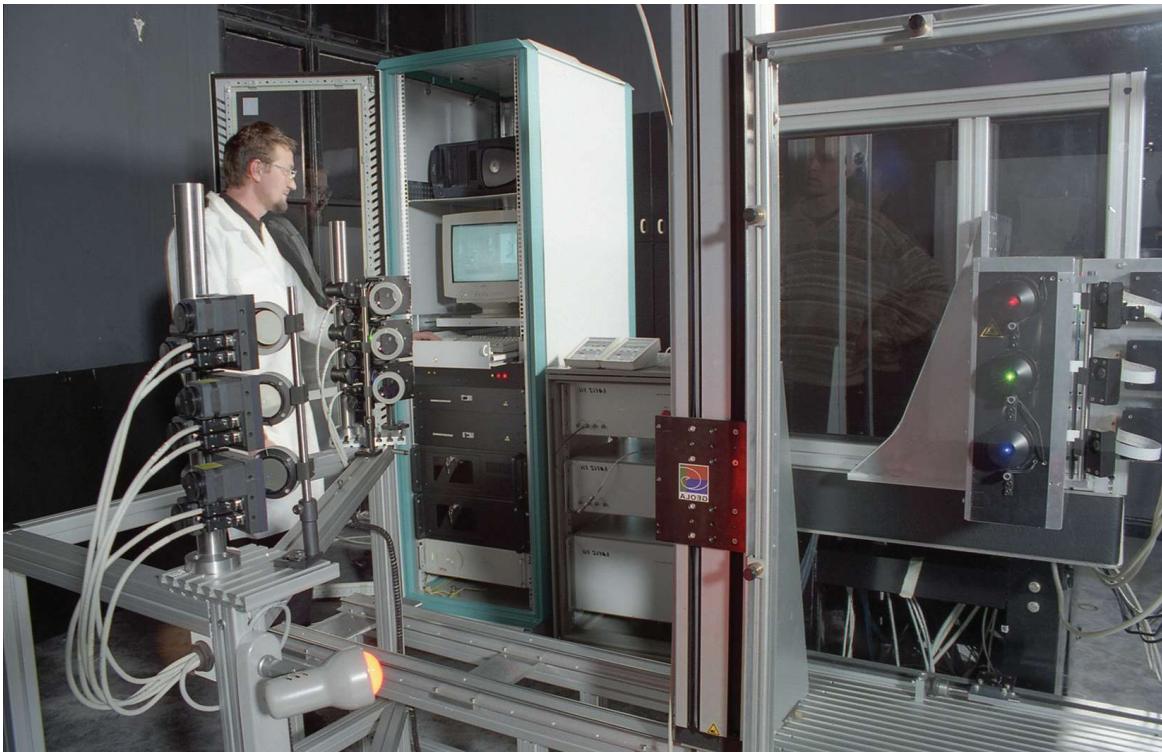


FIGURE A2.7 Very first RGB pulsed laser digital holographic printer manufactured in Vilnius at Geola UAB for XYZ Imaging Inc., seen here in 2001 with Geola's first electronics engineer, Sergei Cimliakov.

A2.6 Panchromatic Film and Sfera-S

One of the critical areas of concern to XYZ was the availability of a suitable panchromatic material. The DuPont material that Zebra Imaging was using was not available to XYZ or Geola and anyhow there were doubts as to its sensitivity to nanosecond pulses. Eric worked with many people to try and produce a film product. However, only one produced a real commercial product that XYZ could really use for mass production. This was a silver halide emulsion developed by Geola's partner, Yuri Sazonov. Yuri was the ex-Micron Director at Slavich and had worked closely with Geola for some years. Geola had after all been appointed the international Slavich sales office some years before. Yuri formed a Russian company, Sfera-S AO, to commercialise his new material. He then signed an exclusive agreement for the distribution of this product through Geola. Eric decided to order a holographic plate plant through Geola and to license this new silver halide product so that XYZ could have an in-house facility for a suitable photosensitive material. Eric also travelled to Moscow with David to meet Yuri at Sfera-S (Figure A2.8). Later, Yuri and Stanislovas installed the equipment in Montreal and Eric hired chemists to start producing holographic plates for his digital printers. But despite now having a facility to make large holographic plates suitable for pulsed colour holograms, XYZ could not make a film product. For this, a tripartite agreement between Geola, XYZ and Sfera-S was signed.

XYZ's first printers (Figures A2.9) to be installed outside Montreal were in Paris. By 1995, investors were starting to lose patience with the business and there was increasing pressure on Eric to start commercial sales. The company had started to win prestigious awards for its full-colour digital holograms and, spurred on by this success, XYZ decided to accept the first commercial orders it received. In retrospect, this was premature and the company rushed into this process. However, new US investors had recently come on board and were not in the mood to hang around. The company probably would have been shut down by these investors had another decision been taken. Unfortunately, these first printers seemed to have had various flaws, which stopped them from functioning perfectly. XYZ was called on to make many service visits to Paris and eventually this became extremely expensive. At about this time, it became clear that XYZ would have to slow down its programme for printer sales and as such, smaller



FIGURE A2.8 Visit of groups from Geola and XYZ to Slavich and Sfera-S near Moscow in 2003. From left to right, Yuri Sazonov (director of Sfera-S), Eric Bosco (president of XYZ), Olga Gradova (Sfera-S), Lynne Hryniw (XYZ), Stanislovas Zacharovas (director of Geola UAB) and Aleksej Rodin (scientific manager, Geola Technologies Ltd.).



FIGURE A2.9 First commercial RGB pulsed laser digital holographic printer produced by XYZ (2005)—examples were installed in Montreal, Paris and Vilnius.

quantities of film would be required from Sfera-S. Unfortunately, Sfera-S had just scaled up operations for film production to meet its commitments under the tripartite film agreement it had signed with Geola and XYZ. When Eric announced that XYZ would have to renegotiate the contract, there was therefore a sizeable problem. Geola Technologies had also scaled up operations for XYZ in the United Kingdom, and it too now had problems with overhead expenses. Eric, David, Yuri and Stanislovas met up in Sussex, UK, to try and find a solution.

A2.7 Vilnius Digital Printer

The solution the four men came up with was that a new XYZ printer would be made available to Geola Technologies Ltd. This printer would be paid for by a percentage discount of film supplied by Sfera-S though Geola UAB. In return, a new film supply contract would be made showing revised minimum yearly film-order quantities. The printer would then be installed in Lithuania on loan to a new company, 3D Print UAB, which would be owned jointly by Sfera-S and Geola Technologies Ltd. To compensate the Russian side of the business, Sfera-S would commercialise the digital one-step holograms in Russia. Yuri believed that this would represent a sizeable market and so could compensate his losses caused by XYZ's change in orders. David too believed that the XYZ holograms should sell well in the United Kingdom. Geola Technologies therefore formed a subsidiary under the name of Power Imaging Ltd. to commercialise the holograms made by 3D print UAB (Figures A2.10 and A2.11).

This sounded like a good solution. XYZ had a spare printer available. Contracts were exchanged and the printer was sent to the United Kingdom and then on to Vilnius. However, problems arose almost immediately after the XYZ team had installed the printer. Like the problems encountered in Paris, the printer seemed to have various faults. Geola's chief engineer, Ramunas Bakanas (Figure A2.12), immediately contacted XYZ, but the problems could not be resolved easily. Orders started to come in from Power Imaging Ltd., but they could not be fulfilled. XYZ printed some of the holograms—some occasionally were of good enough quality from Vilnius. Slowly, however, the business became poisoned. Clients became disenchanted and ended up disappearing. To make matters worse, a new manager was appointed over Eric. Geola asked XYZ for circuit and mechanical diagrams as they felt they could repair the problems, but XYZ, under the new leadership, was adamant—Geola could only carry out service



FIGURE A2.10 Marcin Lesniewski and David Brotherton-Ratcliffe at Power Imaging Ltd. in 2005 with an XYZ hologram.

procedures according to the instructions of XYZ. This process went on for some months until exasperation amongst Geola's engineers reached boiling point. Ramunas and his team had worked nights and weekends for months with continual pressure for results. They started to understand what was really wrong with the printer but were not allowed to implement their conclusions. Finally, acting CEO, Bill Meder, wrote to Geola, giving the company permission to fix the printer as it thought necessary. Because XYZ would still give no details of its electronics or software for fear of intellectual property loss, Geola was basically forced to replace system after system with its own devices until finally the company had



FIGURE A2.11 Ardie Osanlou, Stanislovas Zacharovas, Hans Bjelkhagen and David Brotherton-Ratcliffe at the 2006 International Symposium on Display Holography in Wales, where some of the first holograms produced by Power Imaging Ltd. were exhibited.

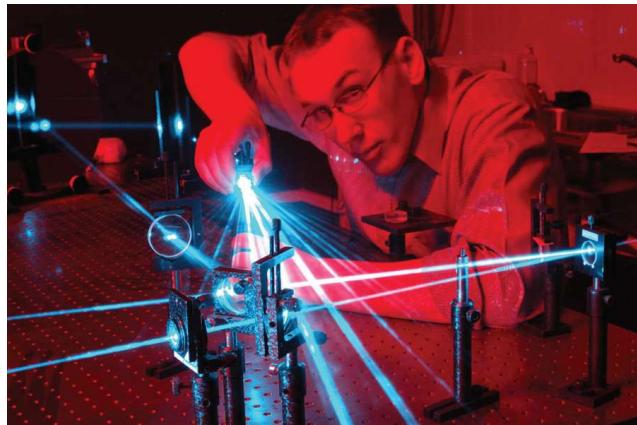


FIGURE A2.12 Ramunas Baranas, chief laser engineer and technical director of Geola Digital UAB in the lab at Geola's Vilnius facilities in 2010.

changed a large part of the printer. At this stage, roughly a year after installation, the printer worked well and predictably. It was too late for Power Imaging Ltd., but nevertheless, the Vilnius printer was now printing and selling digital holograms on a regular basis. In retrospect, despite the problems that appeared, XYZ's first printer was basically well designed. Had the company had access to just a little extra time, there is little doubt that these flaws could have been detected before proper commercial sales commenced. Geola had ended up changing most of the printer systems, not because they were intrinsically bad, but because of the lack of information available during the crisis created by the technical issues in Paris and Vilnius. David and Stanislovas precipitated in the ISDH 2006 Welsh conference, chaired by Hans Bjelkhagen, to promote the large colour holograms produced by Geola (Figure A2.11). The Geola chief laser engineer and technical director Ramunas Baranas is responsible for the lasers produced in the Vilnius lab (Figure A2.12).

A2.8 Sale of XYZ

The French company that had purchased the two digital printers from XYZ was not so lucky as Geola. They did not have engineers capable of replacing the critical printer systems with their own. Instead, the owners decided to bring the case to court in Montreal. This was the final straw for the institutional shareholders backing XYZ. A plan was made to sell the company and despite ostensibly valid inside offers from various shareholders pursuant to the shareholders' agreement, the company was sold to another Canadian company.

The new owners held XYZ for only a short period, after which it was sold and its name was changed to RabbitHoles Media Inc. The new director contacted Geola and asked to meet with David and Yuri in Vilnius to discuss collaboration several months after XYZ's acquisition. Both David and Yuri travelled to Vilnius to meet with him to see if they could work together. RabbitHoles had inherited the old XYZ contracts and was in serious default of some with Sfera-S and Geola UAB. Geola wished to normalise these defaults and was willing to wipe out all debts owed to them. Sadly, however, no agreement could be found. Basically, Geola had come away from the breakup with XYZ with the ability to print and sell one-step and two-step digital holograms and this business, for the first time, was going quite well. RabbitHoles seemed to want to position Geola as their distributor—the two groups obviously had rather different ideas. Despite their differences, however, workers within XYZ and Geola still kept in touch. Roman Rus, one of the original optical engineers at Geola who worked on the first generation printer project in Vilnius, became chief engineer at RabbitHoles and over the coming years, produced some wonderful digital holograms using the XYZ printer technology.

Appendix 3: Active Cavity Length Stabilisation in Pulsed Neodymium Lasers

A3.1 Introduction

The application of digital holographic printing currently requires low-repetition rate (10–100 Hz) pulsed nanosecond laser sources having a very narrow linewidth spectral emission in the visible region, a super-Gaussian or TEM_{00} spatial profile and an exceptionally stable pulse-to-pulse energy and frequency stability.

In the context of lamp-pumped lasers, which are today the most common laser source for digital holographic printing, current technology can only provide the required stability properties using the technique of injection seeding. This technique essentially employs a special thermostabilised small cavity (preferably monolithic) SLM TEM_{00} CW laser to provide a seeding signal for the main lamp-pumped master oscillator. For injection seeding to work, the cavity length of the master oscillator must be actively controlled using a piezo-mounted rear mirror and matched to the longitudinal mode of the seeding signal. This matching process is usually performed by a minimisation of the time for a laser pulse to appear after an electro-optic Q-switch in the master oscillator cavity is opened.

Such injection-seeded systems are quite complex and costly as they require a seeding laser, an expensive and complex electro-optical Q-switching system requiring nanosecond switching of multi-KV signals, a piezo feedback system and complex fast feedback and optimisation electronics.

Here, we describe a much simpler technique [1,2] whereby excellent energy and frequency stability of a lamp-pumped pulsed neodymium laser may be obtained. In its simplest form, thermal heating of either the cavity rear mirror mount, the output coupler or an intracavity glass wedge is used to control the cavity length according to a simple algorithm that either maximises output energy or minimises the standard deviation of output energy. In a slightly more sophisticated version, a piezo element is used to introduce a small amplitude harmonic oscillation into the cavity length and the computed cross-correlation between output laser energy and the test signal allows the cavity length to be maintained at an optimal point.

A3.2 Example of Cavity Stabilisation System Using Heated Rear Mirror Holder

We illustrate the active cavity length stabilisation scheme on the 1064 nm Nd:YAG laser oscillator, which forms the green channel of the RGB laser described in Figure 6.24. This laser has been described in Chapter 6 in detail.

Figure A3.1 shows a diagram of the laser oscillator with the stabilisation system. In the simplest version, the rear cavity mirror (M_{1g}) is attached to a 1.5 cm-long aluminium tube (PZTg) that is in turn attached to a mount through an insulating layer. Heaters embedded in the aluminium tube act to control its temperature. This in turn creates a proportional change in the laser cavity length. The temperature of the aluminium tube is monitored by a thermistor sensor that sends an electric current to a small printed circuit board (PCB). The PCB controls the heating of the heaters in such a way that the temperature reported by the thermistor remains at the given set point temperature to an accuracy of $\pm 0.01^\circ\text{C}$. The set point temperature is defined by a computer that communicates the desired temperature via an RS-485 link to the heating system controller.

The cavity length can also be controlled by mounting the rear cavity mirror on a piezo tube. In this case, the computer then communicates via RS-485 the desired mirror displacement to a piezo controller.

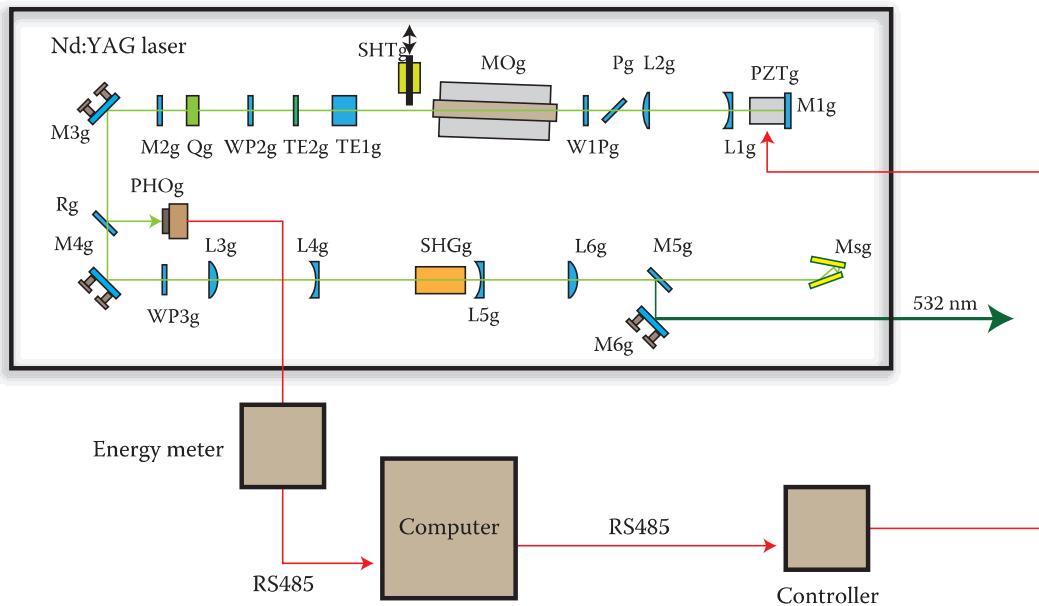


FIGURE A3.1 Schematic diagram of a pulsed Nd:YAG laser oscillator and frequency-doubling optics with active stabilisation scheme. The silicon photodiode PHOg detects the pulse energy of each laser pulse. The signal is processed, calibrated and transferred to a computer in real time by an RS-485 link. The computer then uses an algorithm to control the cavity length through the element PZT_g (a heated mirror mount or a piezo mount) to optimise the lasing stability.

To monitor the energy of each laser pulse, a small uncoated glass wedge (R_g) is used to divert some of the output beam towards a diffuser and a silicon photodiode (PHOg). The signal from this photodiode is used to calculate the energy of the output laser pulse and this information is then transferred to the computer via an RS-485 link.

The laser resonator described should be constructed using a temperature-stabilised super-invar mechanical structure and should be enclosed in a temperature-controlled environment to ensure that the physical cavity length remains as stable as possible. If this is not done, fluctuations in environmental conditions may make it impossible for the active stabilisation system to work. Despite the use of a super-invar mechanical resonator and an active temperature control system, without active stabilisation of the cavity length, the output energy of the laser will inevitably drift from one longitudinal mode to another as the optical path length (rather than the physical length) varies. This is caused by several processes, but the most dominant stems from the fact that the xenon lamp wears out. As this happens, the thermal energy deposited into the Nd:YAG crystal changes, and this causes a change in the cavity optical path.

Lamp wear and other processes cause mode drifting, which leads to bad energy stability at 1064 nm and even worse stability at 532 nm. At 532 nm, in the transition region between stability for one longitudinal mode and another, two modes may oscillate and this causes mode beating and increased frequency conversion efficiency.

In Figure A3.2, we show what happens when the rear mirror mount temperature is changed (thus changing the cavity length)—we plot energy per pulse averaged over 1000 points. Figure A3.3 shows the corresponding graph for the standard deviation of energy calculated over 1000 points.

Clearly, both the energy and standard deviation of energy vary with rear mirror mount temperature. What is interesting is that as one increases the temperature of the rear mirror mount by an amount that gives a thermal expansion equivalent to approximately 1 μm , the graph undergoes a complete cycle. Hence, one can conclude that the high-energy, low standard deviation regions correspond to an optical cavity length that is matched to a given longitudinal mode. As the temperature is increased away from this region, one falls into a region in which one particular mode by itself is not optimised. At the peak standard deviation, one falls into a zone in which the longitudinal mode may change frequently from shot to shot.

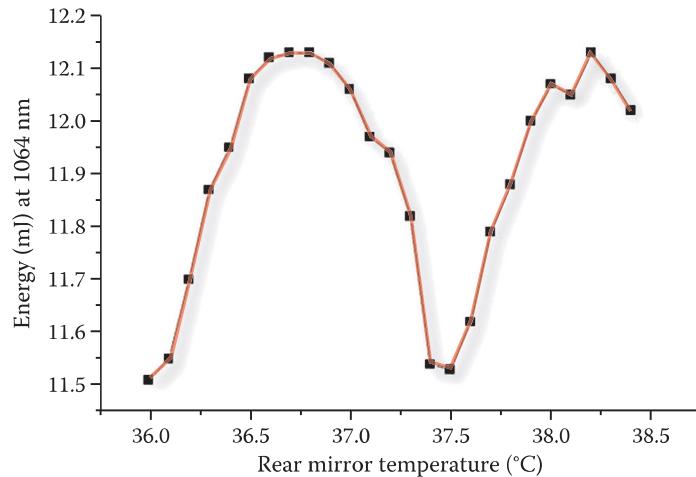


FIGURE A3.2 Graph showing laser output energy at 1064 nm versus rear mirror mount temperature.

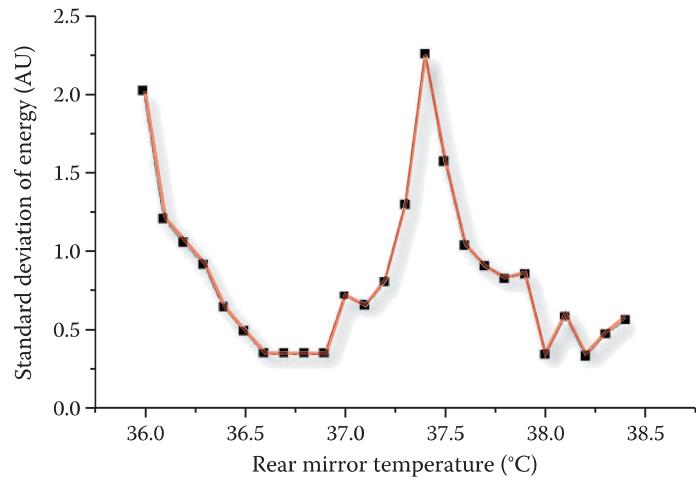


FIGURE A3.3 Graph showing standard deviation of laser output energy at 1064 nm versus rear mirror mount temperature (corresponding to Figure A3.2).

Given that the laser cavity described above exhibits such a clear behaviour concerning energy stability versus rear mirror mount temperature, we use a computer to continually calculate and instruct the thermo-controller to set the optimum temperature of the rear mirror mount such that the averaged energy (more than 1000 points) is a maximum and standard deviation (more than 1000 points) of energy is a minimum. This is the simplest type of algorithm for active cavity stabilisation.

Figure A3.4 shows a typical flow chart of the computer program that we use to iterate the rear mirror mount temperature. Typically, more than 1000 pulses are used for averaging; N_1 is approximately 2000 pulses and N_2 is approximately 1600 pulses.

The function f is usually chosen such that the jump in temperature is smaller as the standard deviation of energy gets smaller. Usually, below a certain standard deviation (typically 0.5%), one wants to put $f = 0$.

Figure A3.4 shows a simple algorithm based on optimisation of the averaged energy.* Figures A3.5 through A3.7 show how the laser responds to this scheme. Figure A3.5 shows a typical plot of output energy at the second harmonic (532 nm) versus time for the case of active stabilisation applied (Zone A)

* One could, of course, optimise the reciprocal of the standard deviation of energy. Alternatively, one could optimise the average energy times a coefficient + the reciprocal of the standard deviation of energy times another coefficient.

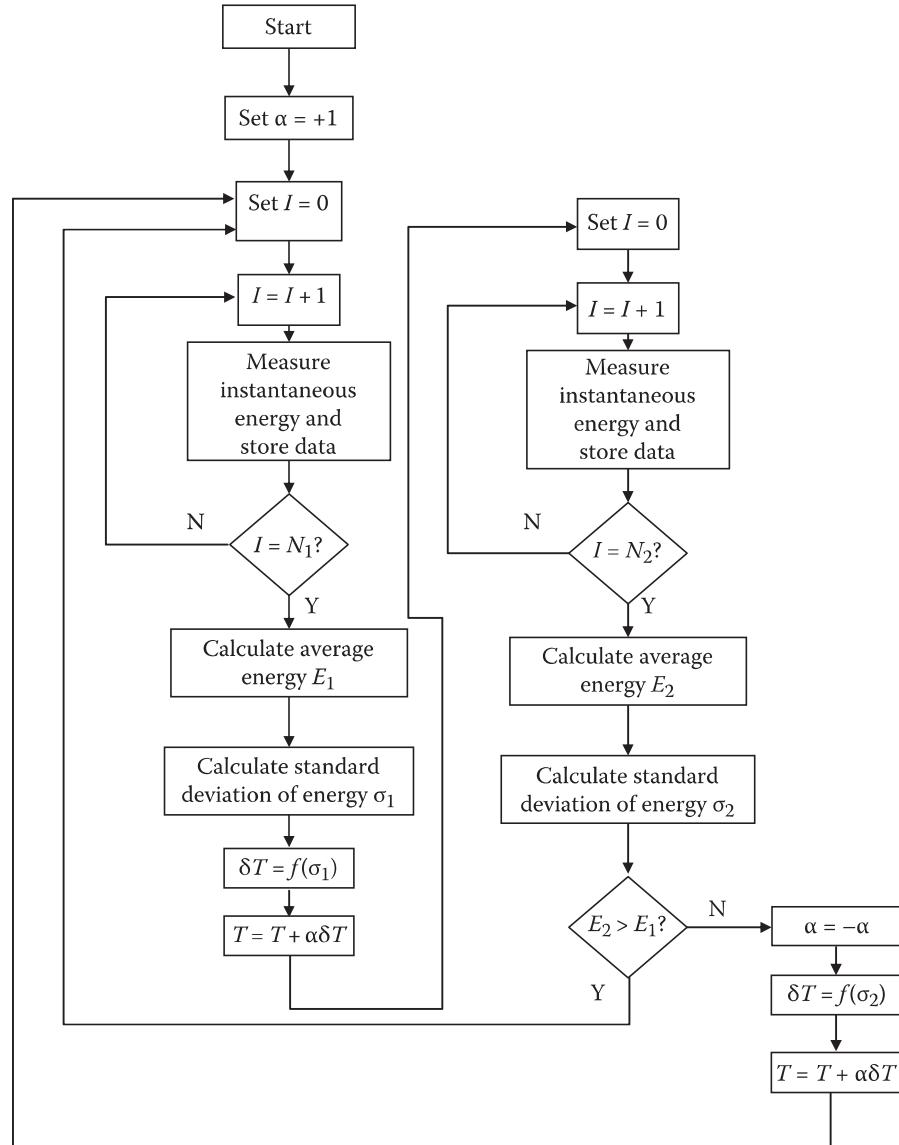


FIGURE A3.4 Flow chart of a simple optimisation algorithm.

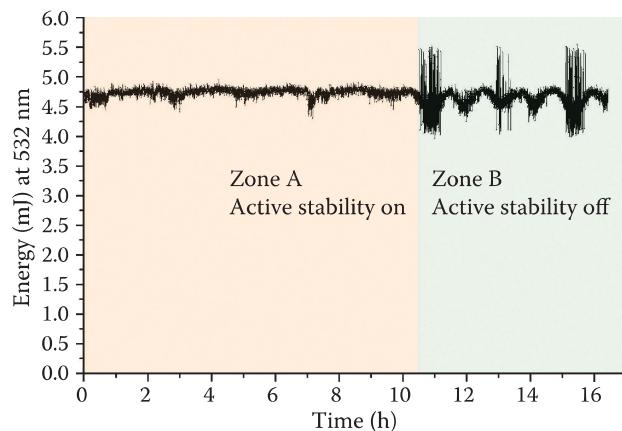


FIGURE A3.5 Typical plot of the output energy at the second harmonic (532 nm) versus time for the case of active stabilisation applied (Zone A) and no active stabilisation applied (Zone B).

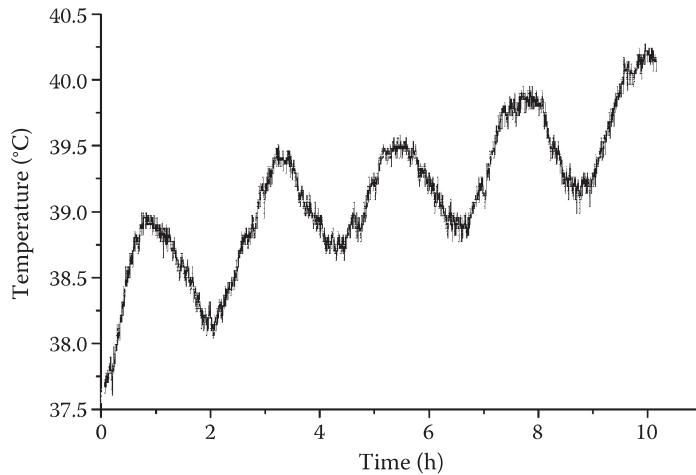


FIGURE A3.6 Plot of temperature of rear mirror mount versus time corresponding to Figure A3.5.

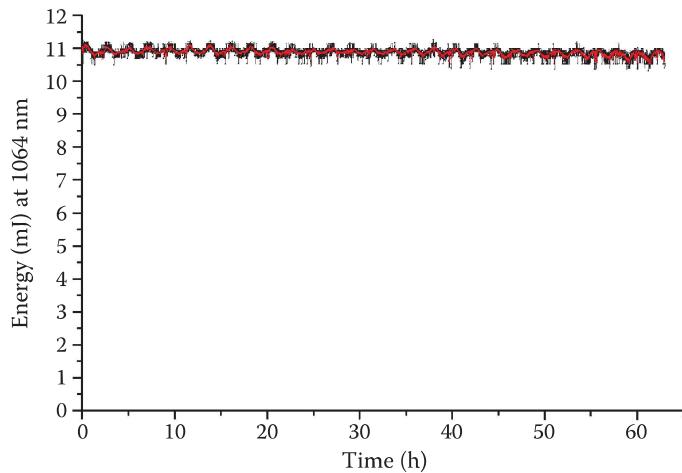


FIGURE A3.7 Typical longer duration plot of output energy (1064 nm) versus time for the case of active stabilisation applied.

and no active stabilisation applied (Zone B). Figure A3.6 shows a plot of the temperature of the rear mirror mount versus time corresponding to Figure A3.5. Figure A3.7 shows a longer plot of output energy (1064 nm) versus time for the case of the active stabilisation applied. Note that the periodic oscillations in temperature that can be observed in all the traces presented (Figures A3.5 through A3.7; period = approximately 2 h) are due to ambient temperature cycling.

Clearly, active stabilisation of the rear mirror mount temperature produces a long-term laser emission of greatly increased energy stability. This is also true of the emission at the second harmonic, where $\pm 4\%$ peak-to-peak energy stabilities and less than $\pm 1\%$ RMS energy stabilities have routinely been obtained over several days. Because the stable energy regime is produced by optimising conditions for a single unique longitudinal mode to resonate, inevitably one also observes a marked improvement in frequency stability. Such stabilities are vital in the applications of writing dot-matrix holographic optical elements, holographic screens and digital holograms.

A3.3 Active Cavity Length Stabilisation by Piezo Element

A piezo element may be used (Figure A3.8) to alter the cavity length of the laser. The principal advantage of this approach over the heated element approach is the faster time response. This is particularly

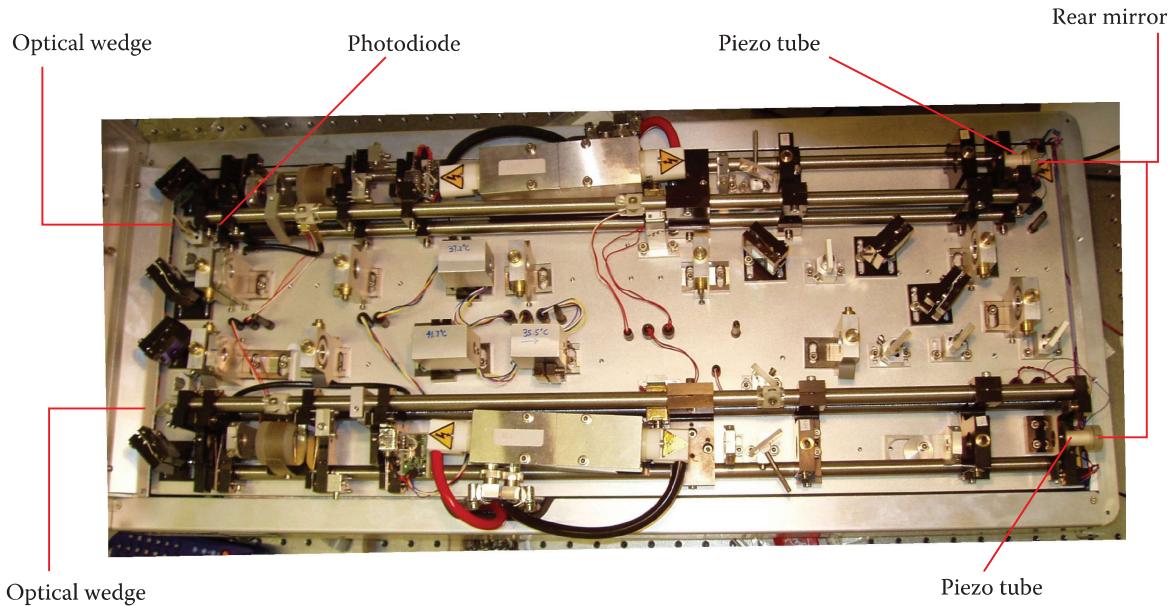


FIGURE A3.8 Photograph of a lamp-pumped pulsed RGB Nd:YAG holography laser with piezo-mounted rear mirrors on both the 1064 nm cavity (upper) and the 1319 nm cavity (lower). Laser manufactured by Geola Technologies Ltd.

important at $1.3\text{ }\mu\text{m}$, as here one needs greater displacements due to the greater wavelength. These greater displacements must be produced either by a greater temperature swing or by lengthening the temperature-sensitive aluminium tube. Greater temperature swings can produce problems themselves and greater tube lengths lead to a poorer temporal response. If the laser system is not of a certain fundamental stability, it may then occur that the heated element approach cannot keep up with the mode drifts and the system is unable to “lock”.

A piezo system, albeit a little more complex, can therefore achieve a better stability. It can also cope with worse environmental conditions (i.e., fluctuations in laboratory temperature) than the simpler heated element approach.

A3.3.1 Statistical Optimisation Algorithm

The fact that the piezo element technique has a much better temporal response allows one to use a statistical method to optimise the cavity length. Specifically, one applies a small test signal to the piezo element:

$$V = V_0 + V_1 \sin \omega t \quad (\text{A3.1})$$

Here, V_0 determines an appropriate midpoint cavity length and $V_1 \ll V_0$. Typically, ω must be chosen to be significantly smaller than the laser repetition rate but large enough to ensure many cycles over the desired cavity length updating time, T . For a 30 Hz laser, T should typically be somewhat more than 10 s.

Every T seconds, the voltage to the piezo element is updated according to the following rule:

$$V^{(n+1)} = V^{(n)} + \alpha \varphi \quad (\text{A3.2})$$

where α is a small stepping constant and φ is an error signal. If α is too small, the iteration will be very stable, but convergence will be slow. On the other hand, if α is too large, the iteration may go unstable.

The error signal is calculated as

$$\varphi = \frac{1}{T} \int_{nT}^{(n+1)T} E(t) V_1 \sin \omega t dt \quad (\text{A3.3})$$

which may be approximated to

$$\delta\phi \sim \frac{V_1 \delta T}{T} \sum_{i=i_0}^{i_f} E_i \sin \omega t_i \quad (\text{A3.4})$$

Here, $E(t)$ is the pulse energy of the laser and E_i is its measurement at the i th pulse. The parameter δT is the time between laser pulses. The sum in Equation A3.4 is taken over every laser pulse between updating intervals. This simple algorithm is very effective. By using an error signal that is based on a first-order cross-correlation between the test signal and the laser energy, spurious effects not induced by the test signal are eliminated. In this way, a very clear error signal is attained, resulting in quick cavity locking.

A3.4 Extension to Other Lasers

Either of the two stabilisation systems we have described above can be easily applied to ring cavity laser oscillator designs. Here an intracavity heated glass wedge may be used instead of a heated rear mirror mount which is usually used for the linear oscillator. The optical path length of the laser cavity is then changed actively by varying the temperature of the wedge. Active cavity length stabilisation can be used effectively on lasers oscillating at many of the 1 μm and 1.3 μm lines of neodymium lines of Nd:YAG. As such, the technique is extremely useful for stabilisation of both channels (1064 nm and 1319 nm) of the RGB-pulsed laser of Figure 6.24.

REFERENCES

1. D. Brotherton-Ratcliffe, *Laser*, US Patent 7,852,887 (filed 2005, granted 2010).
2. C. Cottverte, J.-C. Joseph, J. Kosenro and R. Rus, *Laser*, US Patent 7,596,153 (filed 2005, 2009).

Appendix 4: Aberration Correction by Image Predistortion in Digital Holograms

A4.1 Introduction

In Chapter 11, we derived formulae relating how the light rays recorded in a hologram replay when the hologram is illuminated. We saw that when the holographic emulsion swells or shrinks, when the refractive index changes during processing, or when the illumination geometry and wavelength does not exactly match the recording geometry and wavelength, then aberration in the resulting image is present. This aberration can take the form of chromatic aberration or geometric aberration (or both).

In digital holography, it is possible, to some extent at least, to correct for such aberration by predistorting the digital data. In this appendix, we shall give a very brief outline of how exactly this may be done in one specific and very simple case. In particular, we shall consider the case of a slight emulsion shrinkage on chemical processing in the three-colour reflection hologram and how this may be corrected by predistortion and chromatic rebalancing of the digital data.

A4.2 Mathematical Model

In almost all cases in reflection display holography, the Bragg condition provides a good description of how the hologram replays. According to our discussions in Chapter 11, the grating vector of a recorded hologram is given by

$$\mathbf{G} = \mathbf{k}_r - \mathbf{k}_o \quad (\text{A4.1})$$

Here, the \mathbf{k} parameters are the wave vectors of the rays incident at a given hogel of the hologram being recorded. The subscripts “ o ” and “ r ” stand for the object and reference ray. When there is no change to the emulsion thickness, we can expect that the hologram will replay most efficiently when

$$\mathbf{k}_c - \mathbf{k}_i = \mathbf{G} \quad (\text{A4.2})$$

Here, the subscripts “ c ” and “ i ” denote the reference illumination ray and the image ray at replay. When there is a change in the thickness of the emulsion, the grating vector is changed

$$\mathbf{G} \rightarrow \underline{\underline{\alpha}} \cdot \mathbf{G} \quad (\text{A4.3})$$

with the emulsion deformation matrix usually given by the simple relation

$$\underline{\underline{\alpha}} = \begin{pmatrix} \tau^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{A4.4})$$

where τ is the swelling constant with $\tau < 1$ indicating shrinkage and $\tau > 1$ indicating swelling.

The hologram can then be expected to replay according to the formula

$$\mathbf{k}_c - \mathbf{k}_i = \underline{\underline{\alpha}} \cdot \mathbf{G} \quad (\text{A4.5})$$

We can therefore calculate that a hogel recorded with an object wave vector, \mathbf{k}_o , and a reference wave vector, \mathbf{k}_r , and then suffering an emulsion deformation given by $\underline{\alpha}$ will replay most efficiently with a wave vector, \mathbf{k}_i , when illuminated by a wave vector, \mathbf{k}_c , where

$$\mathbf{k}_i = \mathbf{k}_c - \underline{\alpha} \cdot \mathbf{G} = \mathbf{k}_c - \underline{\alpha} \cdot (\mathbf{k}_r - \mathbf{k}_o) \quad (\text{A4.6})$$

The optimal free-space wavelength of replay, λ_c , will then be given simply by taking the magnitude of this expression:

$$\frac{2\pi n_c}{\lambda_c} = |\mathbf{k}_c - \underline{\alpha} \cdot (\mathbf{k}_r - \mathbf{k}_o)| \quad (\text{A4.7})$$

where n_c is the refractive index of the emulsion after processing, which, in general, will be different from that before processing, n_r .

A4.3 Calculation of Optimal Reference Replay Angle

When the emulsion shrinks (or swells) on processing, the hologram may be illuminated from different angles at different wavelengths. However, there will be an optimal illumination angle at which each hogel is associated with an axially projecting image bundle. Basically, we don't want to illuminate our hologram in such a way that the image is directed downwards or upwards. Rather, we wish the image to project straight out in front.

We will limit our discussion here to considering the case of a digital reflection hologram recorded and replayed using a collimated reference beam. This has two important effects. First, every hogel is the same and, second, we only need to consider angles in the vertical plane. We shall also only consider the case of a paraxial writing objective.

We start by considering the central pixel of the printer spatial light modulator (SLM) (corresponding to one primary colour channel in a triple beam DWDH printer). The k -vector associated with this pixel, \mathbf{k}_o^o is known: the unit vector is given by simply subtracting the Cartesian world coordinates of the pixel location on the projected SLM plane from the hogel coordinates and dividing by the length of this vector—then the actual k -vector can be calculated by multiplying the unit vector by $2\pi n/\lambda_r$. This vector can then be written as (see Figure A4.1)

$$\mathbf{k}_o^o = \frac{2\pi n_r}{\lambda_r} \hat{\mathbf{x}} \quad (\text{A4.8})$$

If we wish the hogel to project axially, then $\hat{\mathbf{k}}_i^o = \hat{\mathbf{k}}_o^o$ and this means that

$$\mathbf{k}_i^o = \frac{2\pi n_c}{\lambda_c} \hat{\mathbf{x}} \quad (\text{A4.9})$$

If we now take

$$\begin{aligned} \mathbf{k}_r &= -\frac{2\pi n_r}{\lambda_r} \{ \cos \theta_r \hat{\mathbf{x}} + \sin \theta_r \hat{\mathbf{y}} \} \\ \mathbf{k}_c &= -\frac{2\pi n_c}{\lambda_c} \{ \cos \theta_c \hat{\mathbf{x}} + \sin \theta_c \hat{\mathbf{y}} \} \end{aligned} \quad (\text{A4.10})$$

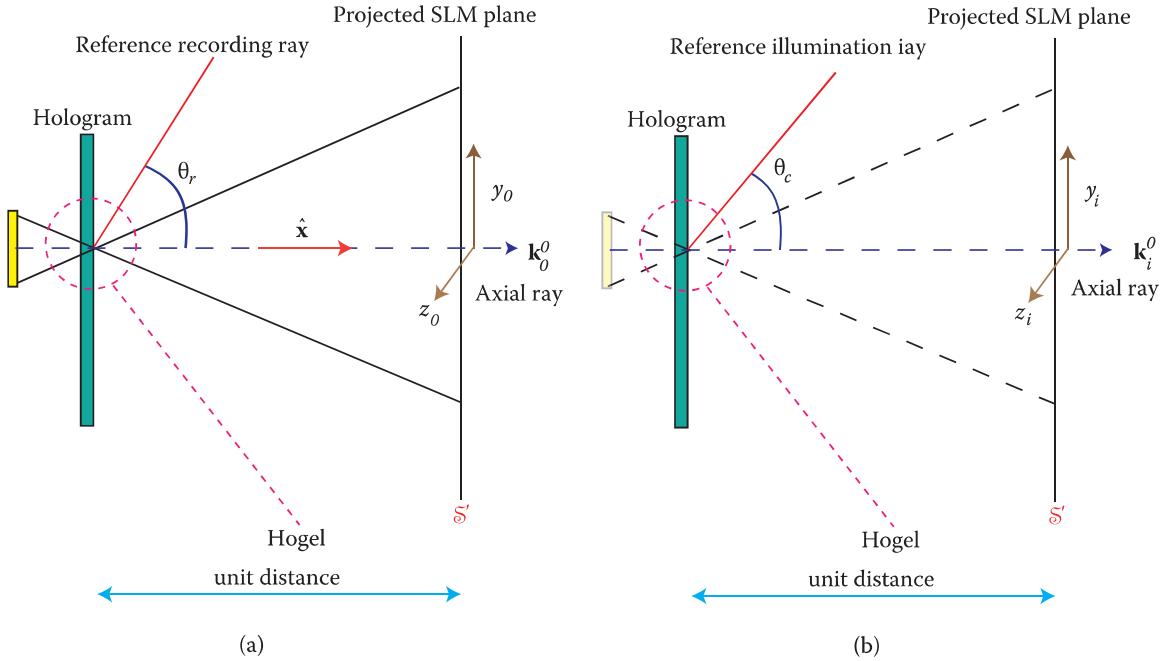


FIGURE A4.1 (a) Recording and (b) playback of the reflection hologram considered in the text. The reference illumination angle is chosen so that the reconstructed image of the axial object ray is axial itself.

the x and y components of Equation A4.6 then tell us that

$$\begin{aligned} \frac{2\pi n_c}{\lambda_c} (1 + \cos \theta_c) &= \frac{2\pi n_r}{\lambda_r} \tau^{-1} (1 + \cos \theta_r) \\ \frac{2\pi n_c}{\lambda_c} \sin \theta_c &= \frac{2\pi n_r}{\lambda_r} \sin \theta_r \end{aligned} \quad (\text{A4.11})$$

Note that these are internal angles. These equations may be solved to give:

$$\begin{aligned} \lambda_c &= 2\lambda_r \frac{n_c}{n_r \tau} \left\{ \frac{1}{(\tau^{-2} + 1) + (\tau^{-2} - 1)\cos \theta_r} \right\} \\ \cos \theta_c &= \frac{(\tau^{-2} - 1) + (\tau^{-2} + 1)\cos \theta_r}{(\tau^{-2} + 1) + (\tau^{-2} - 1)\cos \theta_r} \end{aligned} \quad (\text{A4.12})$$

In terms of external angles they become

$$\begin{aligned} \lambda_c &= \frac{2\lambda_r n_c}{n_r \tau \left[\left\{ 1 + \tau^{-2} \right\} + \left\{ \tau^{-2} - 1 \right\} \sqrt{1 - \frac{\sin^2 \theta_r}{n_r^2}} \right]} \\ \sin \theta &= \frac{2n_c \tau \sin \theta_r}{n_r \left\{ 1 + \sqrt{1 - \frac{\sin^2 \theta_r}{n_r^2}} + \tau^2 \left(1 - \sqrt{1 - \frac{\sin^2 \theta_r}{n_r^2}} \right) \right\}} \end{aligned} \quad (\text{A4.13})$$

This then determines the optimal replay illumination angle for the hologram. Note that this angle of illumination is the same for each primary colour. One should also observe that the optimal replay wavelength for

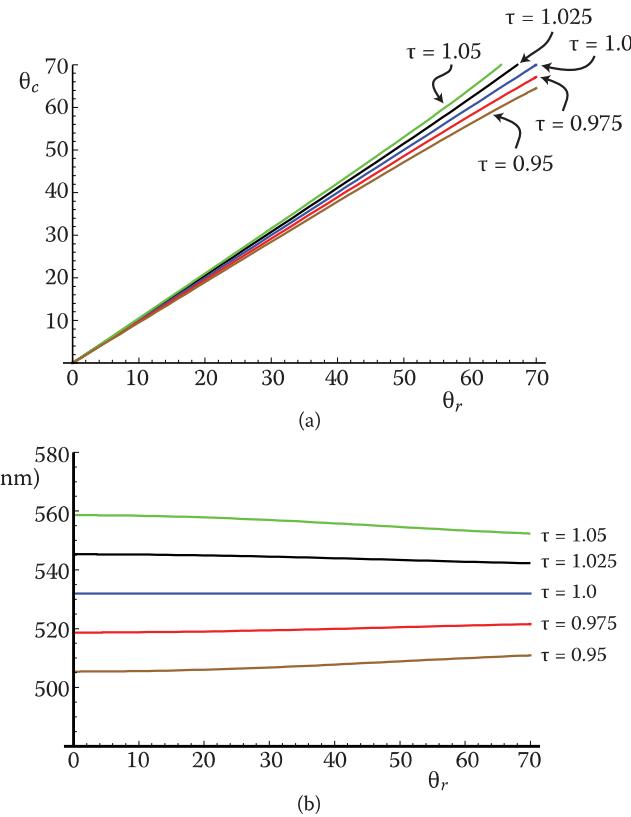


FIGURE A4.2 Optimal replay (assuring an axial projection of the reconstructed image of the axial object ray) of a reflection hologram which has suffered emulsion deformation. (a) Graph of the optimal (external) illumination angle of incidence, θ_c in degrees as defined by Equation 4.13 versus (external) reference recording angle of incidence, θ_r for various values of the emulsion swelling parameter, τ . The refractive indices, n_r and n_c before and after chemical processing of the reflection hologram have both been set to 1.5. The recording wavelength is 532 nm. (b) Corresponding graphs of replay wavelength of the axial ray versus (external) reference recording angle. Note that for more than around 5% of shrinkage or swelling the magnitude of the induced chromatic aberration can make it difficult to rebalance the colours successfully.

the axial ray (given by Equation A4.12) is now different from the recording wavelength (see Figure A4.2). We can anticipate that this optimal replay wavelength will be different for each and every ray in the hogel.

A4.4 Compensation for Geometrical Distortion

Each object ray, characterised by its wave vector, \mathbf{k}_o , will replay under the optimal illumination angle (Equation A4.13) to produce an image ray defined by

$$\mathbf{k}_i = \mathbf{k}_c - \underline{\alpha} \cdot (\mathbf{k}_r - \mathbf{k}_o) \quad (\text{A4.14})$$

We can write the components of this equation as follows:

$$\frac{2\pi n_c}{\lambda_c \sqrt{1+y_i^2+z_i^2}} \begin{pmatrix} 1 \\ y_i \\ z_i \end{pmatrix} = -\frac{2\pi n_c}{\lambda_c} \begin{pmatrix} \cos \theta_c \\ \sin \theta_c \\ 0 \end{pmatrix} + \frac{2\pi n_r}{\lambda_r} \begin{pmatrix} \tau^{-1} \cos \theta_r \\ \sin \theta_r \\ 0 \end{pmatrix} + \frac{2\pi n_r}{\lambda_r \sqrt{\tau^{-2}+y_o^2+z_o^2}} \begin{pmatrix} \tau^{-1} \\ y_o \\ z_o \end{pmatrix} \quad (\text{A4.15})$$

Note that we have used the k -vectors of Equation A4.10. We have also used a unit x distance from the hogel to the projected SLM plane for the calculation of the object and image k -vector. Note finally that angles are internal!

The three equations (Equation A4.15) determine the intersection of image rays with the unit x -distance projection of the SLM, (y_i, z_i) and the optimal replay wavelength, λ_c of each image ray in terms of the corresponding intersection of object rays (y_o, z_o) and the recording wavelength, λ_r . They may be written in the simplified form

$$\frac{1}{\sqrt{1+y_i^2+z_i^2}} \begin{pmatrix} 1 \\ y_i \\ z_i \end{pmatrix} = \varepsilon \left\{ \begin{pmatrix} \tau^{-1} \cos \theta_r \\ \sin \theta_r \\ 0 \end{pmatrix} + \frac{1}{\sqrt{\tau^{-2}+y_o^2+z_o^2}} \begin{pmatrix} \tau^{-1} \\ y_o \\ z_o \end{pmatrix} \right\} - \begin{pmatrix} \cos \theta_c \\ \sin \theta_c \\ 0 \end{pmatrix} \quad (\text{A4.16})$$

where

$$\varepsilon = \frac{\lambda_c n_r}{\lambda_r n_c} \quad (\text{A4.17})$$

If we simplify this further, we can write

$$\begin{aligned} 1 &= (1+y_i^2+z_i^2)(a_x \varepsilon + b_x)^2 \\ y_i^2 &= (1+y_i^2+z_i^2)(a_y \varepsilon + b_y)^2 \\ z_i^2 &= (1+y_i^2+z_i^2)a_z^2 \varepsilon^2 \end{aligned} \quad (\text{A4.18})$$

Adding all three equations together now gives

$$(a_x \varepsilon + b_x)^2 + (a_y \varepsilon + b_y)^2 + a_z^2 \varepsilon^2 = 0 \quad (\text{A4.19})$$

or

$$A\varepsilon^2 + B\varepsilon + C = 0 \quad (\text{A4.20})$$

with coefficients

$$\begin{aligned} A &= a_x^2 + a_y^2 + a_z^2 \\ B &= 2(a_y b_y + a_x b_x) \\ C &= b_x^2 + b_y^2 \end{aligned} \quad (\text{A4.21})$$

This may then be solved by the standard formula

$$\varepsilon = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \quad (\text{A4.22})$$

Practically, there will only be one sensible root to this equation, which will be the nearest one to real unity.

With ε known, we know the replay wavelength of each ray. This will be vital for chromatic rebalancing, which we will treat in the next section. However, now that we know ε , we can easily solve Equation A4.18 to define (y_i, z_i) :

$$\begin{aligned} y_i &= \frac{a_y \varepsilon + b_y}{a_x \varepsilon + b_x} \\ z_i &= \frac{a_z \varepsilon}{a_x \varepsilon + b_x} \end{aligned} \quad (\text{A4.23})$$

Equation A4.23 can be used to formulate a mapping of the SLM data to correct for induced aberration. We start this process by dividing up the SLM into two coordinate systems, (α, β) and (α', β') . We defined

the (α, β) system in Chapter 8. It basically just counts the pixels on the SLM. In Chapter 8, we derived I-to-S transformations, which defined the data $\mathbf{S}_{\alpha\beta}$ at the SLM pixel (α, β) in terms of the raw camera or image data. We can therefore imagine the (α, β) coordinate system corresponding to the (y_o, z_o) coordinate system. Specifically, we can write, following the results of Chapter 8*,

$$\begin{aligned} z_o &= \left\{ \frac{2(\alpha - 1)}{(N_A - 1)} - 1 \right\} \tan\left(\frac{\Psi_{\text{PH}}}{2}\right) \\ y_o &= - \left\{ \frac{2(\beta - 1)}{(N_B - 1)} - 1 \right\} \tan\left(\frac{\Psi_{\text{PV}}}{2}\right) \end{aligned} \quad (\text{A4.24})$$

where Ψ_{PH} and Ψ_{PV} are the printer field of views (horizontal and vertical) as apodised by the SLM.

If we write data, $\mathbf{S}_{\alpha\beta}$ to the SLM at the pixel (α, β) corresponding to the coordinates (y_o, z_o) on the projected SLM plane, \mathcal{S}' at unit distance from the hologram then we know that on replay the ray will effectively move to a new position (y_i, z_i) on \mathcal{S}' . We can now define the system (α', β') to define the effective pixel position this image ray corresponds to

$$\begin{aligned} z_i &= \left\{ \frac{2(\alpha' - 1)}{(N_A - 1)} - 1 \right\} \tan\left(\frac{\Psi_{\text{PH}}}{2}\right) \\ y_i &= - \left\{ \frac{2(\beta' - 1)}{(N_B - 1)} - 1 \right\} \tan\left(\frac{\Psi_{\text{PV}}}{2}\right) \end{aligned} \quad (\text{A4.25})$$

So what we actually want to do is not to write the data $\mathbf{S}_{\alpha\beta}$ to the pixel (α, β) . Rather, we want to write the data $\mathbf{S}_{\alpha'\beta'}$ to the pixel (α, β) . This way, on replay, the data $\mathbf{S}_{\alpha'\beta'}$ ends up in the right place. Another way of saying this is that we should redefine the SLM data so that

$$\mathbf{S}'_{\alpha\beta} = \mathbf{S}_{\alpha'\beta'} \quad (\text{A4.26})$$

where

$$\begin{aligned} \alpha' &= 1 + \left(\frac{N_A - 1}{2} \right) \left\{ 1 + z_i(\alpha, \beta) \cot\left(\frac{\Psi_{\text{PH}}}{2}\right) \right\} \\ \beta' &= 1 + \left(\frac{N_B - 1}{2} \right) \left\{ 1 - y_i(\alpha, \beta) \cot\left(\frac{\Psi_{\text{PV}}}{2}\right) \right\} \end{aligned} \quad (\text{A4.27})$$

In this way, we end up writing an image data byte to a different pixel location on the SLM than we would have done if there were no emulsion thickness change. However, this shift in position is just what is required such that the distortion induced by the emulsion thickness change effectively moves the ray back to the position where it should have been.

The index transformation (Equation A4.27) may be calculated just one time and then applied to all hogels for a given primary colour channel. Typically, bilinear or bicubic interpolation is used (see Appendix 7). Note that the index transformation depends on illumination wavelength, so there will be different transformations necessary for each primary colour.

A4.5 Compensation for Chromatic Aberration

Our reflection hologram is assumed to be illuminated by a broadband white-light source. As such, each ray will be associated with an optimal replay wavelength that is, in general, different from λ_r . This

* Note that we are using a non-conjugate SLM position here as per Figure A4.1.

optimal wavelength is determined by the parameter, ε , in the ray equation (Equation A4.22). For the axial ray of each hogel

$$\varepsilon = \frac{\lambda_c n_r}{\lambda_r n_c} = \frac{2\tau}{1 + \tau^2 + (\tau^2 - 1)\cos \theta_r} \quad (\text{A4.28})$$

If we consider the case of a three colour reflection hologram λ_r will be replaced by the three laser wavelengths used to record the hologram-namely λ_R , λ_G , and λ_B .

With no emulsion change, the tristimulus values associated with a given object/image ray from a given hogel are given by

$$\begin{aligned} X &= k \left\{ \mathbf{S}_G \bar{x}(\lambda_G) + \mathbf{S}_R \bar{x}(\lambda_R) + \mathbf{S}_B \bar{x}(\lambda_B) \right\} \\ Y &= k \left\{ \mathbf{S}_G \bar{y}(\lambda_G) + \mathbf{S}_R \bar{y}(\lambda_R) + \mathbf{S}_B \bar{y}(\lambda_B) \right\} \\ Z &= k \left\{ \mathbf{S}_G \bar{z}(\lambda_G) + \mathbf{S}_R \bar{z}(\lambda_R) + \mathbf{S}_B \bar{z}(\lambda_B) \right\} \end{aligned} \quad (\text{A4.29})$$

Here, \bar{x} , \bar{y} and \bar{z} are the colour-matching functions of the CIE Standard Colorimetric Observer (see, for example, Giorgianni and Madden [1]) and k is a normalising factor. The parameters \mathbf{S}_G , \mathbf{S}_R and \mathbf{S}_B are, respectively, the green, red and blue brightness data written to the three primary-colour SLMs for the case of zero emulsion shrinkage.

When the primary wavelengths change according to Equation A4.22, the tristimulus values will also change:

$$\begin{aligned} X' &= k \left\{ \mathbf{S}_G \bar{x}(\lambda'_G) + \mathbf{S}_R \bar{x}(\lambda'_R) + \mathbf{S}_B \bar{x}(\lambda'_B) \right\} \\ Y' &= k \left\{ \mathbf{S}_G \bar{y}(\lambda'_G) + \mathbf{S}_R \bar{y}(\lambda'_R) + \mathbf{S}_B \bar{y}(\lambda'_B) \right\} \\ Z' &= k \left\{ \mathbf{S}_G \bar{z}(\lambda'_G) + \mathbf{S}_R \bar{z}(\lambda'_R) + \mathbf{S}_B \bar{z}(\lambda'_B) \right\} \end{aligned} \quad (\text{A4.30})$$

This then describes the chromatic aberration of the ray in question. To ensure that there is zero chromatic aberration, we must change the parameters \mathbf{S}_G , \mathbf{S}_R and \mathbf{S}_B and ensure that the tristimulus values are equal to their primed values. Or in other words,

$$\begin{aligned} k \left\{ \mathbf{S}'_G \bar{x}(\lambda'_G) + \mathbf{S}'_R \bar{x}(\lambda'_R) + \mathbf{S}'_B \bar{x}(\lambda'_B) \right\} &= X \\ k \left\{ \mathbf{S}'_G \bar{y}(\lambda'_G) + \mathbf{S}'_R \bar{y}(\lambda'_R) + \mathbf{S}'_B \bar{y}(\lambda'_B) \right\} &= Y \\ k \left\{ \mathbf{S}'_G \bar{z}(\lambda'_G) + \mathbf{S}'_R \bar{z}(\lambda'_R) + \mathbf{S}'_B \bar{z}(\lambda'_B) \right\} &= Z \end{aligned} \quad (\text{A4.31})$$

where X , Y and Z are given by Equation 4.29.

Equation 4.31 can be written in matrix form:

$$\begin{aligned} \begin{pmatrix} \mathbf{S}'_G \\ \mathbf{S}'_R \\ \mathbf{S}'_B \end{pmatrix} &= \begin{pmatrix} \bar{x}(\lambda'_G) & \bar{x}(\lambda'_R) & \bar{x}(\lambda'_B) \\ \bar{y}(\lambda'_G) & \bar{y}(\lambda'_R) & \bar{y}(\lambda'_B) \\ \bar{z}(\lambda'_G) & \bar{z}(\lambda'_R) & \bar{z}(\lambda'_B) \end{pmatrix}^{-1} \times \begin{pmatrix} \bar{x}(\lambda_G) & \bar{x}(\lambda_R) & \bar{x}(\lambda_B) \\ \bar{y}(\lambda_G) & \bar{y}(\lambda_R) & \bar{y}(\lambda_B) \\ \bar{z}(\lambda_G) & \bar{z}(\lambda_R) & \bar{z}(\lambda_B) \end{pmatrix} \begin{pmatrix} \mathbf{S}_G \\ \mathbf{S}_R \\ \mathbf{S}_B \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \mathbf{S}_G \\ \mathbf{S}_R \\ \mathbf{S}_B \end{pmatrix} \end{aligned} \quad (\text{A4.32})$$

The matrix coefficients a_{ij} need only be calculated for each ray one time, as they are identical (for a particular ray) for all hogels. Application of the transformation (Equation A4.32) to all the SLM brightness data will effectively rebalance the chromatic equation and ensure that each ray from each hogel has the correct tristimulus values.

The above analysis assumes that the hologram is rather thick and that the colour-matching functions of the CIE Standard Colorimetric Observer are only sampled at one exact wavelength. In the case that the hologram is thinner, we can calculate, the exact form of the spectral function for each ray using the results of the PSM theory presented in Chapter 12. In addition, one can include the spectral power distribution of the illumination source. In this case, each of the matrix elements in Equation A4.32 are transformed in a similar fashion to

$$\bar{x}(\lambda_G) \rightarrow \int F(\lambda) \eta_G(\lambda, \theta_c, \psi, \varphi) \bar{x}(\lambda) d\lambda \quad (\text{A4.33})$$

Here, $F(\lambda)$ is the spectral power distribution of the illumination source and η_G is the diffractive efficiency in the green channel as given, for example, by Equation 12.90, with the coefficients in Equation 12.101, in the case of illumination by σ -polarised light.

A4.6 Other Corrections

We have considered here only the case of a simple change in emulsion thickness with different refractive indices on recording and processing. This is really the simplest case and serves to illustrate in the most simple and straightforward manner how geometric and chromatic predistortion works. In practice, however, rather more complicated scenarios arise. In particular, often one wants to write a hologram with a collimated reference beam and then replay it with a spot lamp at a certain distance. The mathematics used above can be generalised rather easily to this situation—all that needs to be done is for the vector \mathbf{k}_c in Equation A4.15 to be written in Cartesian form, thus describing a point source at the desired distance—this simply changes b_x and b_y and introduces a new b_z parameter in Equation A4.18. Note, however, that both the geometric and chromatic predistortions are now different for all rays *and* all hogels.

It should also be underlined that predistortion has its limits. If too great a change in emulsion thickness or index occurs or if too great a disparity in the illumination/recording geometry exists, then it may just not be possible to compensate for the induced aberrations. We should also state that it is possible to compensate for slightly larger chromatic aberration if one relaxes somewhat the condition of an axially propagating image ray bundle. In this case one trades chromatic correction for geometric correction. However this scheme can very quickly reduce the vertical field of view.

REFERENCE

1. E. J. Giorgianni and T. E. Madden, *Digital Colour Management—Encoding Solutions*, Addison-Wesley, Reading, MA (1998).

Appendix 5: MAXScript Holocam Program

The following is a full listing of the MAXScript virtual holocam program described in Chapter 10.

```
rollout DoubleParallaxRender "DoubleParallaxRender" width:687 height:361
(
--control definitions
    button BRen "Start Render" pos:[462,13] width:175 height:36
    edittext RenderDir "Render Directory" pos:[248,266] width:373 height:21
    label hOutput "h (mm)" pos:[386,98] width:108 height:18
    spinner Win "W (mm)" pos:[59,288] width:74 height:16 range:[0,5000,600] type:#integer
    spinner Hin "H (mm)" pos:[59,319] width:74 height:16 range:[10,5000,400] type:#integer
    spinner NA "NA" pos:[229,100] width:74 height:16 range:[100,3000,600] type:#integer
    spinner NB "NB" pos:[230,131] width:74 height:16 range:[100,3000,400] type:#integer
    GroupBox grp1 "Calculated Parameters" pos:[371,74] width:268 height:155
    GroupBox grp2 "Hologram Definition" pos:[195,73] width:163 height:155
    GroupBox grp3 "Window Parameters" pos:[3,265] width:173 height:92
    GroupBox grp4 "SLM Definition" pos:[2,73] width:175 height:83
    spinner NM "NM" pos:[56,99] width:74 height:16 range:[100,5000,1280] type:#integer
    spinner NV "NV" pos:[57,129] width:74 height:16 range:[100,5000,1024] type:#integer
    label NKOutput "NKOutput" pos:[385,118] width:66 height:18
    label NGOOutput "NGOOutput" pos:[386,139] width:74 height:18
    GroupBox grp5 "Printer Optics Definition" pos:[2,173] width:174 height:77
    spinner PsiPH "PsiPH" pos:[56,203] width:72 height:16 range:[10,120,89] \
    type:#integer scale:0.1
    spinner NStopLine "Render: Line Number End" pos:[352,326] width:192 height:16 \
    range:[1,5000,1] type:#integer
    GroupBox grp6 "Job Definition" pos:[195,237] width:444 height:120
    spinner NStartLine "Render: Line Number Start" pos:[350,303] width:196 height:16 \
    range:[1,5000,1] type:#integer
    button BMD "Make Directories" pos:[5,13] width:173 height:34
    button BCal "Calculate" pos:[195,12] width:104 height:34
    label PsiPVOutput "PsiPVOutput" pos:[385,158] width:82 height:18
    label DXOutput "DX (mm)" pos:[385,179] width:87 height:18
    label DYOutput "DY (mm)" pos:[386,198] width:86 height:18
    label W_Output "W Used (mm)" pos:[506,98] width:118 height:18
    label H_Output "H Used (mm)" pos:[508,119] width:124 height:18
    label CompleteColOutput "Completed Column" pos:[491,184] width:135 height:18
    GroupBox grp19 "Render Progress" pos:[487,154] width:144 height:63
    spinner Delta "Hogel Size (microns)" pos:[279,162] width:70 height:16 \
    range:[100,3000,800] type:#integer
    button BCam "Make Camera" pos:[330,13] width:104 height:34
    global W,h1,H,DY,NK,NG,TopDirName--global variable definitions
    on BRen pressed do
    (
        --EVENT HANDLER FOR BRen Button Pressed
        for g=NStartLine.value to NStopLine.value do
        (
            LineText=gas string
            LineDirName=TopDirName+{/Line"+LineText
            for k=1 to NK do
            (
                ColText=k as string           -make filename
                FrameFileName=LineDirName+{/Col"+ColText+".jpg"
                $Cam1.pos=[-W/2+(k-1)*Delta.value/1000.0,-h1,-H/2+ \
                (g-1)*Delta.value/1000.0]      -move camera
                $Tar1.pos=[-W/2+(k-1)*Delta.value/1000.0,0,-H/2+(g-1)*Delta.value/1000.0]
                NI=NK+NA.value-1             -prepare frame parameters
                NJ=NG+NB.value-1
                N1x=1+NK-k
                N2x=N1x+NA.value-1
                N1y=g
                N2y=N1y+NB.value-1
            )
        )
    )
)
```

```

    Render\
        OUTPUTFILE: FrameFileName\
        OUTPUTWIDTH:N1 \
        OUTPUTHEIGHT:NJ \
        RENDERTYPE: #RegionCrop \
        REGION: #(N1x,N1y,N2x,N2y) \
        VFB:OFF           -render frame

        gc()           -empty garbage

        OutText=kas string
        CompleteColOutput.Caption=OutText
    )
)
)
)
on NA changed val do
(
NA.value=2.0*floor(NA.value/2.0)      -ensure NA is even
)
on NB changed val do
(
NB.value=2.0*floor(NB.value/2.0)      -ensure NB is even
)

on NM changed val do
(
NM.value=2.0*floor(NM.value/2.0)      -ensure NM is even
)
on NV changed val do
(
NV.value=2.0*floor(NV.value/2.0)      -ensure NV is even
)

)
on BMD pressed do
(
-EVENT HANDLER FOR BMD Button Pressed
TopDirName="C:/"+RenderDir.Text
makedir TopDirName
fori=NStartLine.value to NStopLine.value do
(
    LineText=ias string
    DirName=TopDirName+{/Line"+LineText
    makedir DirName
)
)

on BCAL pressed do
(
-EVENT HANDLER FOR BCAL Button Pressed
h1=Delta.value/2.0/1000.0*(NM.value-1)/tan(PsiPH.value/2.0)
hText=h1 as string
hOutput.caption="h(mm) = "+hText

NK=1+2*floor(Win.value*1000.0/2.0/Delta.value) as integer
NG=1+2*floor(Hin.value*1000.0/2.0/Delta.value) as integer
NKText=NK as string
NGText=NG as string
NKOutput.caption="NK = "+NKText
NGOutput.caption="NG = "+NGText

W=(NK-1)*Delta.value/1000.0
H=(NG-1)*Delta.value/1000.0
WText=Was string
Htext=Has string
W_Output.caption="W(mm) = "+WText
H_Output.caption="H(mm) = "+Htext

DX=(NA.value-1)*Delta.value/1000.0
DY=(NB.value-1)*Delta.value/1000.0
DXText=DX as string

```

```
DYText=DY as string
DXOutput.caption="DX(mm) = "+DXText
DYOutput.caption="DY(mm) = "+DYtext

PsiPV=2.0*atan(float(NV.value-1)/float(NM.value-1)*tan(PsiPH.value/2.0))
PsiPVTText=PsiPV as string
PsiPVOutput.caption="PsiPV = "+PsiPVTText
)

on BCam pressed do
(
    -EVENT HANDLER FOR BCam Button Pressed
    PsiCH=2.0*atan((W+DX)/2.0/h1)
    TargetCamera NAME:"Cam1" \
        SHOWCONE: true \
        FOV:PsiCH \
        POS: [-W/2,-h1,-H/2.0] \
        TARGET: (targetObject NAME:"Tar1" POS: [-W/2,0.0,-H/2.0])
    Viewport.setCamera $Cam1
)
)
CreateDialog DoubleParallaxRender
```

Appendix 6: Design Study of Compact RGB LED Hologram Illumination Source

A6.1 Introduction

We present here a design study of a new light source that is suitable for holography.* This light source is bright, efficient, small, lightweight and has a long life with low heat output, low power consumption and low étendue. The design allows for a wide range of customised performance in terms of beam footprint, beam irradiance distribution at the target, spectral composition of the emitted radiation, spatial coherence and beam divergence. In particular, the light source can be designed with a highly non-rectangular footprint† and with a radiant intensity distribution suitable for even illumination of holograms in which the illumination angle is as large as 70°, and in which the light source is positioned at close proximity. Geometrical illumination efficiency, as defined by the radiant energy used by the hologram (in either the reconstruction of the image or absorption) divided by the total radiant energy emitted by the source, can be as high as 70% for such highly tilted cases. This should be compared with typically less than 10% illumination efficiency which is available from previous light sources. In addition to providing higher geometric efficiency, the new source generally also provides better spatial coherence; in addition, the source can be optimised to produce exceptionally good spatial coherence at the cost of a lesser choice in the definition of the profile of radiant intensity. Finally, the source is characterised by high spectral radiance at one or more wavelengths. These wavelengths may be matched to the recording wavelengths used to write the hologram, leading to a reconstructed image of high brightness and low noise.

The new light source is based on recent developments in semiconductor diodes. These sources (e.g., Philips Lumileds) are available in appropriate wavelengths for display holography and general illumination applications (e.g., 455, 530 and 627 nm), have small emitter sizes (typically 1 mm², emitting within a solid angle of just under 2π rad), are narrow band (typically full-width half-maximum ~30 nm) and emit approximately 145 lm. Due to the Bragg selectivity of a typical reflection hologram made with modern silver halide materials, such diodes produce significantly superior brightness when compared with illumination by a 700 lm, 50 W MR16 halogen lamp. The electrical power consumed by each diode is typically only 1 W, and the operating life is of the order of 100,000 hours. In general illumination applications, approximately five diodes are required to produce the same perception of brightness as a single 50 W MR16 halogen lamp.

Using commercial ray-tracing packages, such as TracePro and Zemax, it is possible to conclude that the light emitted from modern diodes may be very conveniently injected into square light guides of approximately 1mm² sectional area made from glasses with refractive indices in the region of 1.8. ‡Such guides may be somewhat curved and are typically a few centimetres long. The guides serve to transport light from typically nine diodes (three of each colour) to an optical element that combines and redistributes the light into a beam of defined form. The guides also serve to limit and homogenise the angular and spatial distribution of radiation.

The light-emitting diodes (LED) used in the new light source do not require as much input energy as previous light sources, and this makes it possible, when illuminating a self-animating hologram that requires movement of the light source, to supply the power to the light source without the use of wires

* This design study was carried out in 2008 by one of the authors (DBR) in collaboration with John Fleming, whose kind permission to reproduce the study here is gratefully acknowledged by the authors.

† This means that rectangular hologram panels can be illuminated by an exactly matching rectangular illumination.

‡ Plastic light-guides may also be used. In later work these were found to be preferable to glass.

connecting the power source to the light. In addition, the heat output of the new light source is much lower than that from previous devices and this aids in preventing thermally caused chromatic distortions of the image where the light source is mounted close to the hologram being displayed.

A6.2 Monochromatic Light Source

Figure A6.1 shows a diagram of a monochromatic light source comprising three diode sources. The diodes used are LMXL-PM01-0080 Luxeon Rebels emitting at a centre wavelength of 530 nm with a full-width half-maximum bandwidth of 30 nm. The three diodes are optically cemented to square-section light guides 1.27 mm × 1.27 mm made from Schott SF57HHT glass or Ohara S-LAF52. The ends of each light guide comprise a linearly tapered section that increases the collection efficiency and improves the angular distribution of radiation. The LMXL-PM01-0080 diodes are sold with an attached silica gel lens assuring an approximately lambertian emission. This lens is easily removed and the diode is then optically attached directly to the light guide. The size of the light guide is chosen by making it just a little larger than the active area of the diode emitter. The maximum angle of output radiation is then fixed by the angle of total internal reflection of the guide glass, which, in this case, is approximately 33.75°. Having fixed the light-guide size in this way, the overall collection efficiency may be increased by including a tapered section to the smaller diode emission area. This tapered section collects rays that would normally pass outside the 33.75° angle, and so would not normally be collected. The angle of the tapered section redirects such rays to fall within the 33.75° limit. The length and angle of this section can be chosen such that both the collection efficiency is optimal and the output angular distribution of radiation from the guide is closest to a “top hat” distribution. In addition, there exists an optimum refractive index that minimises the size of the light guide for this collection geometry.

Typical collection efficiencies of the light guides (power collected to total power emitted) are approximately 95%. An alternative collection scheme would be to taper and dielectrically coat the guide in the form of a parabolic reflector. However, this method is more expensive and does not produce better

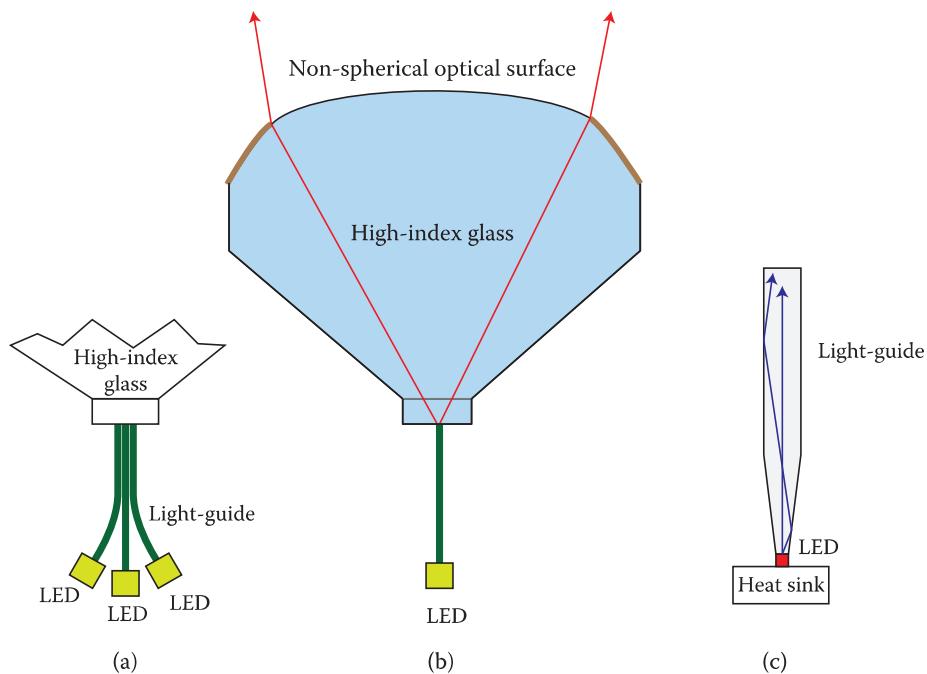


FIGURE A6.1 Monochromatic light source showing details of (a) the coupling of radiation from three diodes to three light guides and (b) subsequent coupling of the light guides to a main high-index optic with non-spherical surface. (c) Details of the tapered light guide matched to the LED.

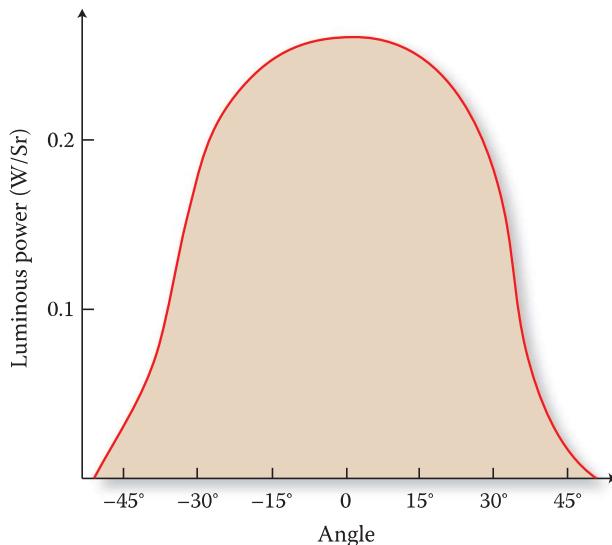


FIGURE A6.2 Radiant intensity distribution emitted from a light guide fed by a LMXL-PM01-0080 diode and coupled into Sumita K-PSFFn202 glass.

results. Each of the light guides may be bent somewhat to transport the radiation from each diode to a desired location on the main optic.

Figure A6.2 shows a typical plot of the radiant intensity distribution emitted at the end of a light guide as calculated by diffractive ray tracing. The light guide is optically attached to the main optic element, which is made from Sumita K-PSFFn202 glass having a refractive index of 2.035 at 530 nm. The radiation emission into the Sumita glass is therefore somewhat reduced in angular extent by refraction at the interface. The overall appearance of the distribution is quite close to a top hat with a width of $\pm 32^\circ$.

Returning to Figure A6.1, the radiation from the light guides is injected into the main optic. This constitutes a cylindrical block of Sumita K-PSFFn202 glass, 41.5 mm long, with a specially designed non-spherical front surface incorporating a graded aperture.

The function of the non-spherical surface is to redirect the radiation coming from each of the three light guides such that the irradiance distribution (usually defined at a target plane) downstream of the lens matches a given target distribution (which, for example, would be the hologram in applications of a light source for display holography). Typically, this target distribution will be a top hat intensity distribution over a rectangular area and the target will usually be tilted. Another function of the refracting surface is to ensure that each point on the target plane “sees” a source of good spatial coherence.

The function of the graded aperture is to eliminate rays impinging on the front surface of the lens at too great an angle without causing a secondary diffractive source. In situations where the light guides are designed to produce a good top hat intensity distribution, a hard aperture is also suitable.

The reason that a single non-spherical refracting surface is used is that it economises the cost of fabrication of the light source (the Sumita glass in question is mouldable) and it allows an efficient and compact design. The geometry of the refracting surface is, however, non-trivial to calculate and this subject will be covered in detail in the following sections.

A6.3 Polychromatic Light Source

The positions at which the three light guides are attached to the main optic in Figure A6.1 must be chosen carefully to optimise the characteristics of the light source. This is because the light guides are of finite size and the lens surface always has optimum performance only for one unique source position. In the case of a bichromatic source, it is sometimes possible to find acceptable positions for two or more light

guides carrying radiation of different wavelengths. In this case, a bichromatic light source is extremely similar to the monochromatic source already described. However, for three colours, this is in practice rarely possible (particularly in the case of hologram illumination for realistic hologram illumination geometries).

The light source that we describe here requires the use of high refractive index glasses. However, it is known that such glasses always have low Abbe numbers and, as a result, the behaviour of the main refracting surface changes with wavelength. In practice, we find that as long as the Abbe number is not too low, then a lens surface designed for one wavelength can be used at another if the position of the light guide on the main optic is changed by a small amount. Nevertheless, this displacement is rarely larger than the size of the light guide.

We are therefore almost always obliged to employ dichroic reflectors to deliver the red, green and blue radiation to the main optic. Figure A6.3 illustrates how this works. The main optic is now split into three Sumita K-PSFn202 glass sections that are optically connected. The top section comprises the main lens surface with graded neutral density filter. The next section contains a 45° dichroic mirror embedded in a cylindrical block of glass. Three light guides carrying red radiation from three LMXL-PD01-0040 diodes are optically attached at a location in point A (side by side). The red radiation is reflected by the dichroic mirror, which is transparent to blue and green light. An additional cylindrical block contains a dichroic mirror that reflects green and transmits blue light. Three light guides carrying green radiation from LMXL-PM01-080 diodes are mounted at location B. Finally, three blue guides are mounted at location C.

This design allows the exact positions of the red, green and blue light guides to be adjusted precisely and independently for the best performance. Often, for highly tilted holographic displays, the lens design implies a much greater sensitivity to the positioning of the light guides on the main optic's surface in one dimension rather than the other. This allows similar-coloured light guides to be stacked in that dimension when higher power is required.

The Sumita K-PSFn202 glass is a particularly suitable glass for the present application because it has a high index (2.035 at 530 nm), relatively high Abbe number ($V_d = 21.5$) and melts at a relatively low temperature (yielding point = 486°C) making it easier to mould and polish than other glasses which typically

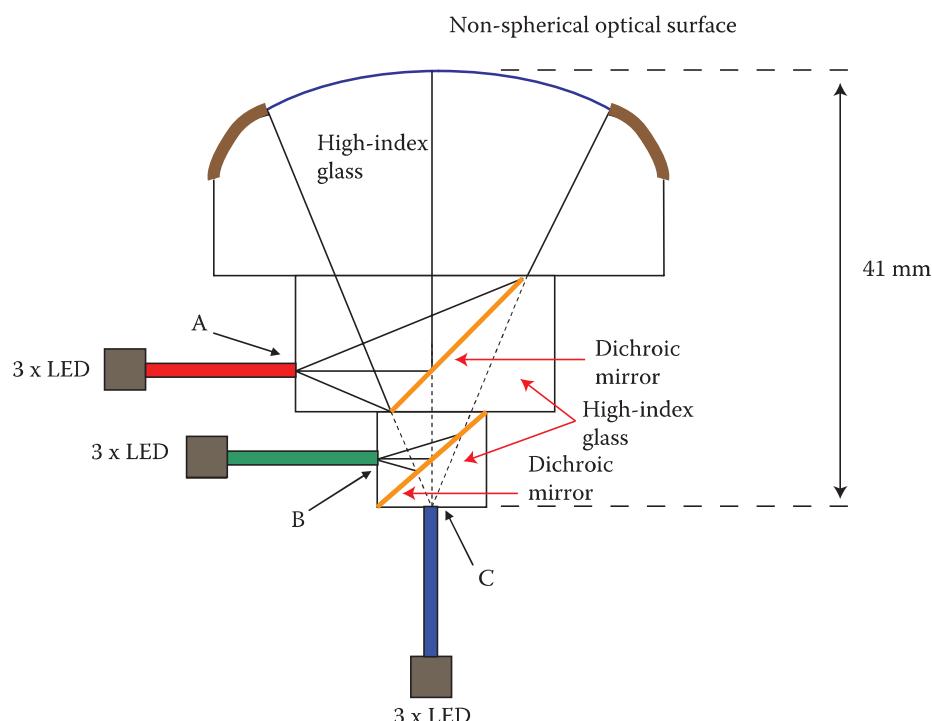


FIGURE A6.3 Trichromatic light source coupled to nine diodes.

have melting points 100°C higher than this. It also has excellent transparency in all visible wavelengths including the blue spectrum.

Other glasses, such as the higher index Sumita K-PSFn214, are unsuitable for polychromatic applications because their transmittance in the blue spectral region is poor and their Abbe number is rather lower. However, their performance is superior for monochromatic applications in the green and red spectral region. Lower index glasses such as SF57 ($n = 1.846$), which have good transparency and Abbe characteristics, can be used for some applications with available diode sources. However, for the majority of applications, glasses with a refractive index of greater than 2 are to be preferred.

A6.4 Design of the Main Lens Surface

The function of the main lens surface is to refract the rays emanating from a light pipe and to redistribute them such that the radiation distribution on a tilted target plane downstream of the lens matches, as closely as possible, a chosen distribution, while the spatial coherence “seen” by individual points on the target plane within a proscribed area remains optimal. Using a geometric optics approach, this may be cast in the form of a standard optimisation problem and may be solved computationally for all cases of interest.

A6.4.1 Monochromatic Light Source

Figure A6.4 (side view) depicts a lighting source optically coupled to a thick lens illuminating a target plane inclined at angle Φ . Figure A6.5 shows the corresponding view from behind the light. The light source emits radiation over a certain solid angle. We assume that the source is point-like to start with; this will be generalised later to a source of finite size. Likewise, we shall assume that the radiant intensity distribution of the source is constant per unit of solid angle and that the emission is in the form of a cone of maximum angle ϕ_m . Later, we will check our results for the real radiant intensity distributions that we calculate explicitly for each light guide.

The source emits rays A, B and C. These propagate through the thick substrate of the lens, hitting the lens surface at points D, E and F, respectively. At the lens surface, the rays are refracted and intersect the target at points H, I and J, respectively. Both Figures A6.4 and A6.5 have a distorted scale for clarity. In practice, the light source and lens are typically some centimetres in size and the target may have dimensions of some tens of centimetres to some tens of metres.

By choosing a general lens surface, all rays emitted from the source can be made to illuminate a target of a chosen shape. Furthermore, by constraining the lens surface more stringently, it is possible to

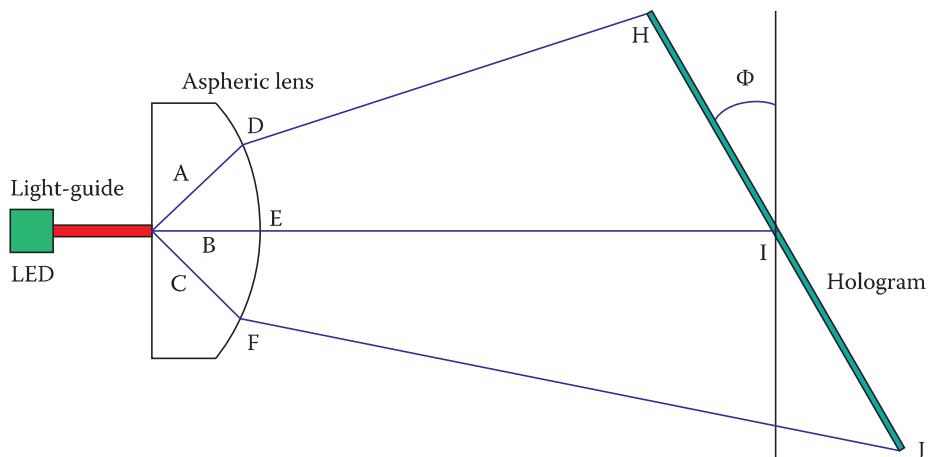


FIGURE A6.4 Side view of a simplified light source comprising a light guide plus the main optic and a tilted target plane.

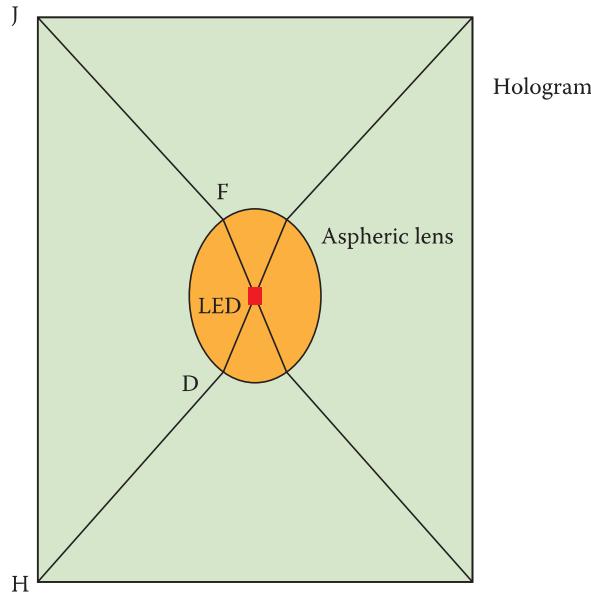


FIGURE A6.5 View of a simplified light source and target plane as seen from behind the light guide.

produce a wide range of power density distributions over the target. Most importantly, it is possible to produce almost exactly rectangular distributions of constant power density on a tilted target as required by the problem of hologram illumination. At extreme parameter values, particularly with hologram tilt, limitations do appear and it may not be possible to then attain exactly flat power distributions. However, numerical studies indicate that there is an extremely wide useful parameter space in which useful light sources may be designed.

A6.4.2 Mathematical Formulation

With reference to Figure A6.6, let the lens surface height, $z(x,y)$, be described by the following polynomial spectral form:

$$z(x,y) = \sum_{\alpha=0}^N \sum_{\beta=0}^{\alpha} A_{\alpha\beta} r^{\alpha} \cos(\beta\theta) \quad (\text{A6.1})$$

where $r \geq 0$, $2\pi > \theta \geq 0$, $y = r \sin \theta$ and $x = r \cos \theta$.

We assume the plane $y=0$ is a plane of symmetry and, accordingly, we omit the sine terms in the expansion. In addition, the fact that z must be single-valued at $r=0$ implies that $\alpha \geq \beta$ and that $A_{0\beta} = 0 \forall \beta \geq 1$.

Assuming a point source at $(0,0,0)$ that emits over a cone angle of $\phi \leq \phi_m$, the intersection (x,y) of a given ray at angle θ with the lens surface may be traced using geometric optics through the lens to its intersection with the target plane. By considering the projection of the outer rays as defined by $\phi = \phi_m$ on the target plane, we may choose the coefficients $A_{\alpha\beta}$ such that the intersection of outer rays most closely matches a given target contour.

More specifically, if we define

$$\Xi \equiv \sum_{\phi=\phi_m} (x_{IO} - x_T)^2 + (y_{IO} - y_T)^2 \quad (\text{A6.2})$$

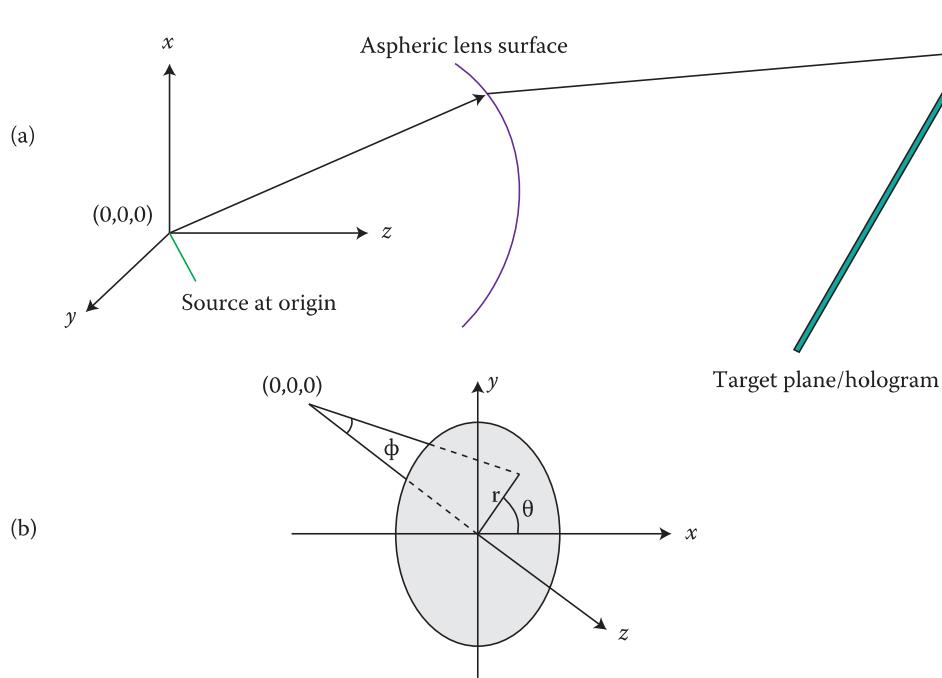


FIGURE A6.6 Geometrical coordinate systems used. (a) Side view and (b) view looking back through the lens towards the LED source.

where x and y represent Cartesian coordinates of the target plane, the subscripts IO and T, respectively, denote outer ray intersections and target contour coordinates, and the sum is carried out over all outer rays, then the equation

$$\frac{\partial \Xi}{\partial A_{\alpha\beta}} = 0 \quad (\text{A6.3})$$

defines a lens that will act to deviate the outer rays in such a manner that they intersect most closely with a given target contour on a given target plane.

The manner in which the target contour coordinates are specified is important. One could, for instance, insist that a given outer ray passes exactly through a certain point on the target contour. However, this is rather restrictive and it is more convenient to define a given target contour coordinate pair as being that coordinate pair whose distance squared to (x_{IO}, y_{IO}) is smallest.

In general, there are too many solutions to Equations A6.2 and A6.3 for such a system to be of use in calculating a real lens. This is because rays “cross over” too easily and most solutions consist of lenses that therefore result in a light source of low spatial coherence. In addition, most solutions are characterised by inappropriate power density distributions on the target plane.

To resolve the inadequacies of the above solution, we introduce a modified form for Ξ :

$$\Xi \equiv \sum_{\varphi=\varphi_m} (x_{IO} - x_T)^2 + (y_{IO} - y_T)^2 + \gamma \sum_{S_m} (\epsilon_{IO} - \epsilon_T)^2 \quad (\text{A6.4})$$

where γ is a Lagrange multiplier, ϵ_{IO} is the calculated power density at a given point on the target plane, ϵ_T is a chosen target power density for that same location, the additional sum is taken over all solid angle elements of the unit sphere surrounding the source and

$$\sum_{S_m} (\epsilon_T) = E \quad (\text{A6.5})$$

where E is the total power emitted by the source.

By introducing this second “energy” term in the error function, solutions in which no rays cross may be preferentially generated and, in addition, tailored power density distributions may be catered for.

It is an interesting empirical fact that the degree of freedom implicit in the problem of the design of a general lens for a point light source is such that many power density distributions and illumination contours may be generated without rays crossing. This situation remains broadly the same with the introduction of a small but finite size to the source.

A6.4.3 Geometrical Ray Tracing in the Point-Source Approximation

Assuming a point source at $(0,0,0)$ that emits over a cone angle of $\phi \leq \phi_m$, the intersection (x,y) of a given ray at angle θ with the lens surface is given by the non-linear equation:

$$\begin{aligned} \sum_{\alpha=0}^N \sum_{\beta=0}^{\alpha} A_{\alpha\beta} r^{\alpha} \cos(\beta\theta) \tan\phi &= r \\ x &= r \cos\theta \\ y &= r \sin\theta \end{aligned} \quad (\text{A6.6})$$

This can be effectively solved using a Newton–Raphson method:

$$\begin{aligned} r_{n+1} &= r_n - \frac{g(r_n)}{\left. \frac{\partial g}{\partial r} \right|_n} \\ g &= \sum_{\alpha=0}^N \sum_{\beta=0}^{\alpha} A_{\alpha\beta} r^{\alpha} \cos(\beta\theta) \tan\phi - r \end{aligned} \quad (\text{A6.7})$$

The normalised propagation vector, \hat{k} , of a given ray emitted from the source and intersecting the lens surface at (x,y) is given by (Figure A6.7):

$$\hat{k} = \frac{1}{(x^2 + y^2 + z^2)^{1/2}} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sin\phi \cos\theta \\ \sin\phi \sin\theta \\ \cos\phi \end{pmatrix} \quad (\text{A6.8})$$

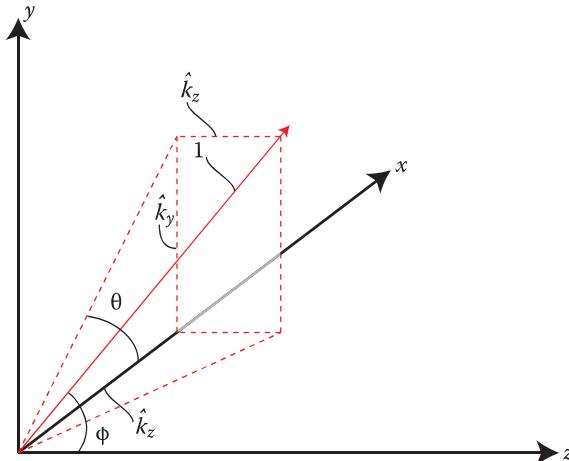


FIGURE A6.7 Diagram showing the unit ray vector \hat{k} in relation to the various coordinates.

At the lens surface, this ray refracts. Denoting all quantities to the right of the surface by starred variables, the vector form of Snell's law states:

$$\hat{\mathbf{k}}^* = \mu \hat{\mathbf{k}} - \hat{\mathbf{N}} \left\{ \mu (\hat{\mathbf{k}} \cdot \hat{\mathbf{N}}) - \sqrt{\left[1 - \mu^2 \right] + \mu^2 (\hat{\mathbf{k}} \cdot \hat{\mathbf{N}})^2} \right\} \quad (\text{A6.9})$$

where $\hat{\mathbf{N}}$ is the unit normal vector to the surface at (x, y) and

$$\mu = \frac{n}{n^*} \quad (\text{A6.10})$$

is the ratio of the refractive indices to the left and right of the lens surface, respectively.

Defining the lens function,

$$\Theta(x, y, z) = \Theta(r, \theta, z) \equiv f(r, \theta) - z = 0 \quad (\text{A6.11})$$

where

$$f(r, \theta) = \sum_{\alpha=0}^N \sum_{\beta=0}^{\alpha} A_{\alpha\beta} r^{\alpha} \cos(\beta\theta), \quad (\text{A6.12})$$

the surface normal vector of the lens may be defined as

$$\mathbf{N} \equiv -\nabla \Theta = \begin{pmatrix} -\nabla \Theta_x \\ -\nabla \Theta_y \\ -\nabla \Theta_z \end{pmatrix} = \begin{pmatrix} -\cos \theta \frac{\partial f}{\partial r} + \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \\ -\sin \theta \frac{\partial f}{\partial r} - \frac{\cos \theta}{r} \frac{\partial f}{\partial \theta} \\ 1 \end{pmatrix} \quad (\text{A6.13})$$

and the unit normal as

$$\hat{\mathbf{N}} = \frac{\mathbf{N}}{\sqrt{\mathbf{N} \cdot \mathbf{N}}}. \quad (\text{A6.14})$$

A6.4.4 Ray Intersection with Target Plane

Figure A6.8 shows a diagram of a target plane inclined at an angle ψ to the (x, y) plane. The intersection (x_H, y_H, z_H) of a ray propagating from the lens at point (x_0, y_0, z_0) having k vector $\hat{\mathbf{k}}^*$ with the target plane is given by

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + \lambda \hat{\mathbf{k}}^* = \begin{pmatrix} x_H \\ y_H \\ z_H \end{pmatrix} = \begin{pmatrix} z_H \cot \psi - d \cot \psi \\ y_H \\ z_H \end{pmatrix} \quad (\text{A6.15})$$

where we have used the defining equation of the target plane:

$$x_H = (\cot \psi)(z_H - d) \quad (\text{A6.16})$$

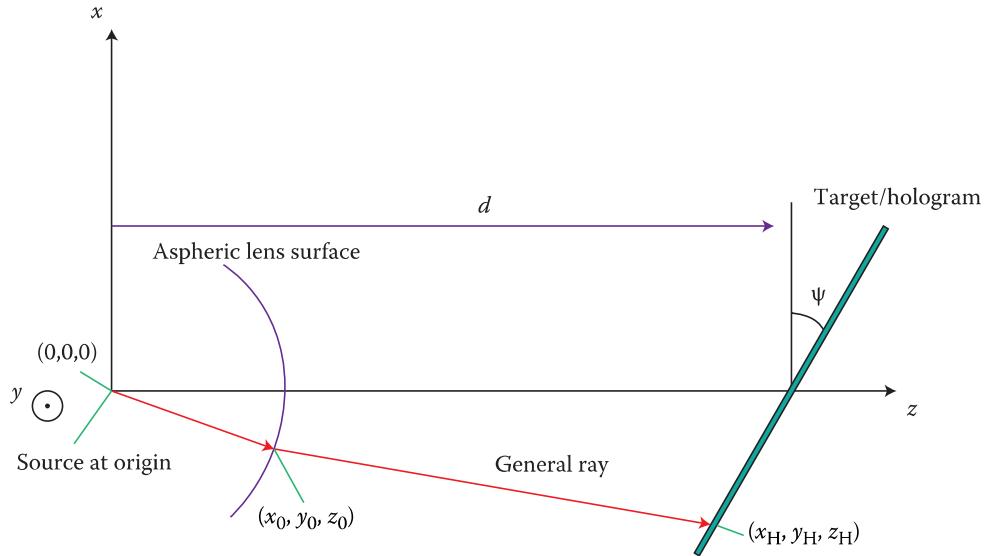


FIGURE A6.8 Diagram showing intersection coordinates on the lens surface and on the tilted target plane.

Solving the above equation yields

$$\begin{pmatrix} x_H \\ y_H \\ z_H \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + \frac{z_0 - d - x_0 \tan \psi}{\hat{k}_x^* \tan \psi - \hat{k}_z^*} \hat{\mathbf{k}}^* \quad (\text{A6.17})$$

In terms of a right-handed coordinate system (x',y') with origin $(0,0,d)$ on the target plane itself, we may write

$$\begin{aligned} x' &= +\sqrt{x_H^2 + (z_H - d)^2} & z_H \geq d \\ x' &= -\sqrt{x_H^2 + (z_H - d)^2} & z_H < d \\ y' &= y_H \end{aligned} \quad (\text{A6.18})$$

This equation defines the intersection point of a given ray with the target plane in target coordinates. Given the spectral polynomial coefficients of a lens and letting θ vary from 0 to 2π and ϕ from a small number, ε to ϕ_m , we are therefore able to calculate the complete pattern of ray intersections on the target plane.

A6.4.5 Calculation of Power Density Distribution at Target Plane

We now assume that the point source emits a constant power, κ , per unit of solid angle within its emission cone. Figure A6.9 shows that portion of the unit sphere surrounding the source through which energy from the source passes. In particular, a general element of solid angle is identified as having a solid angle

$$\delta\Omega = (\cos\phi_N - \cos\phi_{N-1})\delta\theta \quad (\text{A6.19})$$

The four vertices of this element are ray-traced using the equations in the previous section to calculate their intersections with the target plane. These four intersections define a corresponding area element, δA , on the target plane, which may be evaluated using Heron's formula. The energy passing through the element of solid angle, as defined by Equation A6.19, is therefore brought to bear on this elemental area, δA . As $\delta\Omega$ and δA tend to zero, so

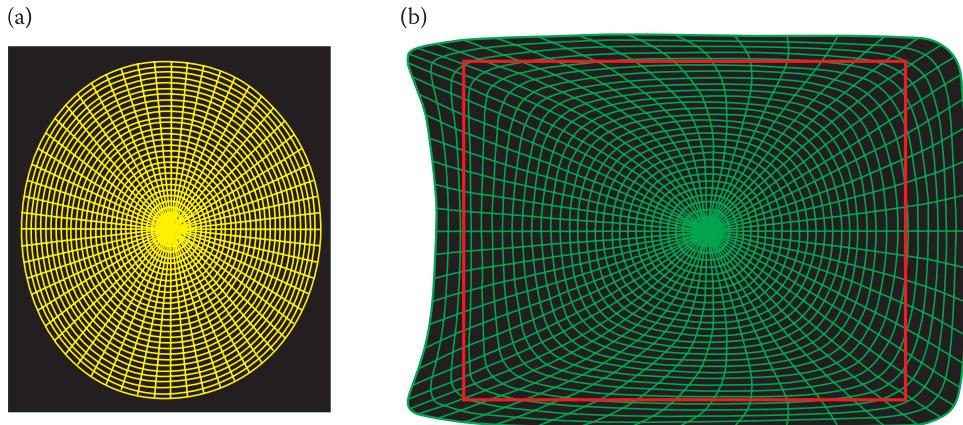


FIGURE A6.9 (a) Lines of equal θ and equal ϕ on the unit sphere and (b) on the target plane b ; target contour shown in red.

$$\lim_{\delta\Omega, \delta A \rightarrow 0} \kappa \frac{\delta\Omega}{\delta A} = \frac{dE}{dA} \equiv \varepsilon \quad (\text{A6.20})$$

where ε is the power density at a point on the target plane due to radiation passing through a given point on the unit sphere.

A6.4.6 Numerical Solution

Equations A6.3, A6.4 and A6.5 are solved numerically using the expansion of Equation A6.1. Ray intersections with the general lens surface are solved using the Newton–Raphson solution presented in Equations A6.6 and A6.7. Ray intersections with the target plane are calculated using Equations A6.8 through A6.18. Power density is calculated using Equation A6.20. For reasons of increased computational speed, the energy summation in Equation A6.4 is usually replaced by a summation along several (typically 10) lines of constant θ . Typically, a value of $N = 8$ to 12 is used. Different (θ, ϕ) meshes from 30×30 to 200×200 are employed.

In general, Equations A6.3 through A6.5 represent a highly non-linear multidimensional system with local minima. It is therefore important to prime the solution with a sensible guess. For many cases, we can use a perfect sphere for this purpose. However, the best initial guess may well depend on each individual case. Where highly tilted target planes are used, solutions should be computed initially for the case with no tilt and then the tilt gradually increased.

Practical experience shows that the best numerical solution is offered by the combination of a quasi-Newton method with a polytrope algorithm. The polytrope algorithm is useful for advancing to the vicinity of a solution, but the quasi-Newton method is often required to find the exact solution. The techniques of Levenberg–Marquardt, conjugate gradients and steepest descents are of little use in the context of the current problem.

A6.4.7 Ray Tracing with Finite Source Size

The best way to calculate useful lenses is by initially using the point-source approximation and then, by checking the results, using a finite-source simulation. Because, in many applications, a lens must be designed for use simultaneously at three component colours, finite-source ray tracing allows a proper characterisation of the lens's behaviour with a change of refractive index with illuminating wavelength. In addition, finite-source ray tracing allows us to compensate for the various effects introduced by the finite source itself or by a different refractive index by varying the lateral position of the source. Indeed, whereas lenses are best calculated using the point-source approximation at a chromatically median

refractive index, finite-source simulation must be used to calculate the exact lateral position of each source of a given wavelength.

A6.4.7.1 Method

We assume that the source is composed of a regular grid of elemental point emitters. Typically, we may use a grid of 10×10 such emitters for a square source. Each point emitter at $(x_i, y_j, 0)$ is assumed to emit over a cone angle of $\phi \leq \phi_m$. The intersection (x_{li}, y_{lj}) of a given ray at angle θ with the lens surface from a given elemental emitter is now given by the following non-linear simultaneous equations:

$$\begin{aligned} g &\equiv \sum_{\alpha=0}^N \sum_{\beta=0}^{\alpha} A_{\alpha\beta} r^{\alpha} \cos(\beta\theta) \tan\phi - R = 0 \\ R^2 + 2R(x_i \cos\theta_s + y_j \sin\theta_s) + x_i^2 + y_j^2 - r^2 &= 0 \\ \tan\theta &= \frac{R \sin\theta_s + y_j}{R \cos\theta_s + x_i} \\ x_{li} &= r \cos\theta \\ y_{lj} &= r \sin\theta \end{aligned} \quad (\text{A6.21})$$

The coordinates (r, θ) refer to the usual polar coordinate system of Figure A6.6 whose origin is at $(x, y) = (0, 0)$. The coordinates (R, θ_s) refer to a shifted polar coordinate system whose origin is at the emitter in question, $(x, y) = (x_i, y_j)$. At the boundary $\phi = \phi_m$, these equations are solved for each i and j using a Newton–Raphson method similar to the one introduced in the previous section:

$$r_{n+1} = r_n - \frac{g(r_n)}{\left. \frac{\partial g}{\partial r} \right|_n} \quad (\text{A6.22})$$

For each emitter (i, j) this leads to a locus of intersection points $[x_{li}(\theta), y_{lj}(\theta)]$ forming a contour at the lens surface. From this locus of points, a non-orthogonal coordinate system (ξ, ζ) is generated to discretise the space interior to the contour. The exact choice of coordinate system is not particularly important. Here we shall choose

$$\begin{aligned} \xi &\equiv r/r_B(\theta) \\ \zeta &\equiv \theta \end{aligned} \quad (\text{A6.23})$$

where

$$r_B(\theta) = \sqrt{x_{li}(\theta)^2 + y_{lj}(\theta)^2} \quad (\text{A6.24})$$

This coordinate system is illustrated in Figure A6.9a for the case of a 30×60 mesh. For calculations, a 200×200 mesh should be used.

For each elemental emitter, the x , y and z coordinates of each node of the (ξ, ζ) system are defined as

$$\begin{aligned} x &= \xi r_B(\theta) \cos(\zeta) \\ y &= \xi r_B(\theta) \sin(\zeta) \\ z &= \sum_{\alpha=0}^N \sum_{\beta=0}^{\alpha} A_{\alpha\beta} (\xi r_B(\theta))^{\alpha} \cos(\beta\theta) \end{aligned} \quad (\text{A6.25})$$

These nodes define a specific (but complete) set of rays emitted from a given elemental emitter, all of which intersect the lens surface. It is a trivial matter, given the nodes and the position of a given elemental source to calculate the set of the normalised k vectors that describe the rays. Once the k vectors are known, then generally each four neighbouring k vectors define an element of area on the lens. The projection of such elemental areas onto the plane of $z = 0$ is of course shown in Figure A6.9a. Using the vector form of Snell's law that we have given in Equation A6.9, we may trace all rays to the target plane where each elemental area will be projected as shown in Figure A6.9b for the case of a 30×60 mesh.

The solid angle of each elemental area on the lens surface is defined as

$$\delta\Omega(\xi, \varsigma) = \iint \frac{\hat{\mathbf{k}} \cdot d\mathbf{A}}{|\mathbf{k}|^2} = \hat{\mathbf{k}} \cdot \hat{\mathbf{N}} \frac{\delta A}{|\mathbf{k}|^2} \quad (\text{A6.26})$$

where $\hat{\mathbf{N}}$ is given by Equations A6.13 and A6.14, and the area element, δA , is calculated by Heron's formula after decomposing the (ξ, ς) element on the lens surface into a pair of triangles.

Because the power emitted per steradian by each elemental source is known, we may easily calculate the power density passing through the lens within this solid angle. In addition, because we know how each elemental area of Figure A6.9a transforms to its corresponding area in Figure A6.9b on the target plane, we may easily work out the power density distribution on the target plane for each elemental emitter. Again, we can use the technique of dividing each elemental area on the target plane into two triangles and applying Heron's formula.

Because each elemental emitter, in general, illuminates the lens in a slightly different way, both the (ξ, ς) coordinate system and its projection onto the target plane will be different for each emitter. For each elemental emitter, we may therefore define a regular Cartesian system at the target plane and interpolate the calculated power density distribution onto this system. In this way, the power density, $P_{ij}(X, Y)$, may be calculated for each elemental emitter (i, j) on the regular Cartesian grid (X, Y) , covering the target plane. Then, the total power density distribution at the target plane due to the finite source is simply

$$P = \sum_i \sum_j P_{ij}(X, Y) \quad (\text{A6.27})$$

A6.4.7.2 Spatial Coherence

In the case of the point-source approximation and if the lens in question has also been designed specifically not to allow rays to cross over on their journey to the target plane, then each point on the target plane will effectively "see" only one illuminating point. This corresponds to the case of a light source of perfect spatial coherence. When we replace the point source, however, with a source of finite extent, each point on the target plane will now "see" a variety of rays impinging at different angles from the source. This corresponds to the case of a source of finite spatial coherence. In applications such as holography, the spatial coherence parameter of a source must usually be as large as possible.

To calculate a measure of the spatial coherence of our source, we should first estimate the average value of the k vector illuminating a given point on the target plane. We do this by defining

$$\langle \hat{\mathbf{k}} \rangle \equiv \frac{\sum_i \sum_j \hat{\mathbf{k}} P_{ij}}{\sum_i \sum_j P_{ij}} \quad (\text{A6.28})$$

The angle that a given ray from a given elemental source makes to this average value is then given by

$$\psi_{ij} = \cos^{-1} \left(\langle \hat{\mathbf{k}} \rangle \cdot \hat{\mathbf{k}}_{ij} \right) \quad (\text{A6.29})$$

And a measure of the spread in illumination angles for the point (X, Y) is given by

$$\langle \delta\psi \rangle \equiv 2 \sqrt{\frac{\sum_i \sum_j (\psi_{ij}^2 P_{ij})}{\sum_i \sum_j P_{ij}}} \quad (A6.30)$$

The parameter $\langle \delta\psi \rangle$ effectively measures the angular illuminating source size “seen” by a point (X, Y) on the target plane. It is related to the spatial coherence, l_s , of the source by

$$l_s = \frac{\kappa\lambda}{\langle \delta\psi \rangle} \quad (A6.31)$$

where κ is a geometrical form factor that is 1.22 for circular sources.

A6.5 Selected Cases: Computational Results

A6.5.1 Case 1

We study here the case of a rectangular reflection hologram 500 mm \times 600 mm in size (in landscape format), illuminated using an overhead light source positioned 500 mm at a diagonal distance from the centre of the hologram. The light source is arranged such that rays strike the centre of the hologram at 45°. This geometry allows an extremely compact lighting arrangement.

We use the design presented in Figure A6.3 with nine Luxeon Rebel diodes comprising three LMXL-PD01-0040 red diodes (assumed 630 nm), three LMXL-PM01-0080 green diodes (assumed 530 nm) and three LMXL-PR01-0275 blue diodes (assumed 455 nm). The light guides employed are 1.27 mm \times 1.27 mm square sections made from Schott SF57HHT glass. The main optic is made from Sumita K-PSFn202 glass. The radiant intensity exiting the light guides at the Schott SF57HHT–Sumita K-PSFn202 interface is calculated by diffractive ray tracing employing the published emission data for the diodes. This is shown in Figure A6.2 for the case of the green diodes for the azimuthal angles of 0, $\pi/2$, π and $3\pi/2$. This data is averaged over the four azimuthal angles, truncated at 34° to model the effect of a hard aperture of this transmission angle and, finally, a top hat distribution independent of azimuthal angle is fitted to the resultant data.

To calculate the shape of the lens surface, we assume initially that the light guide is a point source located at (0,0,0) according to Figure A6.6. We then set up the target geometry to be a rectangle 500 mm \times 600 mm tilted at 45°, again according to Figure A6.6. Using an appropriate Lagrange multiplier in Equation A6.4, assuming an eighth order expansion in Equation A6.6 and employing the fitted top hat emission distribution, Equations A6.3 and A6.4 are solved numerically on a PC for $n = 2.035$. A contour plot of the calculated lens is shown in Figure A6.10a. The lens centre is at 41.5 mm from the point source and contours are spaced at 0.485 mm. The green outer line denotes the 34° boundary.

Figure A6.10b shows a plot of the target (hologram) plane and intersections of the set of rays of constant θ and constant ϕ . The red outline shows half of the target rectangle. As the plot is symmetric about the centre horizontal line, the bottom half of the target rectangle is omitted to show the intersections more clearly. The outer contour corresponds to the contour of $\phi = 34^\circ$. Clearly, it can therefore be seen that the lens is acting to shape the light such that it falls nicely within the target contour. In addition, we can see that no contours cross.

Figure A6.10c shows a plot of the irradiance distribution on the target plane. Different colours indicate a difference in irradiance of 12.5%. Clearly, the radiation power is extremely constant over the target, falling rapidly to zero at the edges. The black hole at the centre of the plot indicates the start of the polar grid over which the equations are solved.

Having determined the form of the lens surface by solving the point-source case for 530 nm, we must verify that this lens can work acceptably with the real size of the light guides and with the red and blue emissions. To do this, we use the equations derived in Section A6.4.7 to compute the irradiance

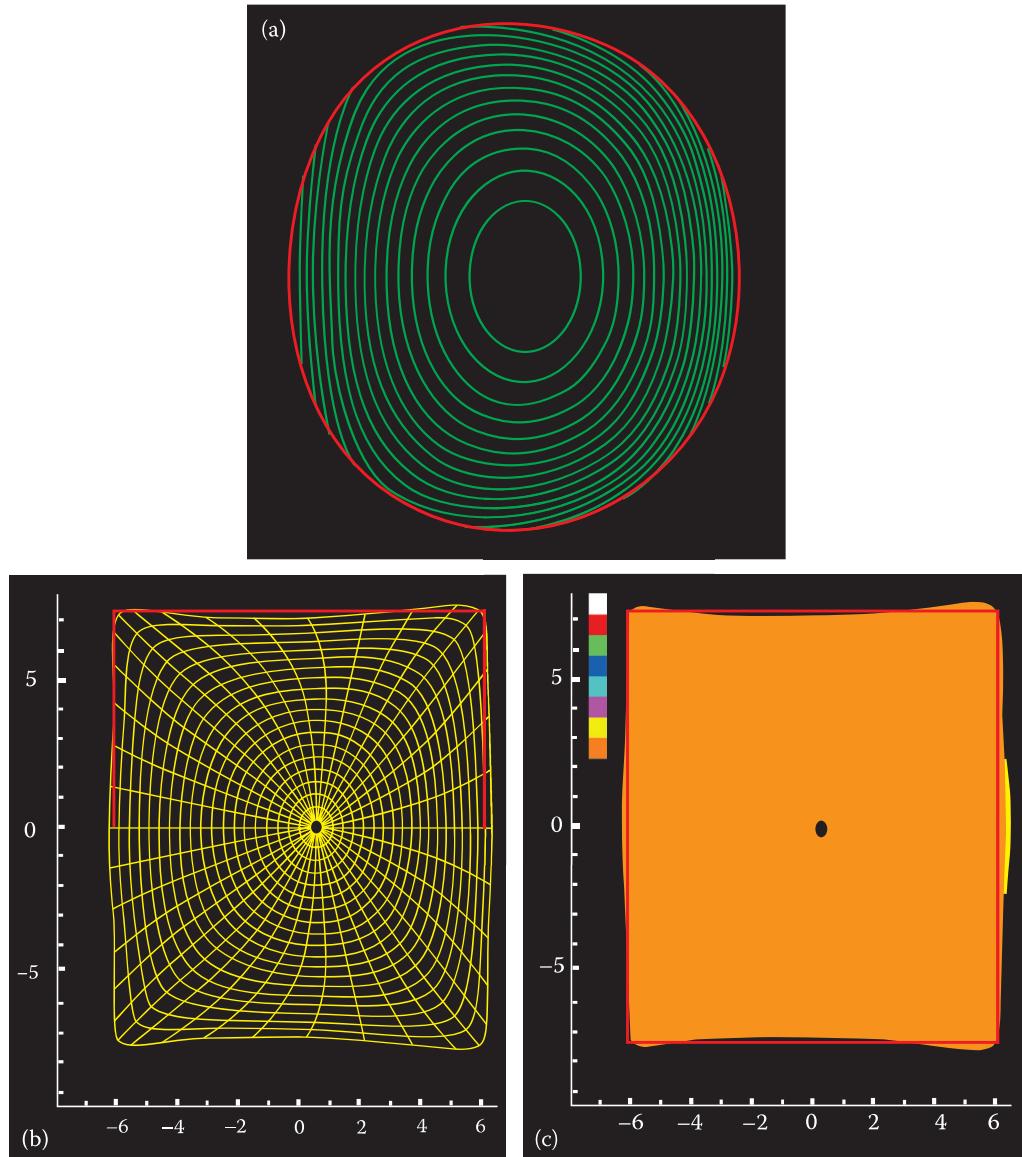


FIGURE A6.10 (a) Calculated height contours of a Sumita K-PSFFn202 glass lens; (b) calculated lines of constant θ and constant ϕ on the target plane with target contour shown in red; and (c) false-colour plot of the calculated irradiance distribution (c). Case of 500 mm \times 600 mm hologram (landscape format) illuminated at 45° from overhead at a distance of 500 mm. Length scales in relative units. Light source positioned to left of diagrams (b) and (c). Power scale: change in colour represents a change in irradiance by 12.5%. Calculations assume a single point source at (0,0,0) and a wavelength of 530 nm.

distribution at the target plane due to the nine actual red, green and blue light guides. Each light guide is divided up for this purpose into 100 equally distributed elemental emitters. Light guides carrying the same colour are stacked symmetrically next to each other in the y direction (Figure A6.8), with the centre diode being located at $y = 0$. For each colour, an optimum offset position in the x direction is calculated for the three guides by solving the equations for the irradiance for a given offset and then re-solving until the best distributions are found for a given offset. Figure A6.11a through c shows graded intensity plots of the optimised irradiance distributions for the green, red and blue emissions, respectively. Differences in shade indicate a 10% difference in irradiance. Clearly, the light source performs extremely well, giving a variation of irradiance over the target of only $\pm 10\%$. Energetic efficiencies as measured by power within the target area divided by total power falling within and outside are 94%, 80% and 92% for the red, green and blue emissions, respectively. Calculations of the spatial coherence, as measured from multiple points on the hologram surface, indicate that the effective size of the source is approximately 5% of its diameter.

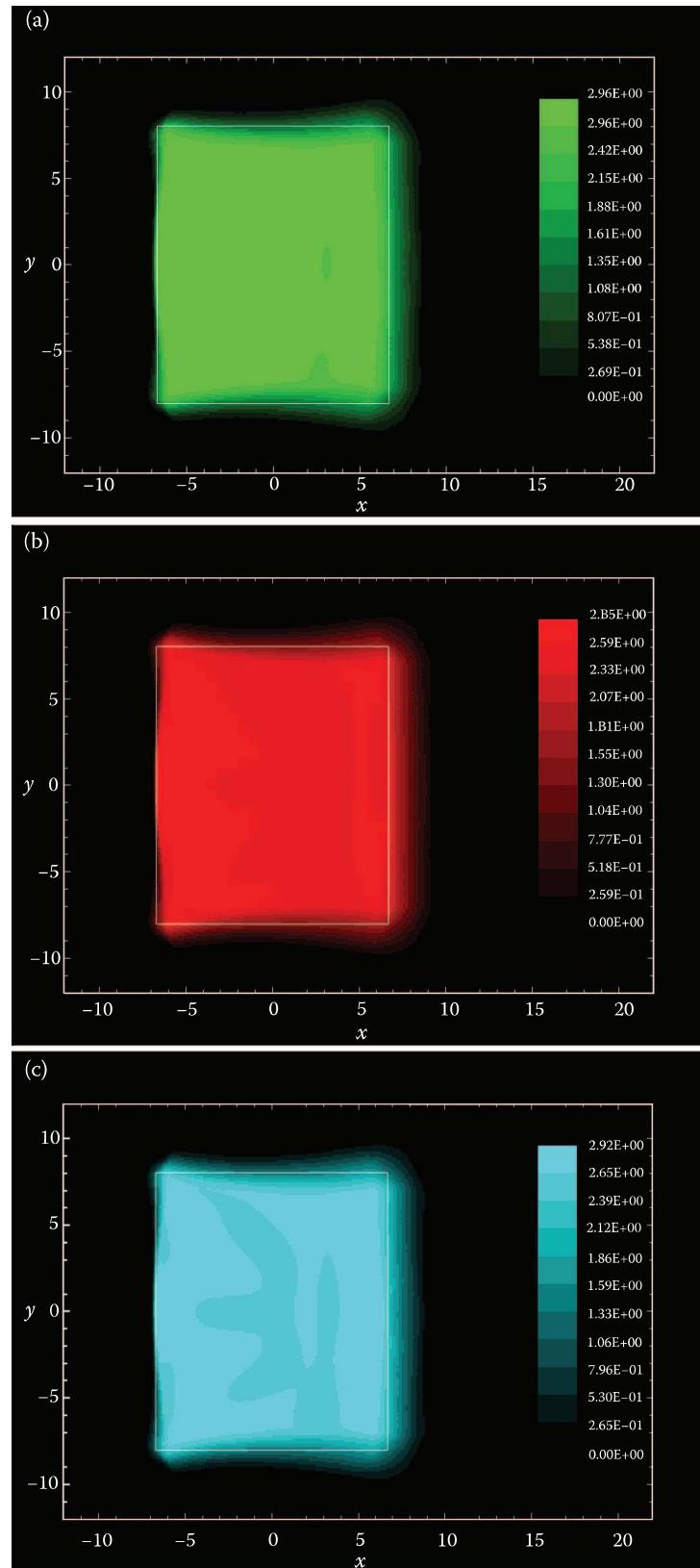


FIGURE A6.11 Graded colour plots of irradiance distribution with hologram outline for (a) green, 530 nm; (b) red, 630 nm; and (c) blue, 455 nm. Case of 500 mm \times 600 mm hologram (landscape format) illuminated at 45° from overhead at a distance of 500 mm. Finite-source calculations for the case of nine diodes. A change in graded colour is equivalent to a 10% change in irradiance. Length scales in relative units.

A6.5.2 Case 2

The next case that we shall study is the case of a side-illuminated hologram measuring 808 mm (H) \times 888 mm (W). The central angle of incidence is 70° to the vertical and the distance of the source to the centre of the hologram is 1306.246 mm. This geometry allows a large reflection hologram to be mounted in a relatively thin box using a flat side mirror to extend the ray paths (Figure A6.12).

Again, we use the exact same diodes and light guides as described in the previous section. We calculate a new lens surface and new optimal offset positions for the current geometry using exactly the same process as in the previous case, again using an eighth order expansion, except for one detail: we set the target contour to be 10% larger than the actual hologram size. Because the angle of incidence is extreme in the present case, this allows us to accept some decrease in irradiance towards the edges. Figure A6.13

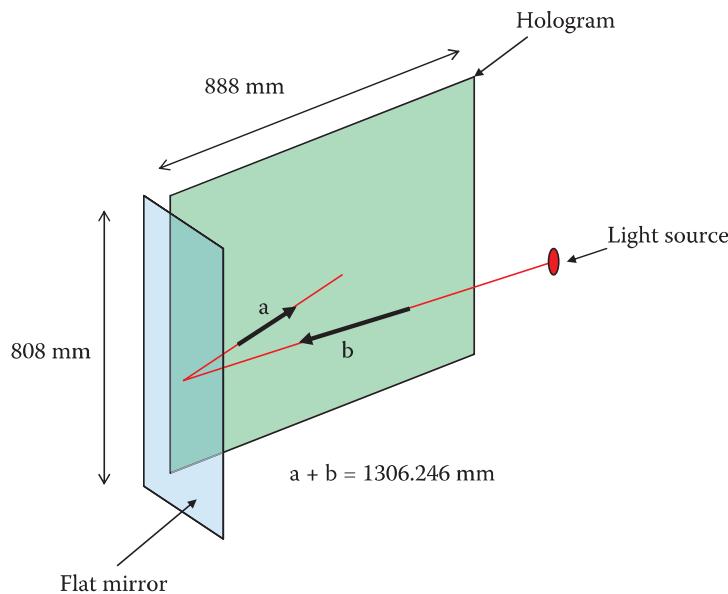


FIGURE A6.12 Hologram (888 mm \times 808 mm) side-illuminated by close light source and plane mirror at large angle of incidence allowing a small depth enclosed display.

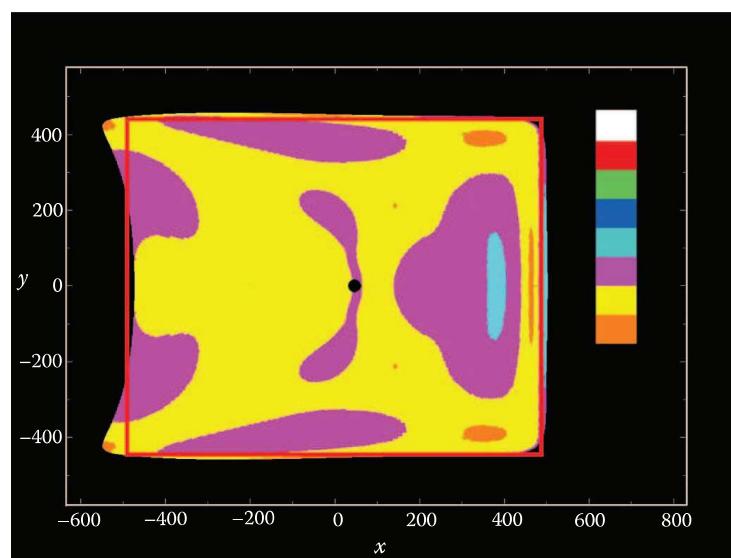


FIGURE A6.13 False-colour plot of the (far field) irradiance distribution on the target plane with target contour shown. Case of 808 mm \times 888 mm hologram (landscape format) illuminated at 70° from the side at a distance of 1306.246 mm. Length scales in relative units. Power scale: change in colour represents a change in irradiance by 12.5%. Calculations assume a single point source at (0,0,0) and a wavelength of 530 nm.

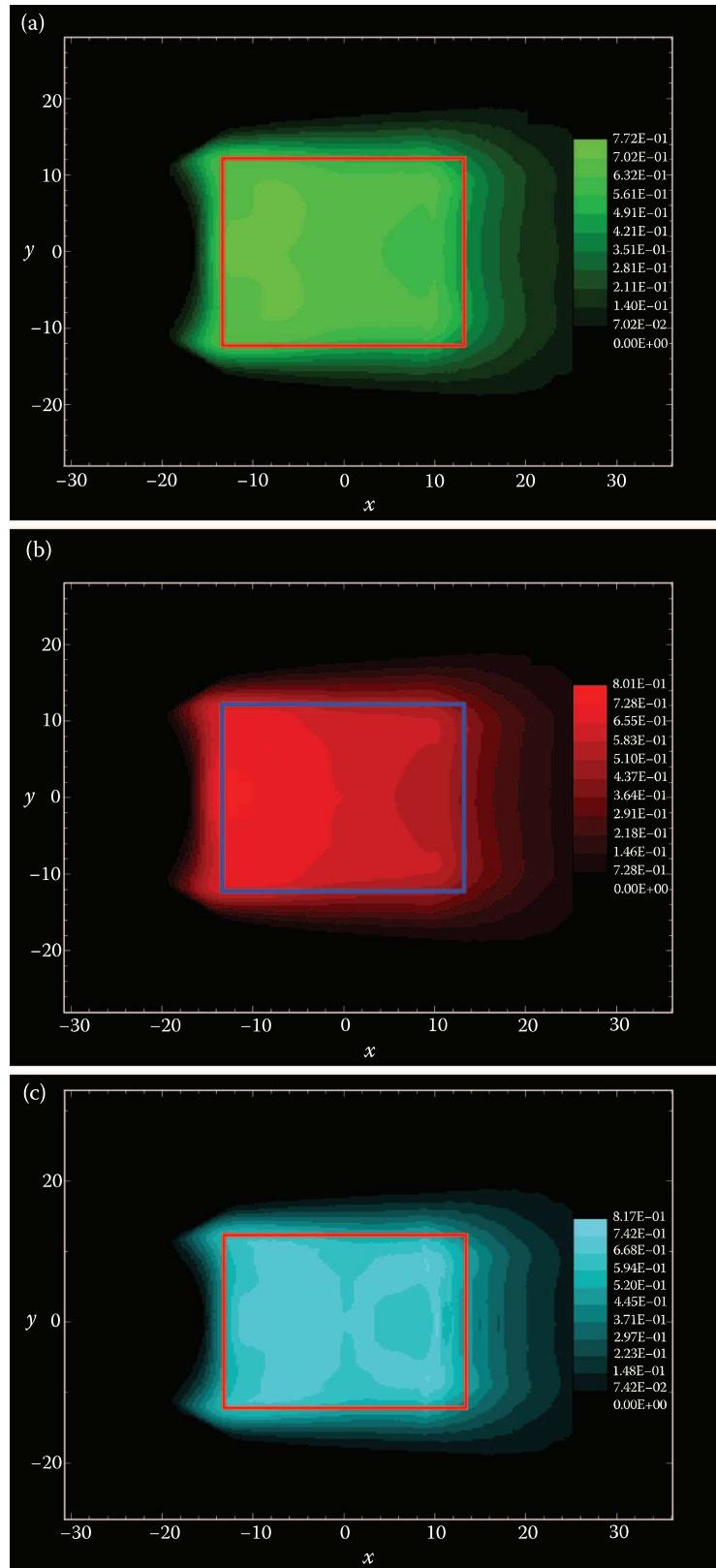


FIGURE A6.14 Graded colour plots of irradiance distribution with hologram outline for (a) green, 530 nm (b); red, 630 nm; and (c) blue, 455 nm. Case of 808 mm × 888 mm hologram (landscape format) illuminated at 70° from the side at a distance of 1306.246 mm. Finite-source calculations for the case of nine diodes. A change in graded colour is equivalent to a 10% change in irradiance. Length scales in relative units.

shows a false-colour map of the calculated irradiance distribution at the target plane, together with an outline of the target contour (note that this is now 10% larger than the hologram size). Different colours indicate a difference in irradiance of 12.5%. Clearly, the result is still good in that the irradiance is relatively flat within the contour and the contour is still well fitted. However, it is clear that the results are not as good as we saw in case 1. This simply shows the limitation of the physics; the 70° angle of incidence is indeed an extreme angle.

Figure A6.14a through c show graded colour plots of the irradiance at the target plane for the three colours using finite-source ray tracing according to Section A6.4.7. The hologram size is also drawn. Differences in shade indicate a 10% difference in irradiance. Clearly, the result is still extremely good for a practical light source. A difference of approximately $\pm 20\%$ in irradiance is apparent within the hologram contour and this falls well into what would be acceptable for a commercial source. Energetic efficiencies for the green, red and blue spectra are, respectively, 72%, 69% and 71%. For the geometry in question, these are excellent figures. Calculations of the spatial coherence, as measured from multiple points on the hologram surface, indicate that the effective size of the source is approximately 10% to 25% of its diameter. By running many cases on the computer, we see that we may accept a reduction in energetic efficiency in return for a flatter irradiance distribution.

A6.5.3 Case 3

The final case that we shall study is one in which a non-rectangular distribution of light is required on the target plane. In particular, we consider a variation of the geometry of case 2, in which the light source is moved up and down while always pointing towards the centre of the hologram. Such a variable geometry can be used to replay repetitive motion in a special type of digital hologram. As such, a reflection hologram may be mounted in a relatively thin box and a light source, which illuminates the hologram from the side and moves up and down continuously, may be incorporated therein: as the light moves, different holographic images are then replayed. This is shown in Figure A6.15. The geometry is the same as for case 2 except that the light source now travels up 47.5 cm and down 47.5 cm.

The geometry of the present case corresponds identically with the geometry of case 2 when the light source is in position A (Figure A6.15). However, if we calculate a lens surface that illuminates the hologram exactly at this position, then when the light source moves to the upper position (position B), clearly not all of the hologram will now be illuminated. The same will be the case at the bottom position (C). To

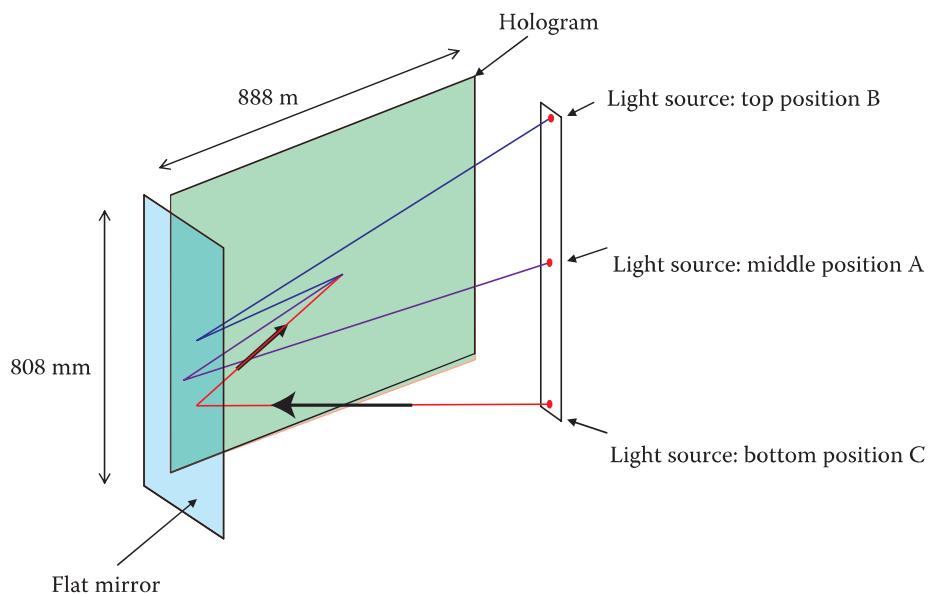


FIGURE A6.15 Hologram (888 mm × 808 mm) side-illuminated by a close light source on a motorised rail allowing motion of the light source in the vertical direction while the light source always points to the centre of the hologram.

choose a target contour that will generally fit the hologram at all light source positions, we use the target contour shown in Figure A6.16.

We now apply the process already described to this case. Figure A6.17 shows the calculated point-source irradiance distribution together with the target contour. Again, different colours indicate a difference in irradiance of 12.5%.

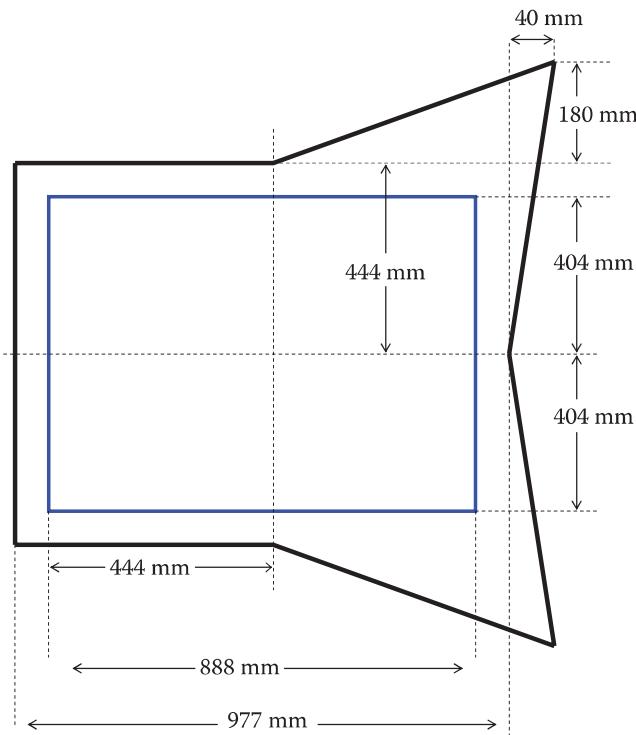


FIGURE A6.16 Target contour used to design a light source appropriate for the geometry of Figure A6.15.

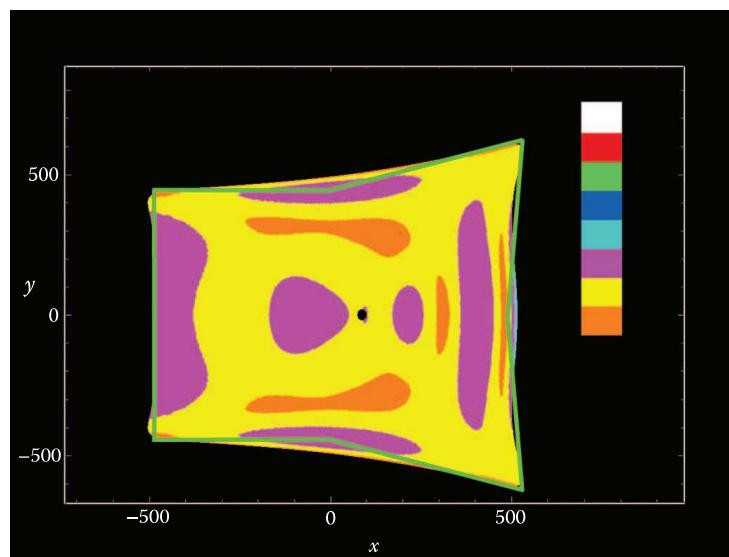


FIGURE A6.17 False-colour plot of the (far field) irradiance distribution on the target plane with target contour of Figure A6.16 shown. Case of 808 mm \times 888 mm hologram (landscape format) illuminated at 70° from the side at a distance of 1306.246 mm. Length scales in relative units. Power scale: change in colour represents a change in irradiance by 12.5%. Calculations assume a single point source at (0,0,0) and a wavelength of 530 nm.

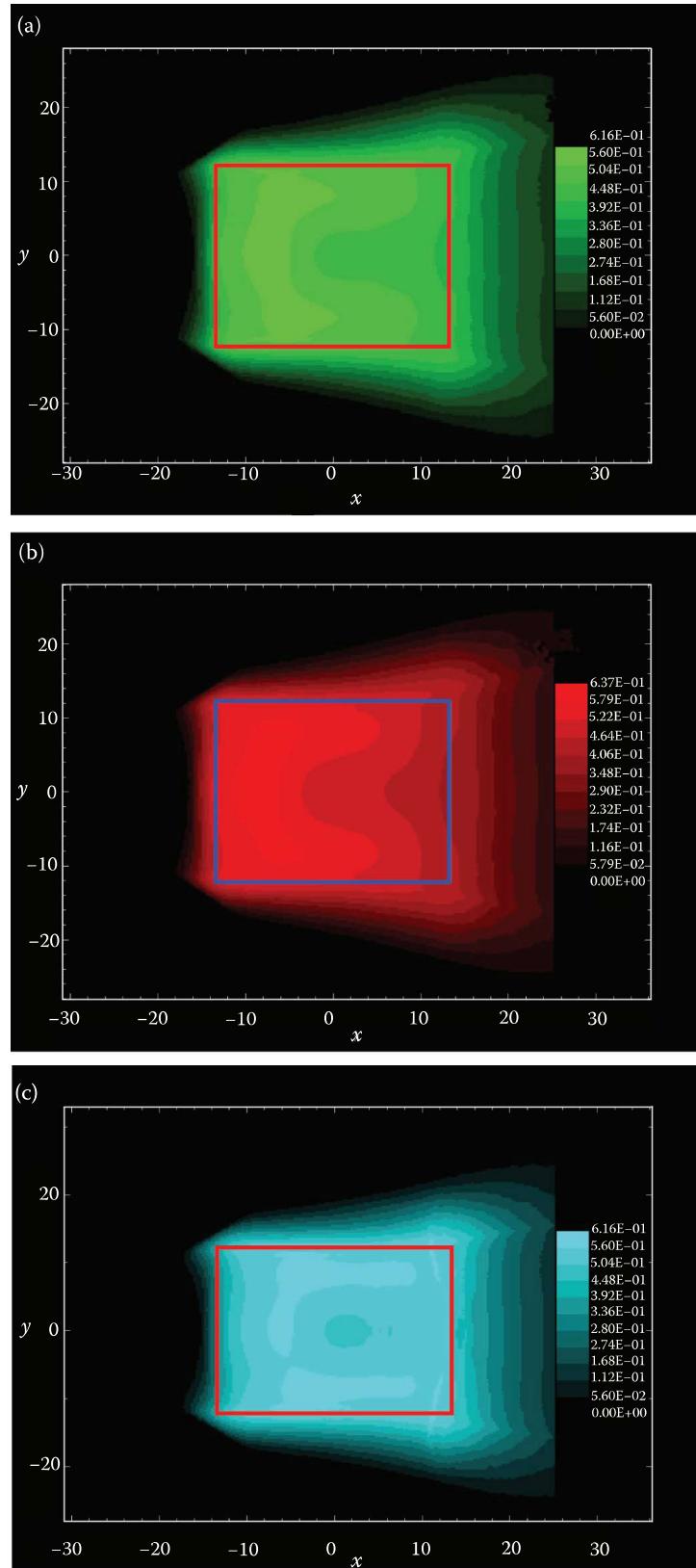


FIGURE A6.18 Graded colour plots of irradiance distribution with hologram outline for (a) green, 530 nm; (b) red, 630 nm; (c) and blue, 455 nm. Case of 808 mm × 888 mm hologram (landscape format) illuminated at 70° from the side at a distance of 1306.246 mm with target contour of Figure A6.16. Finite-source calculations for the case of nine diodes. A change in graded colour is equivalent to a 10% change in irradiance. Length scales in relative units.

Figure A6.18a through c show graded colour plots of the irradiance at the target plane for the three colours using finite-source ray tracing according to Section A6.4.7. The hologram size is also drawn. Differences in shade indicate a 10% difference in irradiance. Clearly, the result is again extremely good for a practical light source. As in case 2, a difference of approximately $\pm 20\%$ in irradiance is apparent within the hologram. Energetic efficiencies for the green, red and blue are, respectively, 61%, 58% and 59%. These are calculated using the real hologram size for a central light source position.

A6.6 Commercial Design of Nine-Diode Framing Light Source

Figure A6.19 shows a commercial design of the light source described conceptually in Figure A6.3. The light pipes and the LEDs are held together in the correct place by a mirrorised plastic holder. The holder touches the light pipes only at sufficient locations to hold and orient it securely, and the mirrored surface ensures that very little light is lost by doing this. The LEDs clip in from the back of the plastic holder and there is optical gel between the light pipes and the surface of the LEDs. The optical gel is matched in refractive index to be the square root of the sum of the squares of the indices of the glass in the light pipe and the surface of the LED. The gel is non-setting and capable of withstanding the surface temperature of the LED dice. In this way, no stress is transmitted from the light pipes to the dice even when there is significant thermal expansion of the dice. The plastic holders in turn push into the lower and upper glass lens parts and are held in place by optical cement, which again is index-matched as described previously. The refractive index of the light pipe glass is chosen so as to provide the minimum light pipe size. The neutral filter fits into the case, and the top glass lens, the dichroic plate and the lower lens are located in the neutral filter and the case that aligns them. Again, optical gel or cement is used between the parts to ensure correct optical performance. The plastic light pipe holders mutually support each other and are in contact with the top lens so that they may be simply pushed down into the case. Heat sink grease is used to make sure there is good thermal contact between the LEDs and the case. The active terminals on the LEDs, which are on the same side as the thermal pad, are arranged to be clear of the case so that an electrical connection can be made.

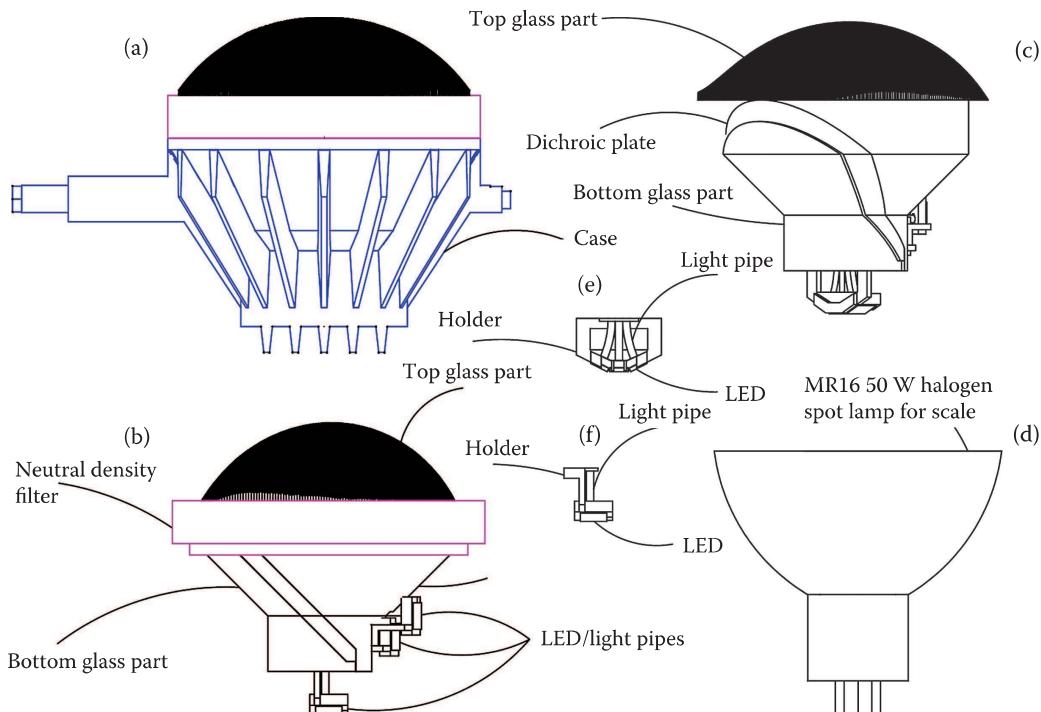


FIGURE A6.19 Elevations of a commercial compact RGB light source (a–c) based on nine LEDs. Details of the LED and light-guide mounting are shown in (e) and (f). A light source (MR16) from previous art is shown to scale for comparison (d).

If appropriate, the top lens can be made into two nesting parts so that a simple change of the top lens will allow field reconfiguration of the beam pattern for many common situations. This feature is something that is impossible with fixed-beam lights such as MR16s.

An important feature of the present design is that the fall-off in the light distribution outside the desired coverage area can be made to be much sharper than that of traditional lamps. These typically have a distribution that falls off in a Gaussian manner, whereas it is easy with the present design to produce a top hat distribution. This means that the light pattern is much more even within the desired area and much less light falls outside that area. It also means that the present design may be used to illuminate abutting holographic panels where different lights illuminate different panels. Previous light sources usually cannot do this.

A scale comparison between the present light and an MR16 lamp is also shown in Figure A6.19d. From this, it can be seen that the physical size of this device is very comparable with an MR16, while at the same time providing all the advantages of power, efficiency, longevity and light distribution pattern described above.

The design lends itself easily to scaling up in size. For instance, 12 diodes may be used instead of 9 in a package broadly the same size as presented here. A version with 30 diodes producing an emission of approximately 3600 lm would be just under twice the size of a standard MR16 when optimised for the display of highly tilted holograms. A version with 90 diodes producing approximately 10,875 lm and optimised for general illumination with an approximately circular footprint would again only be two to three times the size of a standard MR16. This latter example could well be used, when married with a telescopic lens and optional aperture, to replace traditional theatre projectors, which currently use 575-W compact filament halogen lamps giving averages of approximately 12,000 lm (e.g., Source Four Junior Zoom Stage and Studio projector with HPL 575-W CN/240-V bulb made by Ushio) with lifetimes of just 400 h.

A6.7 Additional Considerations

Although the LEDs described are currently state-of-the-art in terms of light output, one of the characteristics of such LEDs is that their lifetime and light output decline significantly with an increase in junction temperature. In the example shown in Figure A6.19, the back of each LED is therefore held firmly against the outer case by the support box. This allows for a good conduction path to the outer case, which is typically die-cast black anodised aluminium and has fins and holes to allow for improved heat dissipation from both the interior and exterior surfaces. This results in a higher light output and a longer life than would otherwise be the case. These particular LEDs have an operating life of up to 100,000 h, which is more than an order of magnitude better than the previous light sources. This is especially important where the light sources are relatively inaccessible for replacement as, for example, in the case where they are used in consumer appliances or in widely dispersed products such as point-of-sale or advertising uses.

Where the light source illuminating a hologram is required to move, it may become difficult to arrange for interconnecting wires. An advantage of the present invention is that because of the low wattages required by the LEDs, it is practical and possible to use a high-frequency electromagnetic coupling between the power source and the LEDs. This removes the necessity for interconnecting wires.

A6.8 Reflective Light-Source

The present light source can also be designed to use a reflector rather than a non-spherical refracting lens. Radiation from one or more diodes may still be delivered to a main reflecting optic via light guides, exactly as we have discussed in previous sections. The shape of the reflector may be parameterised as a polynomial spectral expansion exactly as in the case of a lens. The problem of reflector design may then be cast as a constrained optimisation using one or more Lagrange multipliers exactly as with the refracting lens case; the equations we derived earlier are not changed much except that the vector form of Snell's law must be replaced by the vector reflection law. The numerical solution is then essentially identical.

One advantage of using a reflective geometry is that the vector reflection law does not depend on wavelength. This solution is then better suited to applications in which a truly broad spectrum emission is required and “white” diodes may be employed. As with the design of the lens surface described above, the shape of the reflecting surface is best established by the use of a point-source approximation. Then, the best feeder positions for the light guides are established by ray tracing the case in which all finite-sized light guides are included.

Generally, there exist a number of obvious choices of how to inject light from one or more light guides into a reflective collector. Certain choices (such as where the light guides point themselves in the direction of the target) lead to relatively compact configurations but with the disadvantage of poorer spatial coherence. Other configurations (such as where the light pipes emerge from the centre base of the lamp and are then bent symmetrically to emit radiation at a significant angle to the direction of the target) lead to slightly less compact configurations but with a better spatial coherence.

Appendix 7: Bilinear and Bicubic Interpolation

A7.1 Introduction

As we have seen in Chapters 8 and 9 and also in Appendix 4, digital data are often not available at the exact pixel locations of the image data spatial light modulators (SLMs) in a direct-write digital holography (DWDH) printer. Rather, due to aberration of one kind or another, or non-compliance with the various quantisation rules, exact data are available only at varying distances between the actual pixels. One solution here is simply to use a “nearest integer” interpolation.* Suppose we have a function, $S(x_i, y_i)$, and we want to estimate this function at integer values of x and y . If the data set (x_i, y_i) we are given is dense (in the sense that we always have data near the integer values, i.e., there are no big gaps) but does not contain these integer values themselves, then in nearest integer interpolation, we simply use the estimate $S(\|x_i\|, \|y_i\|)$. So if we know $S(156.1, 293.3)$, then we use this as an approximation to $S(156, 293)$. The problem with this technique is that it creates sudden changes of intensity in S , which leads to interpolation noise in digital holograms. Another way to deal with the situation is to use a single **I-to-S** transformation that implicitly incorporates all sources of aberration. This avoids the use of sequential nearest integer operations.

In this appendix, we briefly mention two alternative techniques to the nearest integer interpolation. These are bilinear interpolation and bicubic interpolation. Both seek to minimise abrupt changes in intensity by creating a two-dimensional analytical model of intensity surrounding each pixel. Both techniques are used with great success in conventional image processing and form a vital part of the digital holographer’s toolkit.

A7.2 Bilinear Interpolation

Bilinear interpolation is an extension of ordinary linear interpolation which is used to interpolate functions of two variables on a regular grid. The interpolated function uses the bilinear product xy . Bilinear interpolation can be broken down into two operations. The first is a standard linear interpolation in one direction, and the second is another standard linear interpolation but in the orthogonal direction. Although each of these steps is linear in the sampled values and in the positions, the composite bilinear interpolation is not linear but quadratic in the sample location. Bilinear interpolation can be trivially extended to trilinear interpolation in three dimensions.

Suppose that we wish to find the value of an unknown function g at a point (x, y) . Suppose further that the value of g is known at the four points surrounding this point. Let us label these points as

$$\begin{aligned} & (x_1, y_1) \\ & (x_1, y_2) \\ & (x_2, y_1) \\ & (x_2, y_2) \end{aligned} \tag{A7.1}$$

where

$$\begin{aligned} & x_2 > x > x_1 \\ & y_2 > y > y_1 \end{aligned} \tag{A7.2}$$

* Also commonly referred to as nearest neighbour interpolation.

The first step is a linear interpolation in the x direction. This gives

$$g(x, y_1) \sim \frac{x_2 - x}{x_2 - x_1} g(x_1, y_1) + \frac{x - x_1}{x_2 - x_1} g(x_2, y_1) \quad (\text{A7.3})$$

and

$$g(x, y_2) \sim \frac{x_2 - x}{x_2 - x_1} g(x_1, y_2) + \frac{x - x_1}{x_2 - x_1} g(x_2, y_2) \quad (\text{A7.4})$$

Next, we use these estimates to interpolate in the y direction:

$$g(x, y) \sim \frac{y_2 - y}{y_2 - y_1} g(x, y_1) + \frac{y - y_1}{y_2 - y_1} g(x, y_2) \quad (\text{A7.5})$$

We can therefore write a single expression for an estimate of the function g at the point (x, y) as

$$\begin{aligned} g(x, y) &\sim \frac{y_2 - y}{y_2 - y_1} \frac{x_2 - x}{x_2 - x_1} g(x_1, y_1) + \frac{y_2 - y}{y_2 - y_1} \frac{x - x_1}{x_2 - x_1} g(x_2, y_1) \\ &+ \frac{y - y_1}{y_2 - y_1} \frac{x_2 - x}{x_2 - x_1} g(x_1, y_2) + \frac{y - y_1}{y_2 - y_1} \frac{x - x_1}{x_2 - x_1} g(x_2, y_2) \end{aligned} \quad (\text{A7.6})$$

In the context of digital holography, we are usually concerned with estimating the SLM data $\mathbf{S}_{\alpha\beta} = \mathbf{S}(x_\alpha, y_\beta)$ in terms of image data ${}^{kg}\mathbf{I}_{ij}$ through index laws of the form

$$\begin{aligned} i &= X(\alpha, \beta) \\ j &= Y(\alpha, \beta) \end{aligned} \quad (\text{A7.7})$$

where X and Y are real functions. We therefore require an estimation of the image data ${}^{kg}\mathbf{I}_{ij}$ at non-integer values of i and j . Equation 7.6 can then be used to provide this:

$$\begin{aligned} I(i, j) &\sim \frac{\lceil j \rceil - j}{\lceil j \rceil - \lfloor j \rfloor} \frac{\lceil i \rceil - i}{\lceil i \rceil - \lfloor i \rfloor} I(\lfloor i \rfloor, \lfloor j \rfloor) + \frac{\lceil j \rceil - j}{\lceil j \rceil - \lfloor j \rfloor} \frac{i - \lfloor i \rfloor}{\lceil i \rceil - \lfloor i \rfloor} I(\lceil i \rceil, \lfloor j \rfloor) \\ &+ \frac{j - \lfloor j \rfloor}{\lceil j \rceil - \lfloor j \rfloor} \frac{\lceil i \rceil - i}{\lceil i \rceil - \lfloor i \rfloor} I(\lfloor i \rfloor, \lceil j \rceil) + \frac{j - \lfloor j \rfloor}{\lceil j \rceil - \lfloor j \rfloor} \frac{i - \lfloor i \rfloor}{\lceil i \rceil - \lfloor i \rfloor} I(\lceil i \rceil, \lceil j \rceil) \end{aligned} \quad (\text{A7.8})$$

A7.3 Bicubic Interpolation

Bicubic interpolation is a more sophisticated version of bilinear interpolation. Instead of taking into account only four points around the point where an interpolation is desired, sixteen points are now effectively considered.

Let us assume that we know the value of a function, f , and its derivatives at the four points of the unit square surrounding the point of interest. The derivatives are typically calculated by finite differences using the points immediately surrounding these four points (making sixteen in total). Then, within the unit square, a smooth interpolated surface may be defined:

$$g(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j \quad (\text{A7.9})$$

The task is to find the coefficient matrix a_{ij} . Matching function values at the four principal points yields

$$\begin{aligned} f(0,0) &= g(0,0) = a_{00} \\ f(1,0) &= g(1,0) = \sum_{i=0}^3 a_{i0} \\ f(0,1) &= g(0,1) = \sum_{j=0}^3 a_{0j} \\ f(1,1) &= g(1,1) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} \end{aligned} \quad (A7.10)$$

In addition, matching the derivatives at the four points yields a further eight conditions, which are

$$\begin{aligned} f_x(0,0) &= g_x(0,0) = a_{10} \\ f_x(1,0) &= g_x(1,0) = a_{10} + 2a_{20} + 3a_{30} \\ f_x(0,1) &= g_x(0,1) = \sum_{j=0}^3 a_{1j} \\ f_x(1,1) &= g_x(1,1) = \sum_{i=0}^3 \sum_{j=0}^3 ia_{ij} \end{aligned} \quad (A7.11)$$

and

$$\begin{aligned} f_y(0,0) &= g_y(0,0) = a_{01} \\ f_y(1,0) &= g_y(1,0) = \sum_{i=0}^3 a_{i1} \\ f_y(0,1) &= g_y(0,1) = a_{01} + 2a_{02} + 3a_{03} \\ f_y(1,1) &= g_y(1,1) = \sum_{i=0}^3 \sum_{j=0}^3 ja_{ij} \end{aligned} \quad (A7.12)$$

Finally, matching cross-derivatives yields the final four necessary conditions:

$$\begin{aligned} f_{xy}(0,0) &= g_{xy}(0,0) = a_{11} \\ f_{xy}(1,0) &= g_{xy}(1,0) = a_{11} + 2a_{21} + 3a_{31} \\ f_{xy}(0,1) &= g_{xy}(0,1) = a_{11} + 2a_{12} + 3a_{13} \\ f_{xy}(1,1) &= g_{xy}(1,1) = \sum_{i=0}^3 \sum_{j=0}^3 ija_{ij} \end{aligned} \quad (A7.13)$$

These relations form a system of linear equations that enable the matrix a_{ij} to be calculated by Gaussian elimination.

Appendix 8: Rigorous Coupled Wave Theory of Simple and Multiplexed Gratings

A8.1 Introduction

Moharam and Gaylord [1] were the first to show how coupled wave (CW) theory could be formulated without approximation. This led to a simple computational algorithm that could be used to solve the wave equation exactly. Although earlier approaches such as the Modal method [2] were also rigorous, they involved the solution of a transcendental equation for which a general unique algorithm could not be defined. This came in contrast to the simple Eigen formulation proposed by Moharam and Gaylord. Here we provide a derivation of the rigorous coupled wave (RCW) theory for the more general spatially multiplexed grating. We shall then show how the resulting equations simplify to Moharam and Gaylord's equations for the simple phase grating. Rather than solving these equations using an Eigen method, we employ an alternative approach using Runge–Kutta integration.*

A8.2 Derivation of RCW Equations

For brevity, we shall limit discussions to the lossless case with isotropic permittivity[†] and we shall employ the σ -polarisation for which the Helmholtz equation may be written:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - \gamma^2 u = 0 \quad (\text{A8.1})$$

where u is the transverse (z) electric field and the parameter

$$\gamma^2 = -\beta^2 - 2\beta \sum_{\mu=1}^N \kappa_{\mu} \left\{ e^{i\mathbf{K}_{\mu} \cdot \mathbf{r}} + e^{-i\mathbf{K}_{\mu} \cdot \mathbf{r}} \right\} \quad (\text{A8.2})$$

defines the multiplexed grating.[‡] Following the notation established in Chapter 12, we now consider the case of illumination of the grating by a wave of the form

$$u(y < 0) = e^{i(k_x x + k_y y)} \quad (\text{A8.3})$$

where

$$\begin{aligned} k_x &= \beta \sin(\theta_{c\mu} - \psi_{\mu}) = \beta \sin(\Phi_c) \\ k_y &= \beta \cos(\theta_{c\mu} - \psi_{\mu}) = \beta \cos(\Phi_c) \end{aligned} \quad \forall \mu \quad (\text{A8.4})$$

* This may be programmed with exceptional ease using Mathematica from Wolfram Research Inc using the NDSolve function.

[†] In 1989, Glytsis and Gaylord extended the RCW theory to anisotropic media and spatially multiplexed gratings.

[‡] This is just the polychromatic version of Equation 12.107.

In both the front region ($y < 0$) and the rear region ($y > d$), the average index is assumed to be n_0 .

Now, the Helmholtz field $u(x, y)$ may be consistently expanded as follows:

$$\begin{aligned} u(x, y) &= \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \sum_{l_3=-\infty}^{\infty} \dots u_{l_1 l_2 l_3 \dots}(y) e^{i(k_x + l_1 K_{1x} + l_2 K_{2x} + \dots)x} \\ &= \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \sum_{l_3=-\infty}^{\infty} \dots u_{l_1 l_2 l_3 \dots}(y) e^{ik_x x} \prod_{\sigma=1}^N e^{il_{\sigma} K_{\sigma x} x} \end{aligned} \quad (\text{A8.5})$$

This expression may be substituted into Equations A8.1 and A8.2. On taking the Fourier transform and applying orthogonality, we then arrive at the following RCW equations:

$$\begin{aligned} &\left\{ \left(k_x + \sum_{\sigma=1}^N l_{\sigma} K_{\sigma x} \right)^2 - \beta^2 \right\} u_{l_1 l_2 l_3 \dots l_N}(y) - \frac{\partial^2 u_{l_1 l_2 l_3 \dots l_N}}{\partial y^2}(y) \\ &= 2\beta \sum_{\sigma=1}^N \kappa_{\sigma} \left\{ u_{l_1 l_2 l_3 \dots (l_{\sigma}-1) \dots l_N}(y) e^{iK_{\sigma y} y} + u_{l_1 l_2 l_3 \dots (l_{\sigma}+1) \dots l_N}(y) e^{-iK_{\sigma y} y} \right\} \end{aligned} \quad (\text{A8.6})$$

A8.3 Simplification in the Case of Simple Non-Multiplexed Grating

For the case of the simple sinusoidal grating, the transformation

$$u_l(y) = \hat{u}_l(y) e^{i(k_y + lK_y)y} \quad (\text{A8.7})$$

reduces Equation A8.6 to the more usual form

$$\begin{aligned} &\frac{\partial^2 \hat{u}_l(y)}{\partial y^2} + 2i(k_y + lK_y) \frac{\partial \hat{u}_l(y)}{\partial y} \\ &= \left\{ (k_x + lK_x)^2 + (k_y + lK_y)^2 - \beta^2 \right\} \hat{u}_l(y) - 2\beta \kappa \left\{ \hat{u}_{l-1}(y) + \hat{u}_{l+1}(y) \right\} \end{aligned} \quad (\text{A8.8})$$

used by Moharam and Gaylord [1].

A8.4 Derivation of Boundary Conditions

In the zones in front of and behind the grating where $\kappa_{\sigma} = 0$, Equations A8.6 is reduced to the simpler constant index equations:

$$\left\{ \sum_{\sigma=1}^N (k_x + l_{\sigma} K_{\sigma x})^2 - \beta^2 \right\} u_{l_1 l_2 l_3 \dots l_N}(y) - \frac{\partial^2 u_{l_1 l_2 l_3 \dots l_N}}{\partial y^2}(y) = 0 \quad (\text{A8.9})$$

These equations define which l modes can propagate in the exterior regions. They have simple solutions of the form

$$u_{l_1 l_2 l_3 \dots l_N} = A_{l_1 l_2 l_3 \dots l_N} e^{i \sqrt{\beta^2 - \sum_{\sigma=1}^N (k_x + l_{\sigma} K_{\sigma x})^2} y} + B_{l_1 l_2 l_3 \dots l_N} e^{-i \sqrt{\beta^2 - \sum_{\sigma=1}^N (k_x + l_{\sigma} K_{\sigma x})^2} y} \quad (\text{A8.10})$$

where the square roots are real for undamped propagation.* Accordingly, we may deduce that the front solution comprising the illumination wave and any reflected modes must be of the form

* Note that there are modes that propagate inside the grating but which show damped propagation outside.

$$u(x, y) = e^{ik_x x} e^{i\sqrt{\beta^2 - k_x^2} y} + \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \sum_{l_3=-\infty}^{\infty} \dots \sum_{l_N=-\infty}^{\infty} u_{l_1 l_2 l_3 \dots l_N} e^{-i\sqrt{\left(\beta^2 - \sum_{\sigma=1}^N (k_x + l_{\sigma} K_{\sigma x})^2\right)} y} e^{i\left(k_x + \sum_{\sigma=1}^N l_{\sigma} K_{\sigma x}\right)x} \quad (\text{A8.11})$$

Likewise, the rear solution comprising all transmitted modes must be of the form

$$u(x, y) = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \sum_{l_3=-\infty}^{\infty} \dots \sum_{l_N=-\infty}^{\infty} u_{l_1 l_2 l_3 \dots l_N} e^{i\sqrt{\left(\beta^2 - \sum_{\sigma=1}^N (k_x + l_{\sigma} K_{\sigma x})^2\right)} y} e^{i\left(k_x + \sum_{\sigma=1}^N l_{\sigma} K_{\sigma x}\right)x} \quad (\text{A8.12})$$

By demanding continuity of the tangential electric field and the tangential magnetic field at the boundaries $y = 0$ and $y = d$, we may now use these expressions to define the boundary conditions required for a solution of Equation A8.1 within the multiplexed grating. At the front surface, these are

$$\begin{aligned} i\sqrt{\beta^2 - k_x^2} (2 - u_{000\dots}(0)) &= \frac{du_{000\dots}}{dy} \Big|_{y=0} \\ -i\sqrt{\beta^2 - (k_x + l_1 K_{1x} + l_2 K_{2x} + \dots)^2} u_{l_1 l_2 l_3 \dots}(0) &= \frac{du_{l_1 l_2 l_3 \dots}}{dy} \Big|_{y=0} \end{aligned} \quad (\text{A8.13})$$

And at the rear surface, they take the form

$$i\sqrt{\beta^2 - (k_x + l_1 K_{1x} + l_2 K_{2x} + \dots)^2} u_{l_1 l_2 l_3 \dots}(d) = \frac{du_{l_1 l_2 l_3 \dots}}{dy} \Big|_{y=d} \quad (\text{A8.14})$$

The modes available for external (undamped) propagation are calculated using the condition

$$\beta^2 > (k_x + l_1 K_{1x} + l_2 K_{2x} + \dots)^2 \quad (\text{A8.15})$$

Note, however, that there are internal propagating modes that nevertheless do not propagate outside the grating, and these must be retained.

A8.5 Numerical Solution of RCW Equations

Moharam and Gaylord [1] solved the single grating equations (Equation A8.8) using a state-space formulation in which solutions are obtainable through the eigenvalues and eigenvectors of an easily defined coefficient matrix. However, as mentioned above, we can also solve the more general equations (Equation A8.6), subject to the boundary conditions in Equations A8.13 and A8.14, using simple Runge–Kutta integration. This is a practical method as long as the number of component gratings within the multiplexed grating is relatively small.

Diffraction efficiencies of the various modes are defined as

$$\eta_{l_1 l_2 l_3 \dots} = \frac{\sqrt{\beta^2 - (k_x + l_1 K_{1x} + l_2 K_{2x} + l_3 K_{3x} + \dots)^2}}{k_y} u_{l_1 l_2 l_3 \dots} u_{l_1 l_2 l_3 \dots}^* \quad (\text{A8.16})$$

where the fields in this equation are defined either at the front boundary in the case of reflected modes or at the rear boundary in the case of transmitted modes. Note that we are treating the lossless case here, so the sum of all transmitted and reflected efficiencies totals to unity.*

* In the case of the front reflected 000... mode, one uses $\eta_{000\dots} = \frac{\sqrt{\beta^2 - k_x^2}}{k_y} (u_{000\dots} - 1)(u_{000\dots} - 1)^*$.

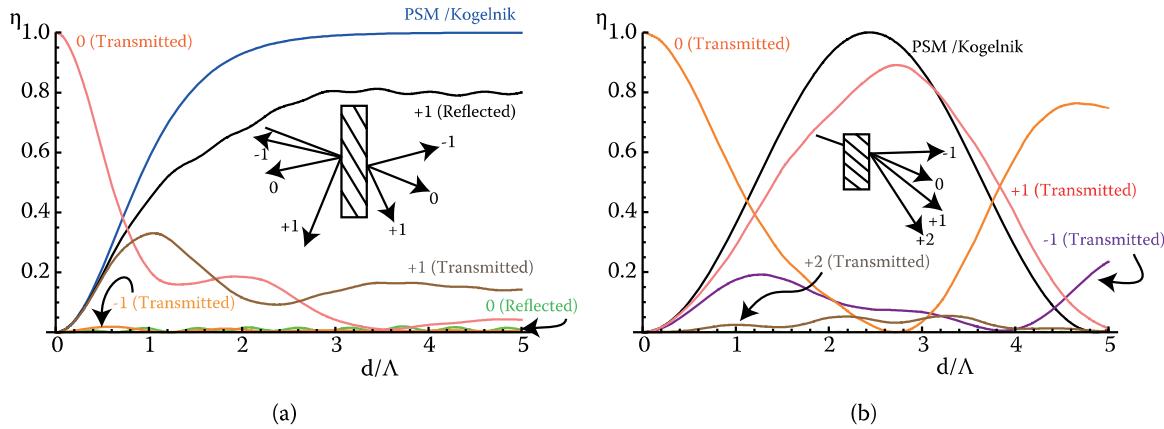


FIGURE A8.1 Diffraction efficiency (η_0) versus normalised grating thickness according to RCW theory and compared to the PSM and Kogelnik theories at Bragg resonance for (a) the simple reflection grating ($n_0 = 1.5$, $n_1/n_0 = 0.331/2$, $\theta_c = \theta_r = 50^\circ$, $\psi = 30^\circ$, $\lambda_c = \lambda_r = 532$ nm) and (b) the simple transmission grating ($n_0 = 1.5$, $n_1/n_0 = 0.121/2$, $\theta_c = \theta_r = 80^\circ$, $\psi = 60^\circ$, $\lambda_c = \lambda_r = 532$ nm).

A8.5.1 Comparison of Kogelnik's Theory and PSM Theory with RCW Theory

Equation A8.6, subject to the boundary conditions in Equations A8.13 and A8.14, is solved by Runge–Kutta integration. This permits the rigorous calculation of the diffraction efficiencies of all modes, which are produced by a general grating. Figure A8.1 shows an example for a simple reflection grating and a simple transmission grating at Bragg resonance.* In the case of the reflection grating, a very high index modulation has been assumed. Nevertheless, the Kogelnik/parallel stacked mirror (PSM) estimation is still only 20% out, and it is clear that most of the “dynamics” of the grating are associated with the +1 reflected mode as both PSM and Kogelnik’s CW theories assume. In the case of the transmission hologram, a relatively high index modulation is assumed and also a large incidence angle with respect to the grating planes. Here we see again only a small departure from the Kogelnik/PSM estimation but also the presence of the +2 mode.

A8.5.2 Comparison of N-PSM Theory with RCW Theory

In order to compare the N-PSM model [4] developed in Chapter 12 with the RCW theory, we investigate the typical spatially multiplexed grating, which is illustrated in Figure A8.2. This grating is composed of two simple gratings that have been sequentially recorded in the same material using the same laser wavelength of 532 nm and the same incidence angle of 30° . The component gratings have differing slants and different grating constants and give rise to the multiplexed grating structure shown in Figure A8.2c.

In Figure A8.3, the diffraction efficiency at Bragg resonance as determined by the N-PSM model† is compared for different grating thicknesses and different index modulations to an RCW calculation. Typically, 14 modes are retained in the Runge–Kutta integration, the higher order modes being many orders of magnitude smaller than the lower ones. Only the reflected modes are plotted. Figure A8.3a shows an extreme case where an index modulation of $n_1 = 0.3$ for each of the two component gratings of Figure A8.2c is assumed. It is clear that in this case, there are quite important differences between the N-PSM model and the RCW model. In addition, quite a few of the higher order modes allowed under the RCW theory start to oscillate at non-negligible amplitudes, including modes that require both gratings present to propagate. However, this is indeed an extreme case, and one might have anticipated this from the results of Moharam and Gaylord [1]. As the index modulation of each grating drops in Figure A8.3b through d, we see a better and better agreement between the two models. Figure A8.3d represents

* Note that at Bragg resonance, the PSM and Kogelnik models give the same predictions.

† Conventional N-CW theory [5] gives an identical prediction to N-PSM at Bragg resonance here.

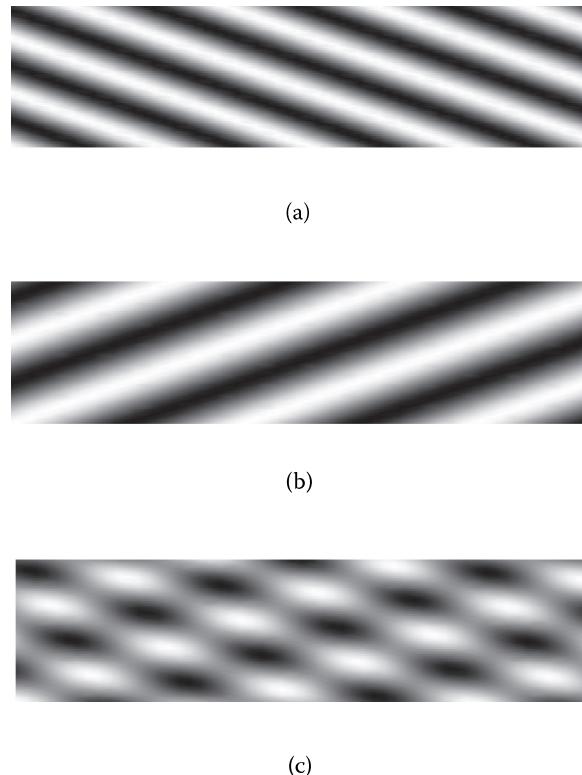


FIGURE A8.2 An example of a spatially multiplexed phase reflection grating. The grating, whose (x,y) index distribution is shown in (c) is formed by the sequential recording of the two simple gratings shown in (a) and (b). Each diagram shows a section of size $0.5 \mu\text{m}$ by $2 \mu\text{m}$. Each simple grating has been recorded with a reference beam angle of $\Phi_c = 30^\circ$ and with a wavelength of 532 nm . One grating has a slope of $\psi_1 = 20^\circ$ and the other has a slope of $\psi_2 = -20^\circ$. Note that the form of the multiplexed grating in (c) is fundamentally different from the characteristic linear shape of its component simple gratings of (a) and (b). Note also that identical index modulations for each of the two component gratings have been assumed in (c).

a typical multiplexed grating made using a modern material such as photopolymer or dichromated gelatine. The N-PSM and RCW models therefore produce extremely good agreement here, and higher order modes only account for less than 1% of the total diffraction.

In Figure A8.4, a comparison of the N-PSM model and the RCW theory for the off-Bragg case is presented. In particular, the diffractive efficiency of the multiplexed grating of Figure A8.2c is investigated as the replay wavelength is changed whilst keeping the replay angle of incidence fixed at 30° . Although a complex analytic solution of the N-PSM equations is available for this problem, for pure convenience, one can solve the N-PSM equations numerically using Runge–Kutta integration. In Figure A8.4a, an index modulation for each of the component gratings of $n_1 = 0.03$ is used; this would be rather typical for diffractive elements made from photopolymer or dichromated gelatin. Agreement between the two theories is clearly excellent, particularly in the primary diffractive band. In Figure A8.4b and c, we plot for comparison the N-PSM diffractive efficiencies when one or another of the two component gratings is removed from the multiplexed element.

It is interesting to note that the diffractive efficiency does not peak in Figure A8.4a for both components at the Bragg angle. This is because the dominant grating within the multiplexed element depletes the reference wave disproportionately around resonance. As the wavelength is changed away from resonance, there rapidly becomes more reference wave available within the grating for signal generation by the second component grating. Even though the intrinsic efficiency of this second grating drops away from Bragg resonance, the increased reference wave left by the first grating more than compensates for this, leading to the characteristic hollow curve with symmetric peaks away from resonance.

Figure A8.4d repeats the case of Figure A8.4a but with individual index modulations of $n_1 = 0.15$. Here the differences between the N-PSM and RCW solutions are somewhat larger as might be expected.

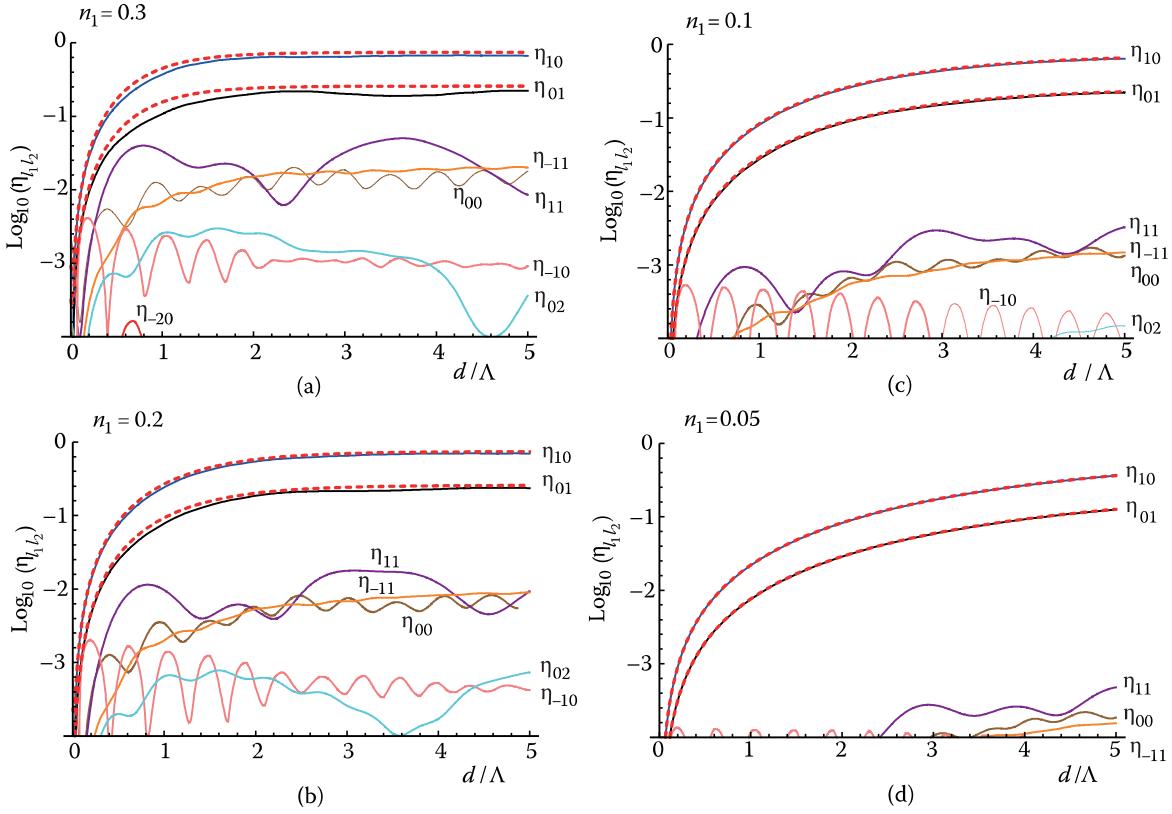


FIGURE A8.3 Diffractive efficiency, η_{σ} , versus normalised grating thickness, d/Λ , as predicted by the PSM model and by an RCW calculation for the case of the twin multiplexed reflection grating of Figure A8.2c at Bragg resonance. The grating is replayed using light of 532 nm at an incidence angle of $\Phi_e = 30^\circ$. The grating index modulation of each of the component twin gratings has been taken to be $n_1 = 0.3$ in (a), $n_1 = 0.2$ in (b), $n_1 = 0.1$ in (c), and $n_1 = 0.05$ in (d). In each case, the average index inside and outside the grating has been set to $n_0 = 1.5$. The dotted lines indicate the S_1 and S_2 modes of the PSM model and the full lines indicate the modes of the RCW calculation. The most prominent RCW modes are the 01 and 10 modes, which correspond to the S_1 and S_2 modes in PSM. Note that Λ refers to the larger of the two grating periods—i.e. to that of the component grating shown in Figure A8.2b.

A8.5.3 Comparison of N-PSM with RCW Theory for Multicolour Gratings

In Chapter 12, we derived simple equations describing the diffractive response of a multicolour spatially multiplexed grating using the N-PSM model. Here we compare this theory with the RCW theory derived in this appendix for the case of a complex grating composed of three component subgratings as shown in Figure A8.5a through c. All gratings are recorded using sequential exposure and a reference incidence angle of 30° . The grating slants are, respectively, -10° , 0° , and 15° . The first two gratings are recorded at 532 nm whilst the last grating is recorded at 660 nm. Index modulations for the three gratings are chosen to be 0.03, 0.02, and 0.035.

In order to reduce computational time for the multidimensional Runge–Kutta integration, we restrict the modes available in the RCW calculation to only, -1 , 0 , and 1 for each of the three l integers. Selective introduction of higher order modes is then used to check convergence. Figure A8.6a shows a graph of the results for the dominant 100 and 010 RCW modes and the corresponding S_1 , S_2 N-PSM modes for replay at 532 nm. Figure A8.6b shows the corresponding graph for 660 nm where the dominant RCW mode is the 001 mode corresponding to the S_3 mode of N-PSM. Clearly, agreement between the two theories is excellent. The higher order modes of the rigorous theory are typically less than a few percent at these index modulations. At higher modulation, however, we would again expect to see differences between the two theories, as per Figure A8.3. What is important though is that the index modulations of Figure A8.6 are typical for modern holographic techniques and that we can therefore be rather optimistic that

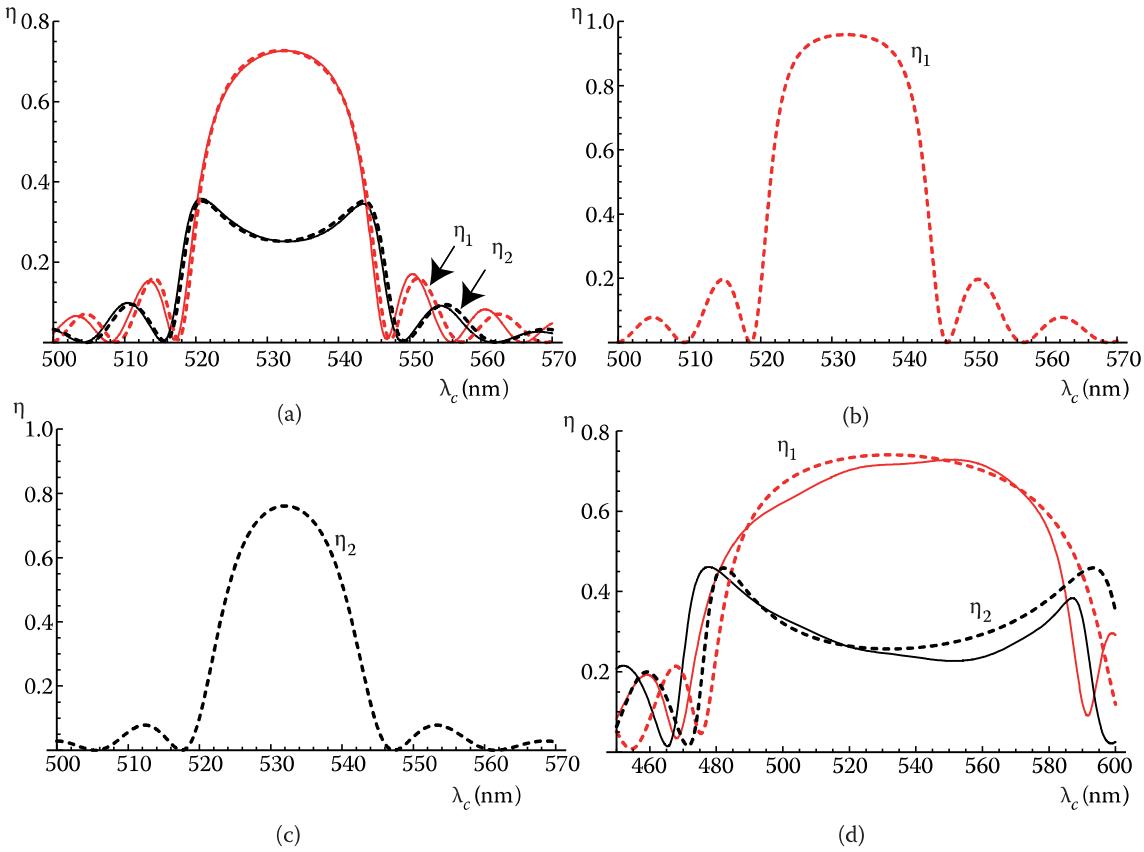


FIGURE A8.4 Diffractive efficiency, η_c , versus replay wavelength, λ_c , as predicted by the PSM model and by an RCW calculation for the case of the twin multiplexed reflection grating of Figure A8.3c at and away from Bragg resonance. The grating is illuminated at its recording angle of $\Phi_c = 30^\circ$. The average index inside and outside the grating has been set to $n_0 = 1.5$. In (a), a grating index modulation of $n_1 = 0.03$ for each component grating is assumed and a grating thickness of $d = 7 \mu\text{m}$ is used. The dotted lines indicate the S_1 and S_2 modes of the PSM model, and the full lines indicate the corresponding 10 and 01 modes of the RCW calculation. Panels (b) and (c) show the PSM diffractive response expected when one or the other of the component gratings is deleted from the diffractive element. Panel (d) shows a similar case to (a) but with higher index modulations ($n_1 = 0.15$ for each grating) and with smaller thickness, $d = 2 \mu\text{m}$.

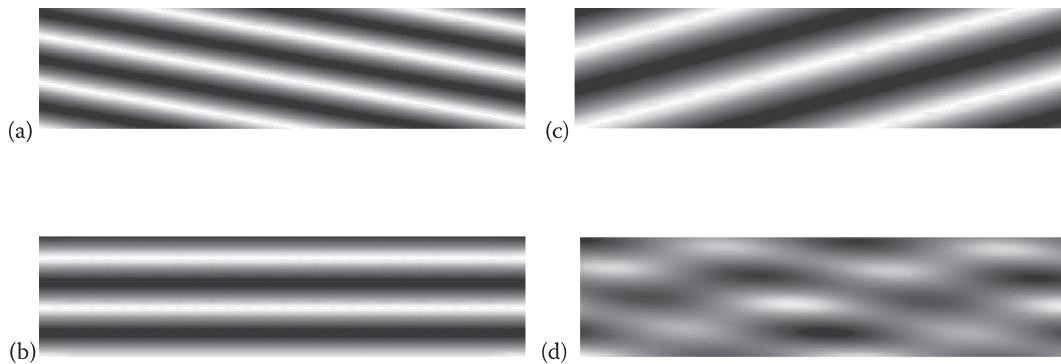


FIGURE A8.5 Example of a panchromatic spatially multiplexed phase reflection grating. The grating, whose (x,y) index distribution is shown in (d), is formed by the sequential recording of the three simple gratings shown in (a) and (b). Each diagram shows a section of size $0.5 \mu\text{m} \times 2 \mu\text{m}$. Each simple grating has been recorded with a reference beam angle of $\Phi_c = 30^\circ$. Gratings 1 and 2 have been recorded at 532 nm whereas grating 3 has been recorded at 660 nm. The gratings have slopes of $\psi_1 = -10^\circ$, $\psi_2 = 0^\circ$, and $\psi_3 = 15^\circ$ and index modulations of 0.03, 0.02, and 0.035, respectively.

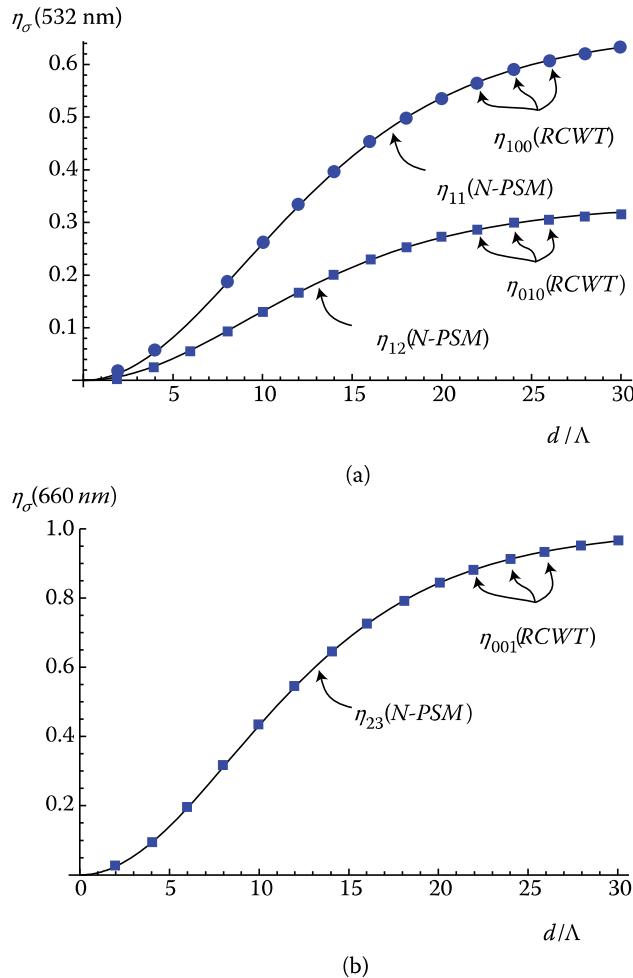


FIGURE A8.6 Diffractive efficiency, η_σ , at 532 nm and 660 nm versus normalised grating thickness, d/Λ , as predicted by the N-PSM model and by an RCW calculation for the case of the triple multiplexed bi-colour reflection grating of Figure A8.5d at Bragg resonance. The grating is replayed using laser light at 532 nm and 660 nm at an incidence angle of $\Phi_c = 30^\circ$. In each case, the average index inside and outside the grating has been set to $n_0 = 1.5$. The full lines indicate the modes of the N-PSM model, and the markers indicate the modes of the RCW calculation. By far, the most prominent RCW modes are the 100 and 010 modes at 532 nm and the 001 mode at 660 nm. Note that Λ refers to the larger of the three grating periods—i.e. to that of the component grating shown in Figure A8.5c.

the N-PSM model at Bragg resonance (or equivalently the N-CW theory of Solymar [5]) is capable of accurately and conveniently describing the multiplexed panchromatic volume phase reflection grating.

REFERENCES

1. M. G. Moharam and T. K. Gaylord, “Rigorous coupled wave analysis of planar grating diffraction,” *J. Opt. Soc. Am.* **71**, 811–818 (1981).
2. R. S. Chu and J. A. Kong, “Modal theory of spatially periodic media,” *IEEE T. Microw. Theory*, MTT-25, (1977) 18–24.
3. E. N. Glytsis and T. K. Gaylord, “Rigorous 3D coupled wave diffraction analysis of multiple superposed gratings in anisotropic media,” *Appl. Opt.* **28**, 2401–2421 (1989).
4. D. Brotherton-Ratcliffe, “Analytical treatment of the polychromatic spatially multiplexed volume holographic grating,” *Appl. Opt.* **51**, 7188–7199 (2012).
5. L. Solymar, “Two-dimensional N-coupled-wave theory for volume holograms,” *Opt. Commun.* **23**, 199–202 (1977).

Appendix 9: Recent Developments

A9.1 New Equipment, Materials, Techniques and Applications

After the manuscript of this book was submitted for typesetting, inevitably new advances in equipment, techniques, and applications came to the authors' attention. Before publication, we were pleased to have the opportunity to include this final appendix, which we believe brings the book properly up to date as of October 2012.

In Chapter 2, we described Lippmann photography and discussed interferential structures in nature. In September 2012, an interesting paper was published revealing that such structures exist not only in insects and butterflies but also in plants. Vignolini et al. [1] present a striking example of multilayer-based strong iridescent colouration in plants, in the fruit of *Pollia condensata* (Figure A9.1). The fruit contains helicoidally stacked cellulose microfibrils that form multilayers in the cell walls of the epicarp. Because the multilayers form with both helicoidicitics, optical characterisation reveals that the reflected light from every epidermal cell is polarised circularly either to the left or to the right, a feature that has never been previously observed.

Important papers were presented at the *9th International Symposium on Display Holography* (ISDH 2012) in June 2012 at MIT, which also need to be mentioned here. The reader is referred to the conference website [2] where video presentations [3] and proceedings can be accessed. Zebra Imaging Inc. described a new simple way of capturing 3D information intended for digital colour holograms. Zebra is now offering a printing service based on the technique for people interested in making their own digital colour holograms.

In association with the ISDH conference, there was a hologram exhibition arranged at the MIT museum. One of the exhibited holograms that drew particular attention was a digital achromatic reflection hologram, printed by Geola. The hologram was a new version of the holographic portrait of Queen Elizabeth II, created by the artist Chris Levine and holographer Rob Munday, and demonstrated clearly that high-quality black-and-white holograms can be made using the DWDH technique described in this book. The Queen's portrait was also featured on a recent postage stamp from Jersey Post.

Progress in solid-state CW and pulsed lasers has been rapid, and accordingly, there are now various improved laser systems on the market that are suitable for recording colour holograms.

Progress in recording materials and recording techniques has been made on two fronts. Firstly, a new recording principle for colour holograms based on *surface plasmon waves* has been reported, although more development work will be needed before the technique can become a practical recording method for colour holograms. Additional progress in recording materials has come from Geola, where high efficiency DWDH holograms have recently been written onto photoresist using single nanosecond pulses at energies 10 times smaller than those usually required using conventional CW exposure.

Holografika introduced a new improved HoloVizio™ 3D holographic system at the *Siggraph 2012* Las Vegas exhibition in August 2012. The technique behind Holografika's products has been mentioned in Chapter 14, but the recent products and improvements are worth including here. The company SeeReal in Germany has also reported progress in their VISIO 20 real-time 3D holographic display system.



FIGURE A9.1 Fruits of *Pollia condensata* conserved in the Herbarium collection at the Royal Botanic Gardens, Kew, United Kingdom. Material collected in Ethiopia in 1974 and preserved in alcohol-based fixative. (Vignolini, S. et al. 2012. Pointillist structural color in *Pollia* fruit. *Proceedings of the National Academy of Sciences of the United States, PNAS* 109 (39), 15712–15715. Copyright [2012] National Academy of Sciences, U.S.A.)

A9.2 Progress at Zebra Imaging

At ISDH 2012, Craig Newswanger of Zebra Imaging Inc. [4] presented a new simplified technique to capture and print digital colour holograms. Zebra's paper [5] outlined several methods based on *photogrammetry* for producing 3D content for holograms, software applications for editing, positioning and lighting. Photogrammetry is the practice of determining the geometric properties of objects from photographic images. A number of free software tools have become available recently that can extract 3D data from collections of photographs using automated photogrammetry. Examples include commercially available tools such as Autodesk 123D Catch [6]. These types of tools can provide surprisingly accurate 3D models from a variety of cameras, including cell phones, tablets, and point and shoot cameras, and promise to simply revolutionise image acquisition for full-parallax digital holography.

Zebra Imaging has developed the Zscape™ print technology based on the hogel concept. The combination of the easy capture method and a free authoring tool plus an easy web interface for ordering a Zscape™ print finally brings holographic technology "within the reach of the masses" according to Newswanger. 123D Catch (originally named *PhotoFly*) is offered for free from Autodesk. It takes in a set of photos of an object or scene and develops a polygonal model from automatic photogrammetric analysis. It is very simple to learn and builds a model starting from a few to a hundred photographs. More photographs do not necessarily result in a better model. The lion example shown below was based on only 17 photos, sufficient to produce a high-quality model.

The desktop version of 123D Catch is intended for photographs taken with a typical digital camera. By employing the software, it is possible to capture photographs and process the 3D model in the same tool. The desktop tool can export an ".obj" file, which is a ubiquitous file format that is compatible with many 3D tools. In Figure A9.2, 17 captured images of a lion sculpture (from Asolo, Italy) are shown.

By employing Zebra's Zscape™ Preview software, it is possible to display a realistic preview of a potential hologram, including viewing angles, blur from the image plane, lighting, and resolution. Preview is not a modeller but a hologram composition tool that allows the user to scale and position multiple polygonal and point cloud data sets in the same hologram. It is possible to import multiple 3D data models and scale and position them prior to printing, for example, to choose the hologram size, colour, and format that best fit the data. At Zebra's website, it is easy to send the 3D information and to order a hologram. Figure A9.3 shows the hologram preview at the Zebra website. A photograph of the final hologram of the lion sculpture is shown in Figure A9.4.

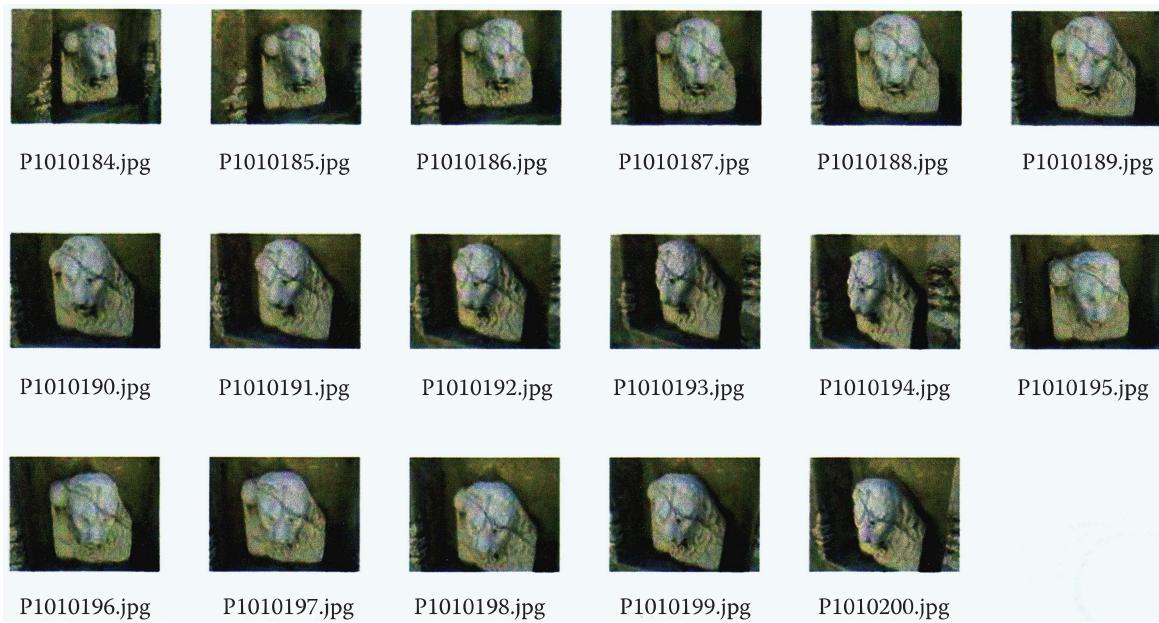


FIGURE A9.2 Collection of 17 photographs of a sculpted stone lion. Two passes were made at different heights. (Courtesy of Zebra Imaging Inc.)

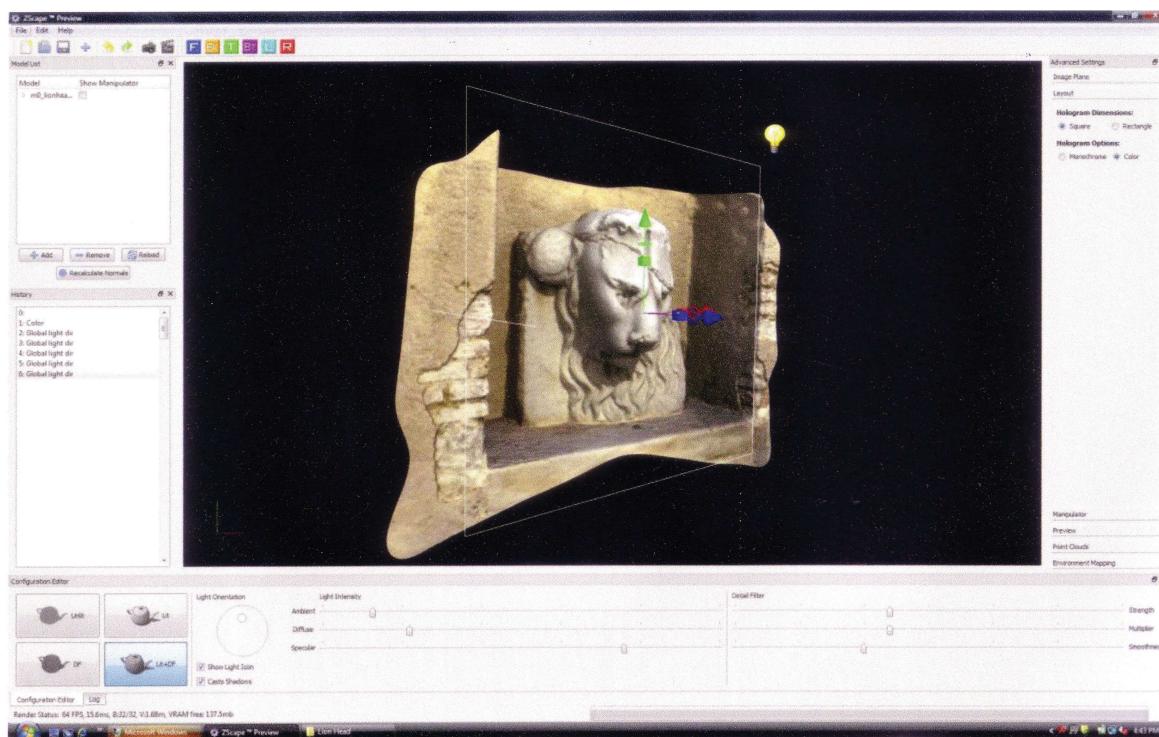


FIGURE A9.3 The “obj” file export from 123D Catch is dropped into Zscape Preview. Note the white frame that indicates the image plane of the hologram. (Courtesy of Zebra Imaging Inc.)

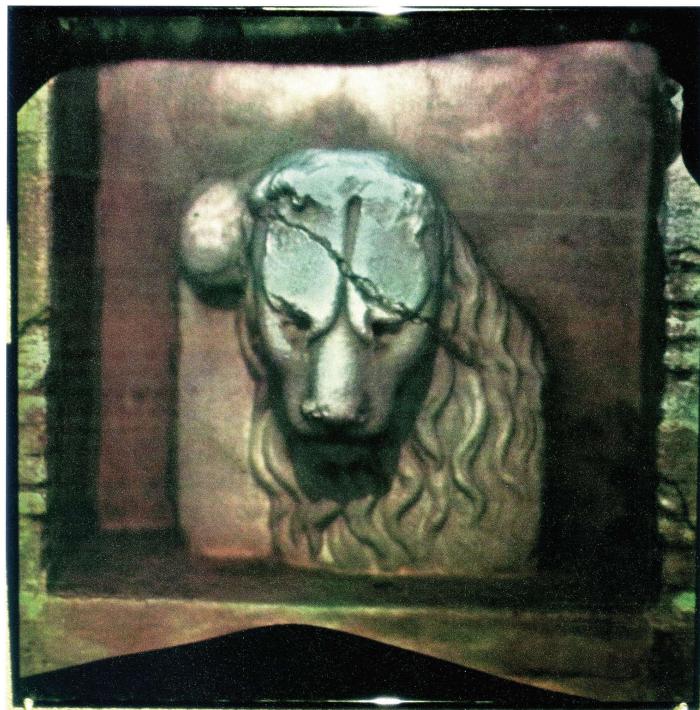


FIGURE A9.4 Photograph of a 60 cm × 60 cm DWDH hologram of the lion sculpture. (Courtesy of Zebra Imaging Inc.)

A9.3 The Queen Elizabeth II Portrait and the Jersey Postage Stamp

A9.3.1 Achromatic Portrait of the Queen

The hologram portrait of Queen Elizabeth II, mentioned in Chapter 1, was commissioned by Jersey Heritage Trust. The first hologram portrait *Equanimity* was created by the artist Chris Levine and holographer Rob Munday. The recording of the portrait took place on 24 March 2004 in the Yellow Drawing Room at Buckingham Palace in London. Rob Munday installed the 3D recording equipment with a moving camera that travels along a horizontal rail for recording a sequence of photographic images to be used for the holographic stereogram. In total, 38 sequences were shot, each consisting of 205 frames and each recording taking 8 seconds. The Queen adopted her inimitable regal pose and remained perfectly still for each of the 38 recordings.

A new version of the portrait, *The Diamond Queen*, by Rob Munday was on display at the ISDH 2012 hologram exhibition at the MIT museum. It was a large format (72 cm × 90 cm) one-step HPO digital achromatic reflection hologram, which was printed at the Geola facility in Vilnius using a pulsed RGB laser and the DWDH technique with a hogel size of 0.8 mm. A photograph of the portrait hologram is shown in Figure A9.5.

A9.3.2 Hologram Postage Stamp of the Queen

Jersey Post issued on 1 June 2012 an embossed version of the Queen hologram portrait on a £10 postage stamp. The value and lettering on the stamp was produced in flat simili silver foiling. This is the first time that a 3D holographic portrait has been used on a postage stamp. The embossed hologram for the stamp was created especially for Jersey Post by Rob Munday, and the stamps were printed by Cartor Security Printing in France. A photograph of the stamp is shown in Figure A9.6.



FIGURE A9.5 Photograph of the 72 cm × 90 cm achromatic DWDH portrait hologram, *The Diamond Queen*, on display at the ISDH 2012 MIT hologram exhibition. (Hologram: Joint creative collaboration between C. Levine, R. Munday, and the Jersey Heritage Trust Commemoration. Photo: Courtesy of R. Munday. Printed at Geola.)



FIGURE A9.6 Jersey Post *Queen Elizabeth II* hologram postage stamp. (Hologram: Joint creative collaboration between C. Levine, R. Munday, and the Jersey Heritage Trust Commemoration.)

A9.4 Lasers for Colour Holography

The progress in solid-state CW and pulsed RGB lasers has been fast—in particular regarding output power.

A9.4.1 New Lasers from Cobolt

The range of ultralow-noise DPSS lasers manufactured by Cobolt AB, Solna, Sweden, has been extended to cover higher power red, green, and blue CW lasers. The lasers, which are suitable for colour holography, are based on the company's 05-01 Series platform. The new Flamenco™ laser operates at a wavelength of 660 nm with an output power of 0.5 W. The new Samba™ laser operates at a wavelength of 532 nm with an output power up to 1.5 W. Cobolt has also a blue DPSS laser, the Calypso™ laser, which operates at a wavelength of 491 nm with an output power of 200 mW. The 491 nm wavelength is, however, not suitable for recording colour holograms.

The Cobolt proprietary laser cavity design provides ultralow-noise performance of typically <0.1% rms and a spectral linewidth of <1 MHz. This extremely narrow linewidth can correspond to a coherence length of more than 10,000 m. The lasers provide a TEM₀₀ spatial mode and are manufactured using the proprietary HTCure™ technology in a compact and hermetically sealed package, which provides a very high level of immunity to varying environmental conditions and exceptional reliability. A compact controller (CDRH or OEM) is supplied with all the lasers, allowing remote operation and monitoring over RS-232 or USB interfaces.

A9.4.2 New Lasers from Coherent

CW lasers from Coherent (including the high-power Verdi laser) were described in Chapter 3. Here we mention the new range of Genesis MX SLM-Series high-power optically pumped semiconductor lasers (OPSL) with single longitudinal mode and TEM₀₀ spatial mode. The two Genesis lasers made by the company are a blue-wavelength laser—the MX 488 nm—with an output power of 1 W and a green-wavelength laser—the MX 532 nm—also with an output power of 1 W.

In addition to the Genesis lasers, Coherent can now also provide lasers with the following wavelengths: 460, 514, 532, 561 and 577 nm with output powers of 0.5 W and, in some cases, up to 1 W.

A9.4.3 New Lasers from Laser Quantum

The Torus SLM lasers from Laser Quantum UK, Stockport, UK, are suitable for colour holography. These lasers have a long coherence length (>100 m) and excellent mode stability provided by active locking. The red-wavelength laser—the Torus 660 nm—has an output power of 200 mW. The green-wavelength laser—the Torus 532 nm—has an output power of 750 mW.

Torus lasers are single-frequency lasers that use intelligent electronics to continually track the longitudinal mode position and to ensure that there are no mode-hops. The patented cavity guarantees single-frequency operation. The photons resonant in the cavity form a travelling wave, removing mode competition, which results in the laser operating at just one frequency. The bandwidths are <5 MHz (the red laser) and <1 MHz (the green laser), and both lasers have a TEM₀₀ spatial mode.

The Torus laser power supply is an integral part of each laser system. It has an LCD screen and monitors component temperatures in the laser head, automatically maintains laser output power, and provides diagnostic analysis.

A9.4.4 Progress in RGB Laser Technology at Geola

Geola Digital UAB has very recently extended its range of single-frequency lamp-pumped pulsed RGB lasers for holography. The RGB- α series (see Table 6.4 in Chapter 6) is now commercially available with both amplification and with the option of a mixing unit for applications that require a single white beam (Figure A9.7). Previously, the company had only offered its RGB lasers unamplified and with three

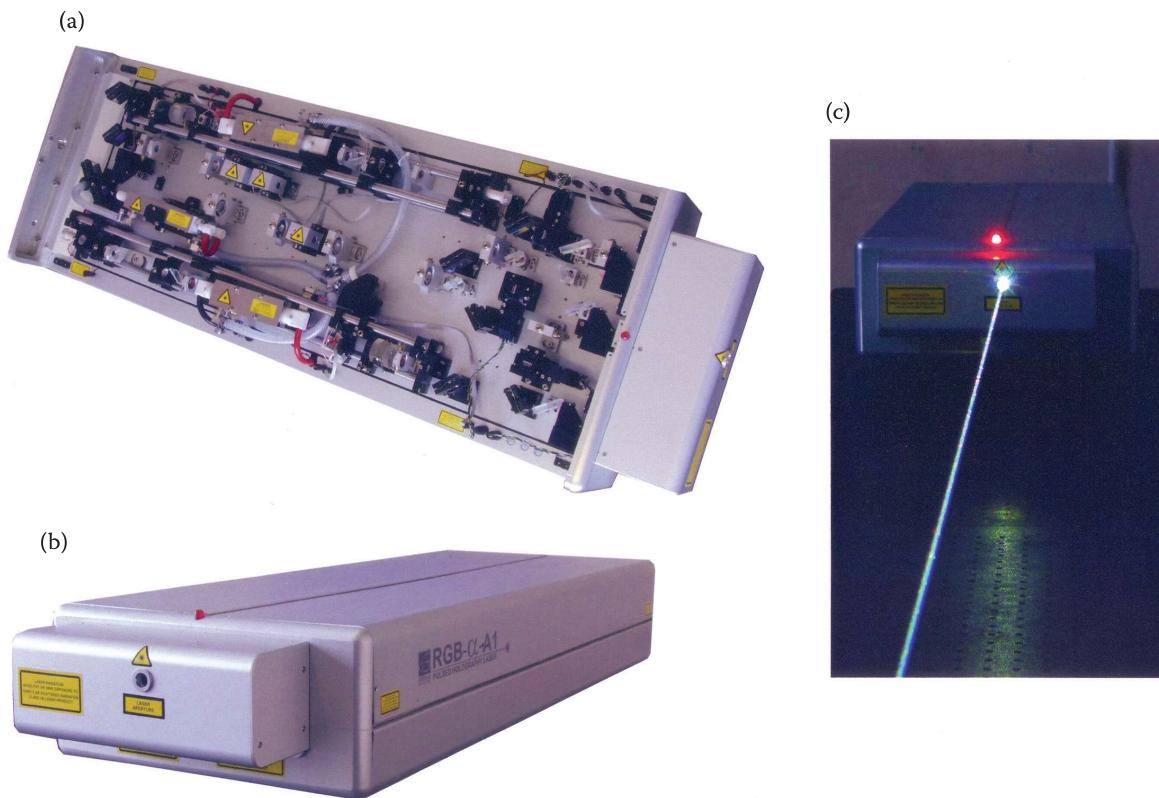


FIGURE A9.7 Several photographs of the new amplified RGB SLM lasers from Geola for holography applications: (a) view of interior optics, (b) external appearance, and (c) laser in action showing output white beam. The laser shown is the RGB- α -A1, which produces just over 0.5 W of white emission at 660 nm, 532 nm, and 440 nm. (Courtesy of Geola Digital UAB.)

separate red, green, and blue output beams. Currently, Geola is able to offer the highest energy output with Nd:YAG. Output wavelengths are 660, 532 and 440 nm. Pulse energies in each colour are up to 15 mJ at a repetition rate of 35 Hz, giving a white beam output of just over 1.5 W. All members of the RGB- α series are available with automatic cavity length stabilisation systems, making these lasers suitable for both digital (single exposure) and analogue (multiexposure) holography. Geola has also released single-colour versions of the technology with power outputs of up to 900 mW at 660 nm (red) or 700 mW at 440 nm (blue). The blue lasers have been extensively tested for the application of writing holograms onto photoresist (see Section A9.5.2).

A9.5 Progress in Recording Materials and Recording Techniques

A9.5.1 Surface Plasmon Waves

On the surface of a metal, there exists a form of slow light in the form of an evanescent wave that is associated with the collective oscillation of free electrons; this is known as a *surface plasmon polariton* (SPP). Such slow light can be used in a variety of applications in nanophotonic materials and devices. Holograms displayed using interactions between light and the collective oscillations of electrons on the surface of a metal are set to transform 3D imaging technology according to Satoshi Kawata and his team at the RIKEN Institute in Japan. They have constructed colour holographic images in a new type of recording material [7]. Plasmons are observed when electrons in a metal collectively oscillate at light wave frequencies. If light falls on a metal with a lower frequency than that characteristic of surface plasmons, it is reflected, while higher-frequency light is transmitted. The characteristic frequencies of surface plasmons propagating in gold and silver are in the visible range and are responsible for the metals'

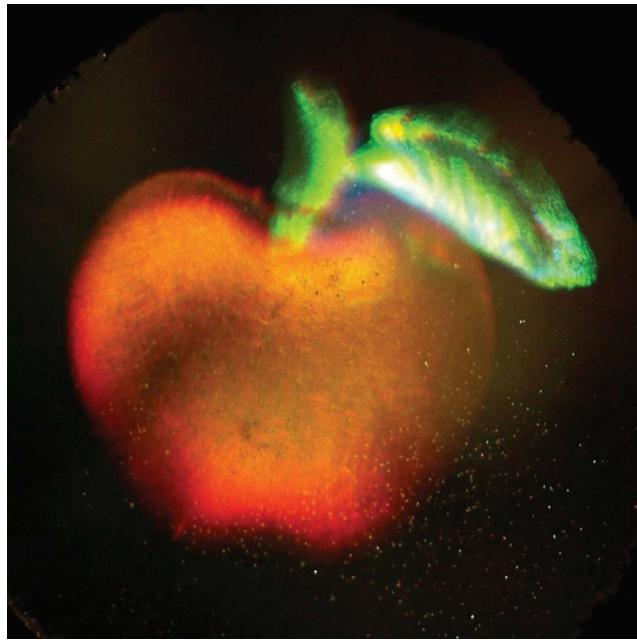


FIGURE A9.8 Surface plasmon hologram of a red apple with a green leaf. (Courtesy S. Kawata, RIKEN.)

distinctive colours. To record a colour hologram in such a recording material, an object is illuminated with red, green, and blue laser light. The light beams are made to diffract onto a glass sheet covered with typically a 150 nm-thick layer of photoresist in which the image is recorded. To make a “plasmon hologram”, it is necessary to cover the photoresist with 55 nm silver and 25 nm glass layers. The hologram is reconstructed using white light by illuminating it through a prism from three different angles, one for each colour, in order to create the 3D colour image. A plasmon colour hologram is shown in Figure A9.8.

A9.5.2 Anomalously High Pulsed Sensitivity of Photoresist

Traditionally, photoresist has not been considered a material suitable for ultrarealistic holographic imaging, and indeed, we have not mentioned this material in Chapter 4. With a typical sensitivity only to blue radiation of approximately 60–90 mJ/cm², it has really only been used extensively by the embossed hologram industry. Now Zacharovas et al. [8] working at the Geola organisation have demonstrated anomalously high sensitivity of Shipley S1813 photoresist to single pulses of blue radiation (440 nm, 50 ns, pulse energy density 6–8 mJ/cm²). DWDH transmission holograms were written using a hogel size of 250 µm and the electroplated gratings studied using an NT-206 atomic force microscope from MTM. Diffractive efficiencies obtained were completely equivalent to typical values obtained on the S1813 material with CW exposure but at total energies 10 times lower. Although photoresist is only sensitive to blue radiation, these latest results may well provoke a reconsideration of this material, which has high diffractive efficiency and excellent stability characteristics, for use in high-quality transmission rainbow and achromatic holograms—particularly when the DWDH technique is used.

A9.6 Printers to Record Digital Colour Holograms

A9.6.1 CGH Composite Reflection Hologram Printer

In 2012, Miyamoto et al. [9] reported a new type of digital holographic printer capable of printing full-colour composite digital reflection holograms using computer-generated holography (CGH). Physically, the printer appears almost identical to a standard DWDH printer. It differs principally in the data displayed on the printer’s spatial light modulator, in the absence of any optics designed to increase the

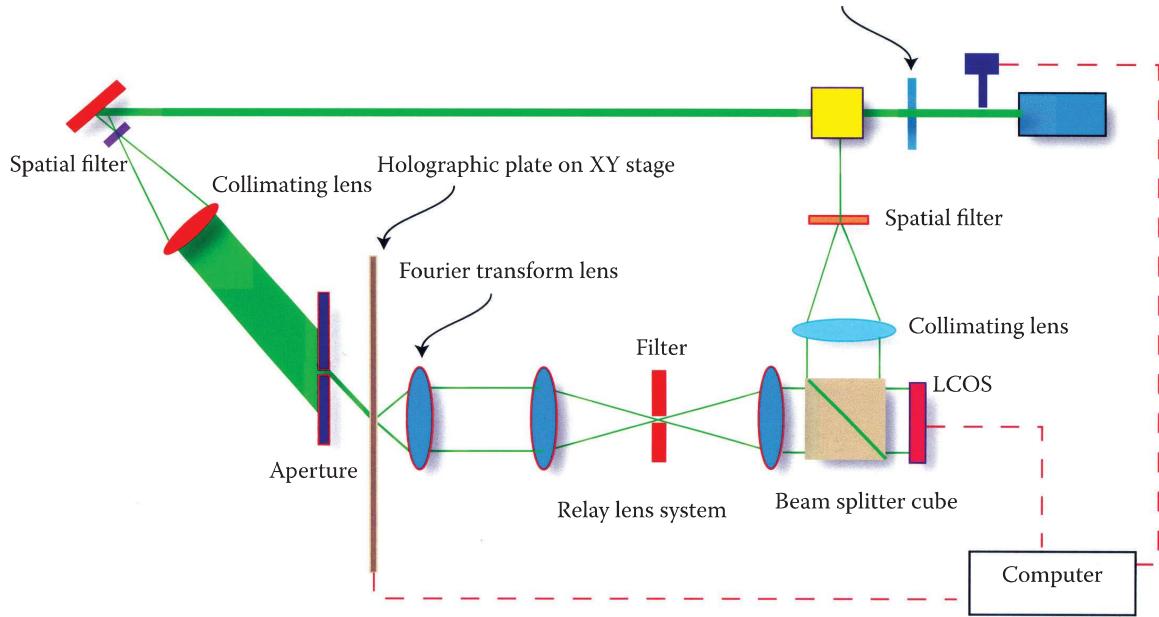


FIGURE A9.9 Simplified optical scheme of the monochrome prototype 2012 CGH reflection hologram printer of Miyamoto et al. [7]. Note that unlike DWDH, first-order diffraction at the SLM is used to create the object beam.

effective étendue of the object beam, which are needed to control hogel size in DWDH, and in its use of first-order diffraction as opposed to the zeroth order that is used in DWDH. In its simplest form (Figure A9.9), a CGH lensless Fourier transform hologram is calculated and displayed on the printer's SLM. This is then illuminated by collimated laser light of high étendue. The first-order diffracted beam produced by the SLM is optically relayed to a Fourier transform lens where it is focused down to the surface of a suitably positioned photosensitive recording material. Here a mutually coherent reference beam co-illuminates the small area illuminated by the diffracted radiation. This results in a small (usually square) elemental reflection hologram of some millimetres in linear dimension being recorded. Further such holograms are then recorded next to this hologram using a step and repeat process. By building up a square matrix of such elemental holograms, a macroscopic reflection hologram is synthesised in a process completely analogous to DWDH.

The main advantage of this new CGH reflection hologram printer over standard DWDH is that each elemental hologram written constitutes effectively a true hologram that contains detailed spatial information exactly as in an analogue hologram. The hogel in DWDH does not have an (x,y) spatial structure, as this is averaged out to increase the fidelity of its main function, which is as a monolithic transmitter of angular information. As such, CGH printers do have the potential to produce better displays than DWDH in applications where close-up viewing of the hologram is paramount. The present state of the technology is, however, far from that of DWDH, which is currently able to produce excellent quality full-parallax full-colour reflection holograms of great depth*—but this may well change.

A9.6.2 Pioneer's Compact Holographic Printer

The Pioneer Corporation in Japan has introduced a compact, full-colour DWDH printer for 3 inch \times 2 inch (about 75 mm \times 50 mm) reflection holograms. The briefcase-shaped consumer electronics product shown in Figure A9.10 has a footprint of 370 mm \times 580 mm and is 130 mm high. The printer recording

* It is often stated that only CGH can produce the correct accommodation cues required in 3D images. This is a common misunderstanding and is misleading. Although CGH is theoretically the technique that is best able to record faithfully in 3D any object or scene with the least optical aberration, from a purely practical point of view, full-parallax DWDH is able to produce full-colour reflection holograms with excellent accommodation cues when hogel sizes are greater than around 0.5 mm.



FIGURE A9.10 New DWDH CW-Laser Colour hologram printer from Pioneer.

wavelengths are 473, 532 and 633 nm. The recording material is the panchromatic photopolymer from Bayer. The holograms are created from graphics images as shown on the unit's monitor.

A9.7 3D Display Systems

A9.7.1 HoloVizio from Holografika

Holografika Kft, Budapest, Hungary [10], demonstrated a new improved HoloVizio system at the *Siggraph 2012* exhibition in Las Vegas. The Hungarian company's 3D holographic technique was mentioned in Chapter 14. A prototype system, the world's first front-projection 140 inch Light Field™ 3D cinema system, was on display. HoloVizio is not a purely holographic system; rather it uses a special holographic screen to redirect light from many small projectors. The pixels of the holographic screen—or, more accurately, voxels as they are intrinsically three-dimensional—emit light beams of different intensity and colour in various directions (Figure A9.11). A light-emitting surface composed of these voxels will act as a digital window or “hologram” and will be able to display 3D information. Many projection engines based on compact LED modules are needed, which are driven by a cluster of nine high-end PCs and sophisticated software. The new large HoloVizio C80 uses the Light Field™ technology, which means that multiple viewers can “look around” objects on the 3.2 m × 1.8 m reflective screen (Figure A9.12). Holografika is working in the framework of various EU projects. The work has been developed in collaboration with Fraunhofer HHI under the MUSCADE European FP7 project. The 3D resolution is 63 megapixels, and the maximum viewing angle is 40°. Recently, Holografika introduced the large HoloVizio 721RC with a 1560 mm × 880 mm 3D display screen with 73 megapixels resolution.

A9.7.2 SeeReal Technologies' Holographic 3D Display

The ultimate 3D television is most likely to be based on the holographic principle; this implies 3D colour images displayed in real time by computer-generated interference fringes generated on an electronic display device of ultrahigh resolution. As we saw in Chapter 14, the great problem here, however, is the insufficient resolution of the available screens.

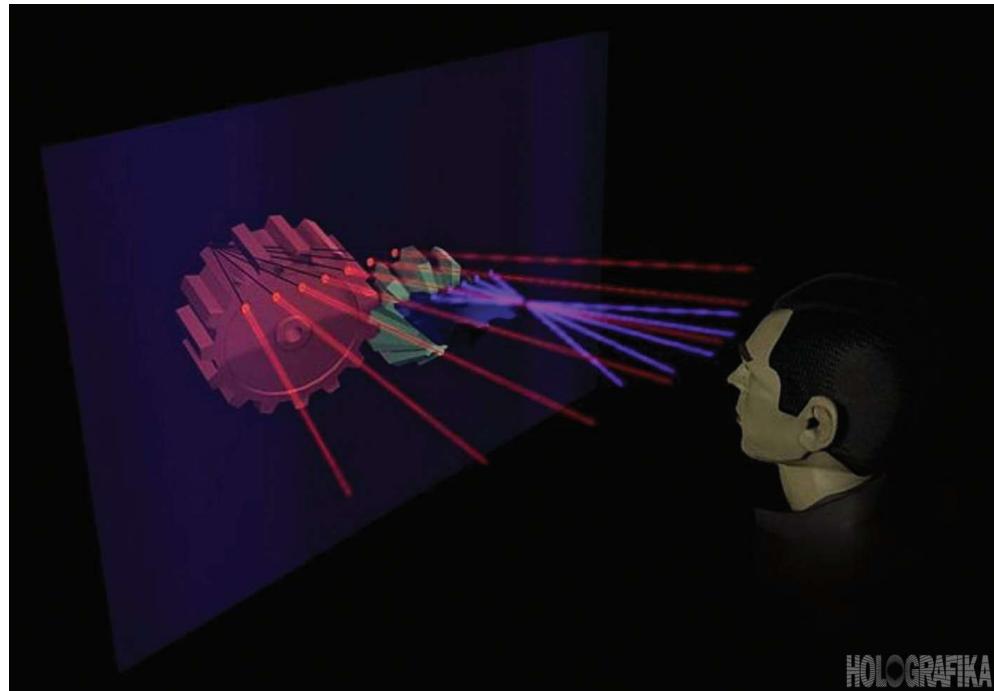


FIGURE A9.11 Principle of the HoloVizio system. (Courtesy of Holografika Kft.)

SeeReal Technologies GmbH, Dresden, Germany, has nonetheless a prototype 3D system up and running that is based on the holographic principle. Eventually, SLMs may well have a resolution high enough to make high-quality full colour real-time holography a reality. The VISIO 20 prototype is, however, based on a real-time computer-generated transmission hologram display with a much lower resolution screen [11–13]. SeeReal's primary goal is to reconstruct the wavefront that is generated by



FIGURE A9.12 HoloVizio 3D cinema system. (Courtesy of Holografika Kft.)

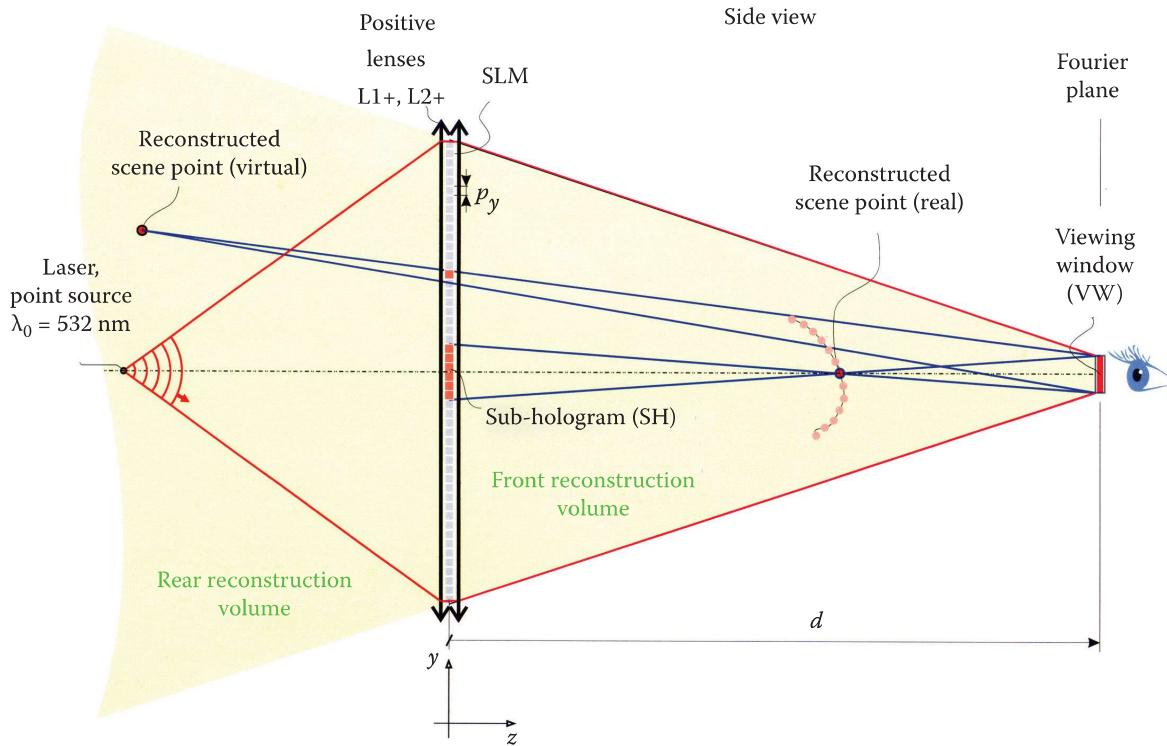


FIGURE A9.13 SeeReal reconstruction geometry. When the hologram (SLM) is illuminated by collimated coherent light ($\lambda_0 = 532$ nm), the 3D scene is reconstructed in the frustum formed by the field lens (L_2+). (From Miyamoto, O. et al. *Proceedings of the SPIE*, 8281, paper 82810N, 1–10, 2012. With permission.)



FIGURE A9.14 Display of a colour 3D image. (Courtesy of SeeReal.)

a real existing object in a viewing window at the eye position. In the company's device, the viewing window is the Fourier transform of the hologram and may be as small as the eye pupil. To get over the lack of resolution available, the viewing window is tracked to the eye position (Figure A9.13). The object size is therefore only limited by the hologram size and not by the hologram resolution. 3D scenes extend from in front of the display to a great depth behind the display. In the colour version of the display, RGB lasers are used with sequential multiplexing so that one laser wavelength at a time illuminates the display screen with its corresponding interference pattern. Figure A9.14 shows an example of a displayed colour image.

A9.8 New Holography Camera from the Hellenic Institute of Holography

The Hellenic Institute of Holography (HiH) in Athens has completed the latest version of its commercial transportable camera Z3^{RGB} (model: ZZZyclops, shown in Figure A9.15). The details of the camera design have been described in Chapter 14. The camera is now ready for use in museums around the world. The ultrarealistic Denisyuk colour holograms recorded by the camera are marketed by HiH as OptoClones®. HiH also reported that they have originated a full-colour DWDH holographic map of the island of Kos in Dodecanese (printed by Geola, size 50 cm × 130 cm) on commission from the Greek Infantry Geographical Services using real GIS data and on permanent exhibition at their museum in Athens. More recent information was provided by Lembessis at the 2012 *Holo-pack-Holo-print* conference in Austria [14].

A9.9 HoloKit from Liti Holographics

Finally, we should mention that Liti Holographics Inc., Newport News, VA [15], is offering a digital hologram printing service for application within the retail signing, merchandising and point of purchase markets. Liti Holographics is also a manufacturer of HoloKits (small complete kits for making holograms), and in 2012, Liti introduced a new kit for making colour reflection holograms. The kit contains three small (RGB) lasers to record 2 inch × 3 inch holograms on a panchromatic photopolymer material.



FIGURE A9.15 HiH transportable camera Z3^{RGB}. (Courtesy of Hellenic Institute of Holography.)

The price of the kit is relatively affordable, which means that artists, students, and amateur holographers may now be able to experience some of the excitement of recording ultrarealistic 3D holographic images.

REFERENCES

1. S. Vignolini, P. J. Rudall, A. V. Rowland, A. Reed, E. Moyroud, R. B. Faden, J. J. Baumberg, B. J. Glover and U. Steiner, "Pointillist structural color in *Pollia* fruit," in *Proc. National Academy of Sciences of the United States, PNAS* **109** (39), 15712–15715 (2012).
2. The 9th International Symposium on Display Holography website: <http://isdh2012.media.mit.edu> (Oct. 2012).
3. The 9th International Symposium on Display Holography video presentations: <http://river-valley.tv/conferences/isdh2012> (Oct. 2012).
4. Zebra Imaging, Inc. USA. www.zebraimaging.com (Oct. 2012).
5. C. Newswanger and M. Klug, "Holograms for the masses," in *Proc. 9th Int'l Symposium on Display Holography*, V. M. Bove Jr. ed., IOP: Conference Series (2013) (to be published).
6. Autodesk®123D® Catch are registered trademarks of Autodesk, Inc. www.123dapp.com/catch (Oct. 2012).
7. M. Ozaki, J.-I. Kato and S. Kawata, "Surface-plasmon holography with white-light illumination," *Science* **332**, 218–220 (April 2011).
8. S. Zacharovas, "Refreshing the concept of security holograms: 3D achromatic masters for manufacturing Embossers," in *Proc. Holo-pack-Holo-print*, www.holopack-holoprint.com (Oct. 2012).
9. O. Miyamoto, T. Yamaguchi and H. Yoshikawa, "The volume hologram printer to record the wavefront of 3D object, in *Practical Holography XXVI: Materials and Applications*. H. I. Bjelkhagen, and V. M. Bove, Jr. eds., *Proc. SPIE*, 8281 (2011) paper 82810N, 1–10.
10. Holografika Kft, Hungary. www.holografika.com (Oct. 2012).
11. S. Reichelt, R. Häussler, G. Fütterer, N. Leicester, H. Kato, N. Usukura and Y. Kanbayashi, "Fullrange complex spatial light modulator for real-time holography," *Opt. Lett.* **37**, 1955–1957 (2012).
12. S. Reichelt and N. Leister, "Computational hologram synthesis and representation on spatial light modulators for real-time 3D holographic imaging," in *Proc. 9th Int'l Symposium on Display Holography*, V. M. Bove Jr. ed., IOP: Conference Series (2013) (to be published).
13. SeeReal Technologies GmbH, Germany. www.seereal.com (Oct. 2012).
14. Liti Holographics Inc, USA. www.litiholo.com (Oct. 2012).
15. A. Lembessis, "Holography for accreditation of works of art," in *Proc. Holo-pack-Holo-print*, www.holopack-holoprint.com (Oct. 2012).

Index

Page numbers followed by f and t denote figures and tables, respectively.

A

Aberration correction, by image predistortion, 561–568
chromatic aberration, compensation for, 566–568
geometrical distortion, compensation for,
564–566
mathematical model, 561–562
optimal reference replay angle calculation, 562–564,
563f, 564f
other corrections, 568
Absolute Imaging Inc., 522
Academia of Media Arts, 540
Académie de Sciences, 43
Advanced Dimensional Displays, 34–35
Advertising displays with holography, 507–511
Agfa-Gevaert, 104
AGHD High Definition Plate, 109
Alfred Nobel Bust, 27, 28f
Alice in the Light World, 537
AP Holographie, 20, 21
Applied Holographics PLC, 30
Argon ion laser, 73, 74f
The Art and Science of the Soviet Union, 537
Art holograms, 515
Australian Holographics Pty. Ltd., 21–22, 538, 541, 543,
544–545, 547
Autodesk 123D Catch, 610
Avatar, 511

B

Bakanas, Ramunas, 550, 551, 552
Bayer Material Science AG, 118
Bayer photopolymer, 419, 420
BB640, 418
Becquerel, Alexandre Edmond, 41
Benton, Stephen, 18–24, 36, 524
Benton, Steven, 198
Benyon, Margaret, 10, 16, 537
Berkhout, Rudie, 24, 24f
Bicubic interpolation, 598–599
Bilinear interpolation, 597–598
Billing, Loren, 540
Bjelkhagen, Hans, 36, 38, 58, 67, 68, 115, 119, 120, 121,
122, 132, 139, 140–142f, 153–155, 193–195,
526–527, 551f, 622
Black-and-white photography, 41
Blender, 317
Blue emission DPSS lasers, 78–80
Blue laser, 187
Blurring
chromatic. *See* Chromatic blurring

digital diffractive, 310
in digital holograms, 420–421
revisited, 419–421
source-size. *See* Source-size blurring
in ultra-thick transmission hologram, 383–384
Bosco, Eric, 507, 509, 547, 548, 549, 550
Boudreau, Joseph, 41
Boulder Nonlinear Systems Lafayette, 524
Boundary conditions
derivation of, RCW theory and, 602–603
reflection holograms, 397–398, 428
transmission holograms, 397–398
Bousigné, Jacques, 20
Bragg condition, 382, 383, 391, 561
for constructive reflection, 389
small departure from, 399–413
lossless reflection hologram, behaviour of,
405–413, 409f–412f
lossless transmission hologram, behaviour of,
401–405, 403f, 404f, 406f, 407f
solving Helmholtz equation with, 395–397, 396f
Bragg planes, 392
Bragg resonance, 428, 429–430, 441–442, 447, 449, 450,
453, 455–456
Bray, Igor, 541, 543
Brightness, 136
Bringing the Artefacts Back to the People project,
499–500
Brotherton-Ratcliffe, David, 39, 69, 154, 193–195,
199, 234–235, 292, 368, 423, 459, 526–527,
541–542, 542f, 544, 545, 546–547, 548, 550,
551f, 552
Bryskin, Vadim, 14
Bubble chamber holography, 161, 494–495
Burns, Joseph (Jody), Jr., 537, 540

C

CAD program. *See* Computer-aided design (CAD)
program
“Calculate” event handler, writing, 363, 364f
Cartesian coordinate system, 239
Casdin-Silver, Harriet, 7
Cavity length stabilisation
in pulsed neodymium lasers, 553–559
by piezo element, 557–559, 558f
statistical optimisation algorithm, 558–559
using heated rear mirror holder, 553–557, 554f–557f
CCD camera. *See* Charge-coupled device
(CCD) camera
Center for Experimental Holography, 14

- Center for the Holographic Arts, 14
- Centred camera configuration, 256–260, 258f, 260f
- full-parallax holograms, 295
 - full-parallax image data acquisition from physical scene, 340–344, 341f
 - image-planed DWDH HPO holograms, 256–260, 258f
- Centred SLM configuration
- image-planed DWDH HPO holograms, 260–263, 260f
- Centre for Laser Photonics in North Wales, 192
- CGH. *See* Computer-generated holography (CGH)
- Charge-coupled device (CCD) camera, 222, 243
- Christakis, Anne-Marie, 16, 540
- Chromatic aberration
- compensation for, 566–568
- Chromatic blurring, 310, 420
- paraxial theory of thin transmission hologram, 379–380, 380f
- Chromaticity, 136
- Chromaticity coordinates, 136
- Chromium forsterite ($\text{Cr}_4\text{:Mg}_2\text{SiO}_4$), 81, 189–191
- injection seeding, 190, 191f
- Circular transmission holograms, 24
- Claudius, Peter, 34
- Cobolt AB, 79
- new lasers from, 614
- Coherent, new lasers from, 614
- Collimated reference, paraxial theory of thin transmission hologram and, 373–376, 374f, 375f
- Colorimetry, 136
- Colour holograms, 143–152
- exposure, 146–148
 - illumination, 142
 - processing, 148–149, 148t
 - recorded and evaluated, 150–152, 151f, 152f
 - reflection hologram, 143–145
 - sealing, 149–150
 - single-layer silver halide emulsions, 132–134
 - transmission hologram, 143
 - wavelength selection, 137–142
- Colour Holographics Ltd., 104–105, 418
- Colour holography. *See also* Colour holograms; Digital colour holography
- colour theory and measurements, 134–137
 - history, 126–128
 - lasers for
 - from Cobolt, 614
 - from Coherent, 614
 - from Lasers Quantum, 614
 - RGB Laser Technology at Geola, 614–615, 615f - monochromatic holography, 124–125
 - overview, 123–124
 - SHSG, 128–132
- Colour portraits, digital holographic, 512–515
- Colour theory and measurements, 134–137
- Colour transmission hologram, 134, 143
- Commercial DWDH printers based on 2001 Geola printer, 215, 215f
- Commercial holocam systems
- image data from physical scene, 331–334, 332f–333f
- Commission Internationale de l'Eclairage (CIE), 24, 134, 135f, 136
- Commonwealth Scientific and Industrial Research Organisation in Canberra, 22
- Complex wave fronts, recording of
- multiple gratings and, 414
- Computer-aided design (CAD) program, 244, 245
- Computer-generated holography (CGH), 197, 198, 199
- CGH composite reflection hologram printer, 616–617, 617f
- Conductron Corporation, 6–7
- Conferences, holographic, 538
- Conservation of energy, 427
- Contact-copy method, 231–232
- Continuous wave (CW) lasers, 71–86, 157, 199, 202
- diode-pumped solid-state (DPSS) lasers, 77–82
 - blue emission, 78–80
 - chromium forsterite, 81
 - green emission, 78
 - microchip lasers, 81
 - OPO, 81–82
 - orange emission, 80
 - red emission, 80
 - yellow emission, 80
- gas lasers, 72–76
- argon ion laser, 73, 74f
 - copper vapour laser, 76
 - gold vapour laser, 76
 - helium–cadmium (HeCd) laser, 74, 75f
 - helium–mercury laser, 75
 - helium–neon (HeNe) laser, 72, 72f–73f
 - helium–selenium, 75
 - krypton ion lasers, 73–74
- semiconductor diode lasers, 82–85
- amplification, 84–85
 - mode structure and coherence, 83
 - operation and construction, 82, 82t, 83f
 - single longitudinal mode, 83–84
 - vs. pulsed lasers, 158
- Control and video image stream system
- RGB-pulsed laser triple-beam DWDH printers, 207–208, 209f
- Coordinate systems
- digital holograms, 238–242, 239f, 240f
- Copper vapour laser, 76
- Coronation Crown of Erik XIV*, 11–12, 11f–12f
- Cotteverte, Jean-Jacques, 548
- Coupled wave theory (CWT), 391, 394–418, 455–456, 601. *See also* Rigorous coupled wave (RCW) theory, N-Coupled wave theory and the Parallel stacked mirrors model (PSM)
- boundary conditions for CWT, 397
- CWT away from perfect Bragg compliance, 399–414
- CWT with perfect Bragg compliance, 395–399
- definition of diffraction efficiency in CWT, 398
- including loss in CWT, 413–414
- one-dimensional CWT, 394–395
- power conservation in CWT, 398, 413
- solving Helmholtz equation in volume gratings. *See* Helmholtz equation
- treatment of the lossless reflection hologram/grating in CWT, 405–413
- treatment of the lossless transmission hologram/grating in CWT, 401–405

- Credit cards, 19, 19f
Critical angle, 312
Cr:LiSAF, 187, 188, 189f
Cross, Lloyd, 198, 540
Cullen, Ralph, 544
CW lasers. *See* Continuous wave (CW) lasers
CX-PRO scanners, 347, 348f
- D**
- Dai Nippon Printing Co., Ltd., 118
Dalí, Salvador, 7
Data predistortion, 285–288, 287f, 561–568. *See also*
 Aberration correction
Dausmann, Günther, 22
Dawson, Paula, 8, 9f, 21, 311
DeBitetto, D. J., 197–198
DeBitetto/King model, 198
Decouvrez l'Holographie, 537
Denisyuk, Yuri, 2, 5, 5f, 26–27, 30
De Saint-Victor, Abel Nièpce, 41
Diamond, Mark, 35
The Diamond Queen, 612
Dichromated gelatin (DCG), 90, 110–115, 161
 CW exposure, 111
 manufacturers, 115
 panchromatic, 112–114
 preparation, 110
 processing, 111–112, 111t
 pulsed laser exposure, 111
 sensitising, 110–111
Diffraction efficiency, 398, 400–414, 428. *See also*
 Coupled-wave theory and Parallel stacked
 mirrors (PSM) Model
of planar gratings
 rigorous chain matrix method, 421–422
 rigorous coupled wave theory and rigorous modal
 theory for, 421
for transmission and reflection holograms, 398–399,
 399f, 425–459, 601–608
Digital colour holography, 197–234
 background, 197–198
 DeBitetto/King model, 198
 digital data processing. *See* Digital data processing, of
 holographic printers
 DWDH, 198–199. *See also* Direct-write digital
 holography (DWDH)
full-parallax holograms, computational methods for.
 See Full-parallax holograms
holographic stereograms, 197–198
modern DWDH printers, 202–230
MWDH printers, 230–232, 231f
overview, 197
recent developments, 616–618
 CGH composite reflection hologram printer,
 616–617, 617f
 Pioneer's compact holographic printer, 617–618, 618f
 simple DWDH printer, 199–202, 200f
Digital data processing, of holographic printers, 237–292
 basic considerations, 237–238
 coordinate meshes, 241–242
- coordinate systems, 238–239, 239f, 240f
full-parallax DWDH “Denisyuk” reflection hologram,
 243–245, 244f
Huygens' principle, 237
image-planed DWDH HPO holograms. *See* Image-
 planed DWDH HPO holograms
independent primary colours, 242
MWDH master hologram, 245, 246f
overview, 237
printer optical objectives, correcting inherent
 distortion in, 283
data predistortion, 285–288, 287f
formalism setting, 283–284, 284f, 285f
 non-paraxial I-to-S transformations, 288–292, 291f
rainbow and achromatic transmission holograms
 DWDH, 276–283
 MWDH, 268–276
 MWDH achromats, 276
 viewing plane, 242–243, 243f
Digital diffractive blurring, 310, 420–421
Digital display holograms (image data), 317–368. *See also*
 Display holography
image acquisition from physical scene (full-parallax
 digital holograms), 339–347
centred camera data set, calculation of, 340–344,
 341f
3D structured-light scanners, 346–347, 347f, 348f
horizontal rotating platform with vertically linear-
 translating camera, 340–346, 341f, 346f
I-to-S transformation, derivation of, 344–346, 346f
image acquisition from physical scene (HPO
 holograms), 317–339, 318f
alternative strategies (rotating the object), 334–339,
 334f–337f
commercial holocam systems, 331–334, 332f–333f
hybrid translating/rotating camera with fixed
 target, 319–326, 320f–322f, 325f
hybrid translating/rotating camera with optimised
 target, 326–329, 326f, 330f
image reduction and magnification, 331
resolution requirements, 330–331
simple translating camera, 318–319, 318f
images derived from virtual computer model (full-
 parallax holograms), 357–368
“Calculate” event handler, writing, 363, 364f
function of MAXScript, 360–361
global variables, 361
“Make Camera” event handler, writing, 364–365,
 364f
“Make Directories” event handler, writing,
 362–363, 362f
MAXScript GUI, creating, 357–359, 357f–359f,
 360t
“Start Render” event handler, writing, 365–368,
 366f, 367f
images derived from virtual computer model (HPO
 holograms), 348–357
creating 20 cm × 30 cm landscape hologram,
 349–357, 350f–356f
 data preparation, 348–349, 349t
overview, 317

- Digital holograms. *See also* specific entries
 blurring in, 420–421
 constraints on
 aberrations of printer optical system, 520
 constraints on digital data, 520
 constraints on illumination source, 520–521
 field of view, 519
 optimal hogel size, 518–519
 on photographic material, 521
 pixel dimensions of printer SLM, 519–520
 printer variable reference beam system, 520
- Digital holographic printing, 508
- Digital image predistortion
 aberration correction by, 561–568
 chromatic aberration, compensation for, 566–568
 geometrical distortion, compensation for, 564–566
 mathematical model, 561–562
 optimal reference replay angle calculation, 562–564, 563f, 564f
 other corrections, 568
 transmission/reflection holograms, 394
- Diode-pumped solid-state (DPSS) lasers, 77–82, 192–193
 blue emission, 78–80
 chromium forsterite, 81
 green emission, 78
 microchip lasers, 81
 OPO, 81–82
 orange emission, 80
 praseodymium-doped lasers, 80–81, 80f
 red emission, 80
 yellow emission, 80
- Diode pumping, 191–193
 in Nd:YAG, 192–193
 Nd:YVO₄, 193
- Direct-write digital holography (DWDH), 198–199, 293, 348, 374, 597, 609
 advantages, 198
 angular field of view of, 204
 contact–copy method, 231–232
 copying full-colour holograms, 232–234, 233f
 full-parallax DWDH “Denisyuk” reflection hologram, 243–245, 244f
 full-parallax holograms, 300–301. *See also* Full-parallax holograms
 HPO, data preparation issues related, 313–314
 image-planed DWDH HPO holograms. *See* Image-planed DWDH HPO holograms
 larger holograms, 300–301
 modern holographic printers, 202–230
 HPO printers, 219–223
 incorporating variable reference beam systems, 217–219, 222f, 223f, 224f
 LCOS printers, 216, 217f, 218f, 219f
 lens-based printers, 203–204, 204f
 pulsed lasers use in, 202–203
 single-beam RGB printers, 223–226
 speckle in lens-based printers, 205
 transmission hologram printers, 228–230
 triple-beam printers, 205–215. *See also* Triple-beam DWDH printers
 ultra-realistic printers, 226–228
- rainbow and achromatic transmission holograms, 276–283, 374
 RGB laser use, I-to-S transformation and, 278–281, 279f, 280f
 synthesis of vertical slit, 277–278, 277f
 transformations for one-colour laser, 281–283
- simple printer, 199–202, 200f
 deficiencies, 201–202
 image data, 201
 operation, 200–201
 optical scheme, 199–200
 speckle blur, 200
- Dispersion equation theory, 418–419
- Display holography. *See also* Digital display holograms
 (image data); High-fidelity display holograms
 historical origins, 537–540
 commercial and educational entities involved in, 538–540
 companies producing holograms, 538–539
 educational institutes, 540
 hologram exhibitions, 537–538, 538f
 hologram galleries, 539, 539f
 holographic museums, 540
- The Dream*, 34
- Dreaming History*, 17
- Driving reference wave, 426
- Dual linear-cavity pulsed RGB Nd:YAG laser, 179–183
- Dual-ring cavity pulsed RGB Nd:YLF/YAG laser, 177–179
- Dublin Institute of Technology, 117
- DuPont Holographics, Inc., 517
- Dutch Holographic Company B.V., 35
- DWDH reflection holograms. *See also* Direct write digital holography (DWDH)
 chromatic blurring, 310
 digital diffractive blurring, 310
 resolution requirements of, 310–311
 source-size blurring, 310
 speckle, 310
- DWDH transmission holograms, 311–312
 advantage, 312
 printers, 228–230, 229f–230f
- Dye lasers, 76–77
 liquid, 76
 solid-state, 77
- E**
- Edhouse, Simon, 21, 544–545, 547
- Edmund Scientific company, 539
- Educational institutes, 540
- Electrical conductivity, of dielectric grating
 Helmholtz equation and, 413–414
- Electric Fireplace*, 517
- Electro-optical Q-switching system, 553
- Emulsions, 91, 93–96
 DCG, sensitising, 110–111
 demands on, 93–94, 94f
 film substrates, 102, 102f, 103
 glass plates, 102, 102f, 103
 grain sizes, 91, 92t
 hypersensitisation, 100

- image resolution, 94–96, 95f, 96f
latensification, 100–101
Lippmann photography, 49–53, 58
manufacturing companies, 104–105
MTF, 91–92, 92f
OTF, 91
sensitivity, 93, 104
SilverCross project, 105–109
substrates, 102–103
swelling, thick transmission hologram, 393–394
Endoscopic holography/endoholography, 497–499
E-632 Physics Experiment, 494
Equanimity, 612
Exhibitions, hologram
 historical origins, 537–538, 538f
Exposure, colour holograms, 146–148
- F**
- Farquhar, Yasmin, 21
Fedorenchik, Vladimir, 548
Fenton, John, 541–542
Fermilab, 10, 10f
Fermilab pulse-stretching, 161–162, 162f
Field of view (FOV), 318–319
Film substrates, 102, 102f, 103
Finite-source-size ray-tracing for illuminant design, 583–586
 method, 584–585
 spatial coherence, 585–586
Finite thickness, holograms of, 390–393, 392f
Fire bed hologram, 517
40-J pulsed ruby laser facility, at Fermilab, 10, 10f
4DDynamics, 347
Fournier, Jean-Marc, 7, 8f
FOV. *See* Field of view (FOV)
Fox, Geoff, 21, 547
Fresnel amplitude coefficients, 438
Fresnel amplitude reflection, 426
Fresnel rules, 426
Frieb, Matthias, 540
Fringe Research, 540
Fuel-effect electric fires, holographic, 517
Full-colour DWDH holograms, copying, 232–234, 233f
Full-colour holographic imaging, visual application of
 in England, 504–505
 in Greece, 501–504
 of museum artefacts, 499–506
 of oil paintings, 505–506
 virtual museum exhibitions, 499–501
Full-parallax digital holographic portraits, 514
Full-parallax DWDH “Denisyuk” reflection hologram, 243–245, 244f, 293
Full-parallax holograms, 293–314
 changing image plane, strategies for, 294
 camera definition, 295
 full-parallax paraxial objective transformations, 298–300
 larger holograms, 300–301
 non-paraxial printer objectives, 300
rectangular viewing windows, 301–310, 302f–307f, 309f
two-step I-to-S transformations, 296–298, 297f, 298f
data preparation issues related to, 313–314
3D model, 294
DWDH reflection holograms, resolution requirements of, 310–311
DWDH transmission holograms, 311–312
HPO holograms *vs.*, 313–314
image data from physical scene, 339–347
 centred camera data set, calculation of, 340–344, 341f
3D structured-light scanners, 346–347, 347f, 348f
horizontal rotating platform with vertically linear-translating camera, 340–346, 341f, 346f
I-to-S transformation, derivation of, 344–346, 346f
images derived from virtual computer model, 357–368
“Calculate” event handler, writing, 363, 364f
function of MAXScript, 360–361
global variables, 361
“Make Camera” event handler, writing, 364–365, 364f
“Make Directories” event handler, writing, 362–363, 362f
MAXScript GUI, creating, 357–359, 357f–359f, 360t
“Start Render” event handler, writing, 365–368, 366f, 367f
MWDH reflection holograms, 312–313, 313f
 overview, 293–294
Full-parallax transmission holograms, 6
Furst, Anton, 537
- G**
- Gabor, Dennis, 3–4, 4f, 6, 7
Galleries, hologram, 539, 539f
Gallery 1134: Fine Arts Research and Holographic Center, 540
Gas lasers, 72–76
 argon ion laser, 73, 74f
 copper vapour laser, 76
 gold vapour laser, 76
 helium–cadmium (HeCd) laser, 74, 75f
 helium–mercury laser, 75
 helium–neon (HeNe) laser, 72, 72f–73f
 helium–selenium, 75
 krypton ion lasers, 73–74
Gateaux Gabor, 12, 13f
Gelatin plates. *See* Dichromated gelatin (DCG)
General Motors, 6
General Optics Pty. Ltd., 541, 543, 544, 545
Gentet, Yves, 16
Geola Digital UAB, 199, 206, 227, 228, 230, 232, 311, 331, 348–349
 incorporation of, 545–546
printing parameters for HPO reflection holograms, 349t
progress in RGB laser technology at, 614–615, 615f

- Geola Laboratories, 17
 Geola organisation, 538, 609
 beginning, 541–542
 first meeting in Vilnius, 543–545
 first Romanian exhibition of large-format holography, 547
 history of, 541–552
 incorporation of Geola UAB, 545–546
 panchromatic film and Sfera-S, 549–550
 Vilnius digital printer, 550–552
 XYZ Imaging Inc.
 sale of, 552
 start of, 547–548
 Geola Technologies Ltd., 548, 549, 550
 Giger, H. H., 17, 545
 Glasses-free autostereoscopic screens for 3D visualisation systems, 522
 Glass plates, 102, 102f, 103
 Glass thickness, 103
 Global variables
 MAXScript program, 361
 Goldberg, Larry, 538–539
Gold Collar, 27, 27f
Goldsmiths' College of Art, 540
 Gold vapour laser, 76
 Grating planes, 387
 Gratings
 multiple. *See* Multiple gratings
 planar. *See* Planar gratings
 response to a plane reference wave, 425
 volume. *See* Volume gratings
 Grating theory. *See* Coupled-wave theory, Kogelnik's theory, Parallel stacked mirrors (PSM) model, N-PSM theory, N-coupled wave theory, Rigorous coupled wave theory, Dispersion equation theory and Diffraction efficiency
 Grating vector, 387, 396f, 446, 456, 561
 High virtual volume (HVV) holograms, 226–227, 228, 518–521
 Green emission DPSS lasers, 78
 neodymium crystals, 78
 Verdi 532 nm laser, 78, 79f
 Green lasers, 183–187
 Grichine, Mikhail, 17, 544, 545, 546, 547, 548
 Grosmann, Michel, 16
- H**
- Hand in Jewel*, 6
 Hariharan, Parameswaran, 18, 19, 22, 31
 Harman Technologies, 104, 109
 Hart, Stephen, 25
 Heated rear mirror holder
 cavity length stabilisation in pulsed neodymium lasers using, 553–557, 554f–557f
 Heliochromes, 41
 Heliochromy, 41
 Helium–cadmium (HeCd) laser, 74, 75f
 Helium–mercury laser, 75
 Helium–neon (HeNe) laser, 72, 72f–73f
 Helium–selenium lasers, 75
- Hellenic Institute of Holography (HiH)
 new holography camera from, 621, 621f
 Helmholtz equation, 427, 601
 for the normal-incidence reflection grating, 429–430
 Runge–Kutta integration of, 430
 solving, in volume gratings (coupled wave theory), 394–419
 boundary conditions, 397–398
 with Bragg compliance, 395–397, 396f
 complex wave fronts and multiple gratings, recording of, 414
 diffraction efficiency, 398–399, 399f
 dispersion equation theory, 418–419
 effect of loss in dielectric, 413–414
 multiple colour gratings, recording, 418
 multiple gratings generated by many object points, 414–418, 415f
 one-dimensional coupled wave theory, 394–395
 power conservation, 398
 small departure from Bragg condition, 399–413, 403f, 404f, 406f, 407f, 409f–412f
 Henderson, Ricky, 35, 36f
 Heron's formula, 582–583
 Herschel, John, 41
 Heytesbury Holography Bell Resources Ltd., 8
 High-fidelity display holograms, 369–422
 blurring revisited, 419–421
 in digital holograms, 420–421
 diffractive efficiency of planar gratings
 rigorous chain matrix method, 421–422
 rigorous coupled wave theory and rigorous modal theory for, 421
 emulsion swelling and change in refractive index on processing, 393–394
 finite thickness holograms, 390–393, 392f
 Helmholtz equation solution in volume gratings
 (coupled wave theory), 394–419
 boundary conditions, 397–398
 with Bragg compliance, 395–397, 396f
 complex wave fronts and multiple gratings, recording of, 414
 diffraction efficiency, 398–399, 399f
 dispersion equation theory, 418–419
 effect of loss in dielectric, 413–414
 multiple colour gratings, recording, 418
 multiple gratings generated by many object points, 414–418, 415f
 one-dimensional coupled wave theory, 394–395
 power conservation, 398
 small departure from Bragg condition, 399–413, 403f, 404f, 406f, 407f, 409f–412f
 laser transmission holograms and problem of speckle, 380–381
 non-paraxial behaviour and digital image predistortion, 394
 overview, 369
 reflection holograms, 384–386, 385f
 thick-phase hologram, PSM model of, 386–390, 387f–389f

- thick transmission hologram, three-dimensional theory of, 381–384
blurring in, 383–384
Snell's law at the air–hologram boundary, 382–383
- thin transmission hologram, three-dimensional
paraxial theory of, 370–380, 371f
chromatic blurring, 379–380, 380f
collimated reference and object beams, 373–376, 374f, 375f
source-size blurring, 376–379, 377f, 378f
- High-intensity reciprocity failure (HIRF), 97, 98, 99
- High Power Laser Energy Research (HiPER) Facility, 192
- High virtual volume (HVV) holograms, 226–227, 228, 518–521
- HiH. *See* Hellenic Institute of Holography (HiH)
- Hill, Levi L., 41
- Hill's matrix technique, 431
- HIRF. *See* High-intensity reciprocity failure (HIRF)
- Hlynksy, David, 540
- HOEs. *See* Holographic optical elements (HOEs)
- Hogels, 157–158, 198, 199
- Hogel-writing sequence
triple-beam DWDH printers, 205–206, 206f, 207f
- The Holox Corporation*, 538–539
- Holicon Corporation, 20
- Holocam systems, 331–334, 332f–333f
- HOLOCO, 537
- Holocopier AHS1, 30
- HOLOEXPO 2011*, 538
- HOLO FX green-sensitive plates, 109
- HOLO FX red-sensitive plates, 109
- Holografi: Det 3-Dimensionella Mediet*, 537
- Holografika Kft, 609, 618
- Hologram. *See also* specific types
defined, 3
- Hologram exhibitions, historical origins, 537–538, 538f
- Hologram galleries, 539, 539f
- Hologram Industries*, 539
- Holograms of finite thickness, 390–393, 392f
- Holographic Images, Inc., 31
- Holographic medical imaging application, 25
- Holographic microscopy, 496–497
- Holographic Museum in Chicago*, 540
- Holographic museums, 540
- Holographic Ocean Laboratories*, 540
- Holographic optical elements (HOEs), 90, 197
- Holographics North, 20
- Holographic stereograms, 197–198
- Holography
in advertising, 507–511
art, 515
colour portraits, 512–515
fuel-effect electric fires, 517
full-colour holographic imaging, visual application of, 499–506
glasses-free autostereoscopic screens for 3D
visualisation systems, 522
holographic windows, 517–521
in mapping and architectural design, 511–512
multiprojector autostereoscopic systems, 522–523
real-time true holographic displays, 523–525
scientific applications of, 493–499
smaller full-colour holograms, 515–517
super-realistic 3D static displays, 517–521
updateable holographic display, 521–522
- Holography '75: The First Decade*, 537
- Holography Institute*, 540
- Holography Marketplace*, 516
- Holography or Reality at the Holographic World*, 538
- Holography School of the Tamara Art College*, 540
- HoloKit, from Liti Holographics, 621
- Holomation, 36
- Holopixels. *See* Hogels
- Holorad LLC, 25
- HOLOS Gallery*, 539
- Holotec GmbH, 115
- HoloVizio™ 3D holographic system, 609, 618, 619f
- Holowood*, 540
- Horizontally mounted overhead-illuminated holograms, 511–512
- Horizontal parallax-only (HPO) printers, 209, 219, 293
HPO DWDH, data preparation issues related, 313–314
- HPO 3D glasses-free display systems, 522
- HPO holograms, 374
image data from physical scene, 317–339, 318f
alternative strategies (rotating the object), 334–339, 334f–337f
commercial holocam systems, 331–334, 332f–333f
hybrid translating/rotating camera with fixed target, 319–326, 320f–322f, 325f
hybrid translating/rotating camera with optimised target, 326–329, 326f, 330f
image reduction and magnification, 331
resolution requirements, 330–331
simple translating camera, 318–319, 318f
image-planed DWDH HPO holograms. *See* Image-planed DWDH HPO holograms
- images derived from virtual computer model, 348–357
creating 20 cm × 30 cm landscape hologram, 349–357, 350f–356f
data preparation, 348–349, 349t
vs. full-parallax holograms, 313–314
- HPO printers. *See* Horizontal parallax-only (HPO) printers
- Hue, 136
- Huygens' principle, 237, 243, 294
- HVV holograms. *See* High virtual volume (HVV) holograms
- Hybrid translating/rotating camera
image data by HPO holograms from physical scene
with fixed target, 319–326, 320f–322f, 325f
with optimised target, 326–329, 326f, 330f
- Hypersensitisation, 100
- I**
- Ilford Ltd, 30
- Illumination, colour holograms, 142
- Image data, from physical scene
full-parallax digital holograms, 339–347
centred camera data set, calculation of, 340–344, 341f
3D structured-light scanners, 346–347, 347f, 348f

- horizontal rotating platform with vertically linear-translating camera, 340–346, 341f, 346f
- I-to-S transformation, derivation of, 344–346, 346f
- HPO holograms, 317–339, 318f
- alternative strategies (rotating the object), 334–339, 334f–337f
 - commercial holocam systems, 331–334, 332f–333f
 - hybrid translating/rotating camera with fixed target, 319–326, 320f–322f, 325f
 - hybrid translating/rotating camera with optimised target, 326–329, 326f, 330f
 - image reduction and magnification, 331
 - resolution requirements, 330–331
 - simple translating camera, 318–319, 318f
- Image data, from virtual computer model
- full-parallax holograms, 357–368
 - “Calculate” event handler, writing, 363, 364f
 - function of MAXScript, 360–361
 - global variables, 361
 - “Make Camera” event handler, writing, 364–365, 364f
 - “Make Directories” event handler, writing, 362–363, 362f
 - MAXScript GUI, creating, 357–359, 357f–359f, 360t
 - “Start Render” event handler, writing, 365–368, 366f, 367f
 - HPO holograms, 348–357
 - creating 20 cm × 30 cm landscape hologram, 349–357, 350f–356f
 - data preparation, 348–349, 349t
- Image-planed DWDH HPO holograms, 245, 247
- advantages, 246
 - centred camera configuration, 256–260, 258f
 - centred SLM configuration, 260–263, 260f
 - fundamental integer constraints, 263–268, 264f, 265f, 267t
 - general rectangular viewing window, 248–256, 248f, 252f, 254f
 - simple translating camera, 248–256, 248f, 252f, 254f
 - printer, camera and viewing window options, 247–248
- Image reduction and magnification
- HPO holographic image, 330–331
- Images in Time and Space*, 538
- Independent primary colours
- digital holograms, 242
- Infortechnika UAB, 543
- Injection seeding, 190, 191f
- Integer constraints
- image-planed DWDH HPO holograms, 263–268, 264f, 265f, 267t
- Interferential imaging, 2–3, 41–42
- examples, in nature, 42–43, 42f
 - recordings, 41–42
- International Symposium on Display Holography*, 540
- 9th International Symposium on Display Holography* (ISDH), 609
- Into the Rainbow*, 22, 24
- Iridescence, 45
- Isachenkov, Lev (Leon), 541, 542f, 543, 545
- Isel, 331
- I-to-S transformation(s)
- full-parallax image data acquisition from physical scene, 340–344, 341f
 - for monochromatic DWDH HPO reflection hologram, 253–254, 256, 258, 259, 262, 263
 - for MWDH achromatically tilted H_1 rainbow hologram, 275–276
 - non-paraxial, 288–292, 291f
 - RGB laser and, 278–281, 279f, 280f
 - for three-component colour H_1 holograms, 271
 - two-step, changing image plane in full-parallax holograms, 296–298, 297f, 298f
- J**
- Jackson, Rosemary (Posy), 537, 540
- Jeong, Tung H., 540
- Jersey Heritage Trust, 612
- Jersey Postage stamp, 612, 613f
- JK Lasers, 160, 161f
- John the Baptist*, 10
- Journal of the Optical Society of America*, 3
- Jung, Dieter, 22, 24, 540
- Juyong Lee, 17, 21
- K**
- K1A Plate, 109
- Karaganov, Viktor, 543–544
- Kate McGougan & Stephen Jones Hat*, 24, 25f
- Kaufman, John, 31
- “Keystoning,” 340
- Kodak, 104
- Kogelnik’s coefficients, 428
 - slanted reflection gratings, 447
- Kogelnik’s constant, 427
- Kogelnik’s coupling constant, 395, 413
- Kogelnik’s equations
 - for normal-incidence unslanted sinusoidal grating, 428
 - for single-colour grating, 447
 - for sinusoidal gratings, 441
 - for slanted grating in three dimensional space, 449
- Kogelnik’s theory, 425, 443–445, 448, 450, 453. *See also*
- Coupled wave theory
 - for slanted gratings, 450–451
 - vs. PSM theory with RCW theory, 604
- Komar, Victor, 33
- Konovalov, Dmitry, 541, 543, 544
- Kosenko, Eugene, 548
- Kranseder & Cie A.G., 49
- Krantz, Eric, 36
- Krypton ion lasers, 73–74
- L**
- Lake Forest College, 540
- Lamp-pumped pulsed neodymium lasers
 - cavity length stabilisation in, 553–555
 - by piezo element, 557–559, 558f
 - statistical optimisation algorithm, 558–559
 - using heated rear mirror holder, 553–557, 554f–557f

- Laser and optical system
RGB-pulsed laser triple-beam DWDH printers, 210, 211f
Lasergruppen Holovision AB, 27, 27f–29f
Laser Megajoule, 192
Lasers for colour holography
from Cobolt, 614
from Coherent, 614
from Lasers Quantum, 614
RGB Laser Technology at Geola, 614–615, 615f
Lasers Quantum, new lasers from, 614
Laser stability
RGB-pulsed laser triple-beam DWDH printers, 214–215
Laser transmission holograms
problem of speckle and, 380–381
Latensification, 100–101
internal, 101
Lauk, Matthias, 540
La Villette Science Park, in Paris, 21
LCDs. *See* Liquid crystal displays (LCDs)
LCOS displays. *See* Liquid crystal on silicon (LCOS) displays
LED. *See* Light-emitting diodes (LED)
Lehmann, Hans, 51–52
Leibnitz's chain rule, 445
Leith, Emmett, 3, 4f
Lens-based DWDH printers, 203–204, 204f
speckle in, 205
Lesniewski, Marcin, 548, 551
Levine, Chris, 16, 609, 612
LGN-222 He-Ne laser, 73f
Lieberman, Larry, 31
Light Dimensions, 537
Light-emitting diodes (LED), 573
Light Fantastic, 539
Light Fantastic I, 537
Light Fantastic II, 537
Lightness, 136
Light-source
monochromatic, 574–575, 574f–575f
nine-diode framing light source, commercial design of, 594–595, 594f
polychromatic, 575–577, 576f
reflective, 595–596
Light Wave Gallery, 539, 540f
Light Years Ahead, 537
Lippmann, Gabriel Jonas, 2–3, 43–45, 46f. *See also* Lippmann photography
Lippmann photography, 2–3, 456–458, 609
emulsions, 49–53, 58
monochromatic recording, 46
monochromatic replay, 46–47
photopolymers, 59–60, 60f
polychromatic recording, 47
polychromatic replay, 48–49
recording and processing, 53–57, 58
silver halide, 60–63, 61f–62f, 63f–64f
Liquid crystal displays (LCDs), 198
Liquid crystal on silicon (LCOS) displays, printers based on, 216, 217f, 218f, 219f
LIRF. *See* Low-intensity reciprocity failure (LIRF)
Liti Holographics, HoloKit from, 621
LMC France Instruments SARL, 544, 547
LXML-PM01-0080 Luxeon Rebels, 574
Lossless reflection hologram, behaviour of, 405–413, 409f–412f
Lossless transmission hologram, behaviour of, 401–405, 403f, 404f, 406f, 407f
LOTIS TII, 189
Low-intensity reciprocity failure (LIRF), 97, 99
LS-2132 Nd:YAG laser, 189, 190f
LT-2212G, 189, 190f
Lumière, Auguste, 49–50
Lumière, Louis, 49–50
Luminance factor, 136
Luminous Presence hologram, 515
- ## M
- Maiman, Theodore “Ted” H., 160
Main lens surface, design of
RGB LED hologram, 577
geometrical ray tracing in point-source approximation, 580–581, 580f
mathematical formulation, 578–580, 579f
monochromatic light source, 577–578, 577f–578f
numerical solution, 583
power density distribution calculation at target plane, 582–583, 583f
ray intersection with target plane, 581–582, 582f
ray tracing with finite source size, 583–586
“Make Camera” event handler, writing, 364–365, 364f
“Make Directories” event handler, writing, 362–363, 362f
Making Faces, 7
Man on the Motorbike, 10, 10f
Manufacturers
dichromated gelatin (DCG), 115
photopolymers, 118
silver halide, 104–105
Marching Band Hologram, 30, 30f
Masks, 34
MasterCard, 19, 19f
Master oscillator power amplifier (MOPA) architecture, 159
Master-write digital holography (MWDH) printers, 198–199, 230–232, 231f, 293, 348
data-processing problem, 314
image data, 245
rainbow and achromatic transmission holograms, 268–276
achromatically tilted component-colour H_1 masters, 275
achromatically tilted RGB H_1 master, 276
component-colour H_1 masters, 271–274, 273f
vertically aligned RGB H_1 master, 268–271, 269f, 271f, 272f
MAXScript GUI, 317
creating, 357–359, 357f–359f, 360t
function of, 360–361
global variables, 361
virtual holocam program, 561–563

- Maxwell, James Clerk, 42
 Maxwell's equations, 237, 386, 388, 391, 394–395, 426, 429
 for arbitrary index profile, 427
 Mazzero, François, 20
 McCormick, Sharon, 34
 McDonnell Douglas Electronics Co., 6, 14, 30
 Mechanical plate displacement system
 RGB-pulsed laser triple-beam DWDH printers, 208–209
 Meder, Bill, 551
 Meyruies, Patrick, 16
 Microchip lasers, 81
MIT Museum, 540
 Modulation transfer factor, 92
 Modulation transfer function (MTF), 91–92, 92f
 Monochromatic holography, 124–125
 Monochromatic light source, 574–575, 574f–575f
 design of main lens surface, 577–578, 577f–578f
 Monochromatic recording, Lippmann photography, 46
 Monochromatic replay, Lippmann photography, 46–47
 Monochrome reflection holograms, 26–27
 Moore, Lon, 31
 MOPA. *See* Master oscillator power amplifier (MOPA)
 architecture
 Moreau, Jean-François, 20
 Moree, Samuel, 540
 Morris, Philip, 537
Mother and Child, 10
 Movie films, 33–34
 MPG H . *See* Multiple photo-generated holography (MPGH)
 MTF. *See* Modulation transfer function (MTF)
 Multicolour gratings, 418
 N-PSM theory *vs.* RCW theory for, 606–608
 Multiple gratings
 generated by many object points, 414–418, 415f
 RCW theory of, 601–608
 recording of
 complex wave fronts and, 414
 multiple colour gratings, 418
 Multiple photo-generated holography (MPGH), 35, 36
 Multiplex holograms, 34–36
 Multiprojector autostereoscopic systems, 522–523
 Munday, R., 16, 36, 609, 612
Musée Française de l'Holographie, 540
Museum für Holographie & Neu Visuelle Medien, 540
Museum of Holography, 540
 Museums, holographic, 540
 Muth, August, 32
 MWDH. *See* Master-write digital holography (MWDH)
 MWDH achromats, 276
 MWDH reflection holograms, 312–313, 313f
- N**
- Nanosecond-pulsed lasers
 use in DWDH printers, 202–203
 National Centre of Photography, 537
 National Ignition Facility (NIF), 192
 Nauman, Bruce, 7
- N-Coupled wave theory, 414–418. *See also* N-PSM theory
 Nd³⁺-doped ceramic YAG, 165
 Nd:YAG. *See* Neodymium-doped yttrium aluminum garnet (Nd:YAG)
 Neodymium-doped glasses (Nd:glass), 164–165, 165t
 Neodymium-doped yttrium aluminium perovskite (Nd:YAP), 164
 Neodymium-doped yttrium aluminum garnet (Nd:YAG), 163–164, 163f
 Neodymium-doped yttrium lithium fluoride (Nd:YLF), 164
 Neodymium-doped yttrium vanadate (Nd:YVO₄), 193
 Neodymium (Nd) lasers, 162–172
 amplification, 167–169
 applications, 171–172
 cooling, 170
 design, 165–167, 166f
 higher energy, 170–171
 master oscillator, 167
 for portraiture applications, 16–17
 pump electronics, 169–170
 RGB lasers, 176–187
 blue lasers, 187
 dual linear-cavity pulsed RGB Nd:YAG laser, 179–183
 dual-ring cavity pulsed RGB Nd:YLF/YAG laser, 177–179
 green lasers, 183–187
 red lasers, 187
 Neuhauss, Richard, 50–51
The New York School of Holography, 540
 Nicholson, Ana Maria, 14, 15f
 Nicholson, Peter, 14, 15f
 Nine-diode framing light source, commercial design of, 594–595, 594f
 Nippon Paint Co., Ltd., 118
 Non-paraxial behaviour
 transmission/reflection holograms, 394
 Non-paraxial I-to-S transformations, 288–292, 291f
 Normal-incidence reflection grating, 426
 N-PSM theory, 455, 456
 vs. RCW theory, 604–605
 for multicolour gratings, 606–608
- O**
- Object beams
 paraxial theory of thin transmission hologram, 373–376, 374f, 375f
 Object beam subsystem
 RGB-pulsed laser triple-beam DWDH printers, 210–211, 211f
 Oblique incidence, unslanted reflection grating at, 435–445
 comparison of Kogelnik's theory with PSM theory, 443–445
 for multiple colour gratings, 442–443
 π-polarisation, differential equations for, 438–439
 σ-polarisation, differential equations for, 435–438
 PSM equations, 439–440
 for sinusoidal gratings, 440–442

- Off-axis transmission hologram, 6
“Off-Bragg” parameter, 400
Ohm’s law, 394–395
Olson, Bernadette, 16
Olson, Ron, 16
One-colour laser, transformations for, 281–283
One-dimensional coupled wave theory, 394–395
One-step CW laser HPO DWDH monochrome reflection hologram, 224f
One-step digital holograms. *See* Direct-write digital holography (DWDH)
123D Catch, 610
OPO. *See* Optical parametric oscillator (OPO)
OPSL. *See* Optically pumped semiconductor lasers (OPSL)
Optically pumped semiconductor lasers (OPSL), 614
Optical parametric oscillator (OPO), 81–82
Optical transfer function (OTF), 91
Optimal reference replay angle, calculation of, 562–564, 563f, 564f
Orange emission DPSS lasers, 80
Orr, Edwina, 16, 24, 540
ORWO FilmoTec GmbH, 109, 115
Osanlou, Ardie, 551
OTF. *See* Optical transfer function (OTF)
- P**
- Painted Margot*, 16, 16f
Panchromatic DCG, 112–114
Panchromatic film, 549–550
Panchromatic photopolymers, 117, 118
Parallel stacked mirrors (PSM) model, 425–459. *See also* N-PSM theory, Coupled wave theory and N-Coupled wave theory
permittivity distribution, 425
PSM and Kogelnik’s theory compared with RCW theory, 604
PSM treatment of slanted monochromatic and polychromatic gratings, 445–451
PSM treatment of the unslanted reflection grating at normal incidence, 425–435
PSM treatment of the unslanted reflection grating at oblique incidence, 435–445
simplified PSM of the thick-phase hologram, 386–390, 387f–389f
Parc des Folies à la Villette, 21, 21f
PB. *See* Polarising beam splitter (PB)
Perry, Hart, 34
Perry, John, 20
Petricevic, Vladimir, 81
Phillips, Nicholas (Nick), 8, 26, 311, 512, 537
PhotoFly, 610
Photogrammetry, 610
Photographer’s Gallery, 537
Photography in Natural Colours (Zenker), 42
Photolytic silver, 98
Photopolymer holograms, 515–517
Photopolymers, 90, 161, 203
advantages, 115
DAROL films, 118
Lippmann photography, 59–60, 60f
manufacturer, 118
panchromatic, 117, 118
parts, 115
with pulsed lasers, 116–117
types, 115
Photoresist, anomalously high pulsed sensitivity of, 616
Physical scene, image acquisition from. *See* Image data, from physical scene
PicoScan, 347, 347f
Piezo element
cavity length stabilisation in pulsed neodymium lasers by, 557–559, 558f
Pioneer’s compact holographic printer, 617–618, 618f
Pixel Systems Inc., 547
Planar gratings
diffractive efficiency of
rigorous chain matrix method, 421–422
rigorous coupled wave theory and rigorous modal theory for, 421
Point-source approximation, geometrical ray tracing in, 580–581, 580f
Poitevin, Alphonse-Louis, 41
 π -polarisation, 438–439
 σ -polarisation, 435–438
Fresnel amplitude coefficients for, 435–436
PSM equations for, 446, 452
Taylor expansion for, 437
Polarising beam splitter (PB), 199
Pole, R. V., 197
Pollia condensata, 609, 610f
Polychromatic hologram, 454
Polychromatic light source, 575–577, 576f
Polychromatic recording, Lippmann photography, 47
Polychromatic replay, Lippmann photography, 48–49
Polychromatic slanted reflection grating, 451
Polyester, 103
Polyethylene film, 117
Polygramma, 118
Portrait
full-colour pulsed, 152–153
neodymium (Nd) lasers, 16–17
Reagan, Ronald, 17, 17f
Power conservation, holographic grating and, 398
Power density distribution
calculation at target plane, 582–583, 583f
Power Imaging Ltd., 550, 551, 552
Poynting’s theorem, 398
Praseodymium-doped lasers, 80–81
Prespeck, 97
Primary colours, 136
Princess of the Iris (Nolte), 150, 151f
Printer optical objectives, correcting inherent distortion in, 283
data predistortion, 285–288, 287f
formalism setting, 283–284, 284f, 285f
non-paraxial I-to-S transformations, 288–292, 291f
Pseudo-colour reflection holograms, 31–33

Pseudo-colour transmission technique, 18, 19
PSM theory. See Parallel stacked mirrors (PSM) model
Psychedelic Amy hologram, 515
 Pulsed holographic portraits, 14–17, 152–153
 neodymium lasers, 16–17
 Reagan, Ronald, 17, 17f
 Pulsed lasers, 157–193
 application, 157–158
 chromium forsterite ($\text{Cr}_4\text{:Mg}_2\text{SiO}_4$), 189–191
 DCG, 111
 diode-pumped, 191–193
 neodymium (Nd) lasers, 162–172
 amplification, 167–169
 applications, 171–172
 cooling, 170
 design, 165–167, 166f
 higher energy, 170–171
 master oscillator, 167
 for portraiture applications, 16–17
 pump electronics, 169–170
 overview, 157–158
 photopolymer holograms, 116–117
 RGB lasers, 176–187
 ruby ($\text{Cr}^{3+}\text{:Al}_2\text{O}_3$) lasers, 158–162
 design, 160–161, 160f, 161f
 energy level diagram, 158, 159f
 parameters, 160t
 pulse lengthening, 161–162, 162f
 pumping bands, 159
 Q-switched, 159, 161
 single-shot high-energy, 191
 stimulated Raman scattering (SRS), 172–176
 red–green pulsed laser, 173–176
 titanium sapphire ($\text{Ti:Al}_2\text{O}_3$) laser, 187–189
 Cr:LiSAF, 187, 188, 189f
 Czochralski method, 187
 energy level diagram, 187–188, 188f
 use in DWDH printers, 202–203
 vs. CW lasers, 158
 Pulse lengthening, in ruby lasers, 161–162, 162f

Q

QinetiQ system, 523–524, 525
 Q-switching systems, 203
 Queen Elizabeth II portrait
 achromatic portrait, 612, 613f
 Jersey Postage stamp, 612, 613f

R

RabbitHoles Media Inc., 509, 511, 552
 Rail-mounted camera systems, use of, 334, 339
 Rainbow and achromatic transmission holograms, 18–24
 DWDH, 276–283, 374
 RGB laser use, I-to-S transformation and, 278–281, 279f, 280f
 synthesis of vertical slit, 277–278, 277f
 transformations for one-colour laser, 281–283

MWDH, 268–276
 achromatically tilted component-colour H_1 masters, 275
 achromatically tilted RGB H_1 master, 276
 component-colour H_1 masters, 271–274, 273f
 vertically aligned RGB H_1 master, 268–271, 269f, 271f, 272f
 MWDH achromats, 276
 Rallison, Richard, 32
 Raman-Nath diffraction, 402
 Ratiu, Ion, 538
 Raven Holographics Ltd., 32
 Ray intersection, with target plane, 581–582, 582f
 Rayleigh, Lord, 42
 RCW theory. *See Rigorous coupled wave (RCW) theory*
 Reagan, Ronald, 17, 17f
 Real image, in thin transmission hologram, 373
 Real-time true holographic displays, 523–525
 Recent developments, 609–622
 achromatic portrait of the Queen, 612, 613f
 CGH composite reflection hologram printer, 616–617, 617f
 3D display systems
 HoloVizio from Holografika, 618, 619f
 SeeReal Technologies holographic 3D display, 618–621, 620f
 hologram postage stamp of the Queen, 612, 613f
 HoloKit from Liti Holographics, 621–622
 lasers for colour holography
 from Cobolt, 614
 from Coherent, 614
 from Lasers Quantum, 614
 RGB Laser Technology at Geola, 614–615, 615f
 new equipment, materials, techniques, and applications, 609
 new holography camera from HiH, 621, 621f
 photogrammetry, 610
 Pioneer’s compact holographic printer, 617–618, 618f
 recording materials and recording techniques, progress in
 anomalously high pulsed sensitivity of photoresist, 616
 surface plasmon waves, 615–616, 616f
 Zebra Imaging, progress at, 610
 Reciprocity failure, 97–99
 HIRF, 97, 98, 99
 holographic, 99
 LIRF, 97, 99
 Recording materials/techniques, 89–90
 dichromated gelatin (DCG), 90. *See also* Dichromated gelatin (DCG)
 photopolymers, 90. *See also* Photopolymers
 recent developments
 anomalously high pulsed sensitivity of photoresist, 616
 surface plasmon waves, 615–616, 616f
 silver halide, 89. *See also* Silver halide
 Red emission DPSS lasers, 80
 Red lasers, 187
 Reference beam subsystem
 RGB-pulsed laser triple-beam DWDH printers, 211–212

- Reflection colour holograms, 143–145
Reflection holograms, 384–386, 385f
 boundary conditions, 397–398
 diffraction efficiency for, 398–399, 399f
 emulsion swelling and change in refractive index on processing, 393–394
 lossless, behaviour of, 405–413, 409f–412f
 non-paraxial behaviour and digital image predistortion, 394
Reflection holographic grating, 425
Reflection holography, 310–311
 vs. transmission holography, 312
Reflective light-source, 595–596
Refractive index, change in
 thick transmission hologram, 393–394
Resolution, 94–96
 recording material, determination by, 95–96
 requirements, of HPO holography, 330–331
Reuterswärd, Carl Fredrik, 7, 12, 13f, 15f
RGB laser(s), 176–187
 dual linear-cavity pulsed RGB Nd:YAG laser, 179–183
 dual-ring cavity pulsed RGB Nd:YLF/YAG laser, 177–179
 illumination, 523
 I-to-S transformation, 278–281, 279f, 280f
 progress in technology at Geola, 614–615, 615f
 short-cavity
 blue lasers, 187
 green lasers, 183–187
 red lasers, 187
RGB LED hologram illumination source, 573–596
 computational results (case studies), 586–594
 main lens surface, design of, 577
 geometrical ray tracing in point-source approximation, 580–581, 580f
 mathematical formulation, 578–580, 579f
 monochromatic light source, 577–578, 577f–578f
 numerical solution, 583
 power density distribution calculation at target plane, 582–583, 583f
 ray intersection with target plane, 581–582, 582f
 ray tracing with finite source size, 583–586
 monochromatic light source, 574–575, 574f–575f
 nine-diode framing light source, commercial design of, 594–595, 594f
 overview, 573–574
 polychromatic light source, 575–577, 576f
 reflective light-source, 595–596
RGB pulsed laser digital holographic printer (XYZ), 550f
RGB-pulsed laser triple-beam DWDH printers, 206–207, 208f, 209f
 alignment, 213
 commercial DWDH printers based on 2001 Geola printer, 215, 215f
 conjugate and non-conjugate operation geometries, 213–214, 214f
 control and video image stream system, 207–208, 209f
 energy requirements, 213
 laser and optical system, 210, 211f
 laser stability issues, 214–215
 lens parameters for telecentric afocal reversing system and Fourier transform objective in, 220t–221t
 mechanical plate displacement system, 208–209
 object beam subsystem, 210–211, 211f
 reference beam subsystem, 211–212
 system architecture, 209–210
Richardson, Martin, 16
Richmond Holographic Studios, 540
Rigorous chain matrix method
 diffractive efficiency of planar gratings, 421–422
Rigorous coupled wave (RCW) theory, 455–456
 boundary conditions, derivation of, 602–603
 diffractive efficiency of planar gratings, 421 equations
 derivation of, 601–602
 numerical solution of, 603–608
Kogelnik's theory vs. parallel stacked mirror theory with, 604
N-PSM theory vs., 604–605
 for multicolour gratings, 606–608
of simple and multiplexed gratings, 601–602
 simple non-multiplexed grating, 602
Rigorous modal theory
 diffractive efficiency of planar gratings, 421
Rinehart, Robert, 6
Ritzau, Teit, 538
Rodin, Aleksej, 17, 545, 546, 547, 548
Romanian exhibition of large-format holography, 547
Ross, Jonathan, 538
Rothé, Edmond, 52–53, 53t
Royal College of Art, 540
Ruby ($\text{Cr}^{3+}:\text{Al}_2\text{O}_3$) lasers, 158–162
 design, 160–161, 160f, 161f
 energy level diagram, 158, 159f
 parameters, 160t
 pulse lengthening, 161–162, 162f
 pumping bands, 159
 Q-switched, 159, 161
Runge–Kutta integration, 603
Runge–Kutta method, 431
Rus, Roman, 548, 552
- ## S
- Sapan, Jason, 34
Saturation, 136
Sazonov, Yuri, 104, 508, 549, 550, 552
Schinella, Robert, 6
Schmidt, David, 34
School of the Art Institute of Chicago, 540
Schweitzer, Daniel, 540
Science Year Annual, 7
Screen-based hogel formation systems
 single-beam RGB DWDH printers, 223, 225–226, 225f
Sculptures de Lumière, 537
Sealing, colour holograms, 149–150
Seebeck, Thomas Johann, 41
SeeReal Technologies holographic 3D display, 609, 618–621, 620f

- Semiconductor diode lasers, 82–85
 amplification, 84–85
 mode structure and coherence, 83
 operation and construction, 82, 82t, 83f
 single longitudinal mode, 83–84
- Sfera-S AO, 104, 508, 549–550
- Short-cavity RGB lasers
 blue lasers, 187
 green lasers, 183–187
 red lasers, 187
- Siebert, Lawrence, 6, 7f
- Siegel, Keeve (Kip), 6, 7
- Siggraph 2012* Las Vegas exhibition, 609
- Silver bromide (AgBr), 91
- Silver chloride (AgCl), 41, 91
- SilverCross emulsion project, 105–109
 coating stage, 108, 108f–109f
 precipitation stage, 106, 107f
 washing stage, 107, 107f–108f
- Silver halide, 89
 emulsions, 91, 93–96
 demands on, 93–94, 94f
 film substrates, 102, 102f, 103
 glass plates, 102, 102f, 103
 grain sizes, 91, 92t
 hypersensitisation, 100
 image resolution, 94–96, 95f, 96f
 latensification, 100–101
 manufacturing companies, 104–105
 MTF, 91–92, 92f
 OTF, 91
 sensitivity, 93, 104
 SilverCross project, 105–109
 substrates, 102–103
 grain sizes, 91
- Lippmann photographs, 60–63, 61f–62f, 63f–64f
- manufacturing companies, 104–105
- reciprocity failure, 97–99
 HIRF, 97, 98, 99
 holographic, 99
 LIRF, 97, 99
 types of, 91
- Silver halide-sensitised gelatin (SHSG) technique, 128–132
- Simple non-multiplexed grating
 RSW theory and, 602
- Simple translating camera
 configuration, image-planed DWDH HPO holograms, 248–256, 248f, 252f, 254f
 image data by HPO holograms from physical scene, 318–319, 318f
- Singapore Military, 21
- Single-beam RGB DWDH printers, 223–226
 screen-based hogel formation systems, 223, 225–226, 225f
- Single-layer silver halide emulsions, 132–134
- Skokov, Gleb, 545
- Slanted reflection gratings, 445–448
 single colour, 446–448
 in three dimensions, 448–449
 unprimed Cartesian system for, 445
- Slavich Joint Stock Company, 104, 115
- Slavich VRP-M, 419, 420
- SLMs. *See* Spatial light modulators (SLMs)
- SLM TEM₀₀ CW laser, 553
- Smigelski, Paul, 16, 34
- Snell's law at the air–hologram boundary, 382–383, 390, 417, 439, 581
- Sony Corporation, 222
- Source-size blurring, 310
 paraxial theory of thin transmission hologram, 376–379, 377f, 378f
- Sowdon, Michael, 540
- Space Bomb*, 16
- Spatial Imaging Ltd., 331
- Spatial light modulators (SLMs), 198, 203, 214, 219, 239, 241, 242, 283–285, 562, 597
 centred SLM configuration, of image-planed DWDH HPO holograms, 260–263, 260f
- Spatially multiplexed monochromatic gratings, 451–455
 N-PSM equations for, 452
- Speckle, 310
 problem of, laser transmission holograms and, 380–381
- Spectrol hologram*, 539
- Spierings, Walter, 35, 198
- SPP. *See* Surface plasmon polariton (SPP)
- Standa UAB, 331
- “Start Render” event handler, writing, 365–368, 366f, 367f
- Static holography, 517
- Static reference beam, 217
- Statistical optimisation algorithm
 cavity length stabilisation in pulsed neodymium lasers, 558–559
- Stephens, Anaït Arutunoff, 16
- Stereograms, holographic, 197–198
- Stereographic holograms, 34–36
- Stimulated Raman scattering (SRS), 172–176
 red–green pulsed laser, 173–176
- STOCKHOLMIA'74*, 11
- “Structuredlight scanner,” 346–347, 347f, 348f
- Subbing, 103
- Substrates, for holographic emulsions, 102–103
- Super-realistic high virtual volume (HVV) displays, 518
- Surface plasmon polariton (SPP), 615, 616f
- Surface plasmon waves, 609, 615–616, 616f
- Swedish Post Authority, 11, 12
- Synthetic aperture radar (SAR), 511

T

- Taylor, Rob, 17
- Taylor series, 427
- TEA. *See* Triethanolamine (TEA)
- There's No Place like Home*, 8
- Thick-phase hologram. *See also* Parallel stacked mirrors (PSM) model, Coupled-wave theory and Kogelnik's theory
 PSM model of, 386–390, 387f–389f
 simplified PSM model of, 386–390, 387f–389f
- Thick transmission hologram
 emulsion swelling and change in refractive index on processing, 393–394

- non-paraxial behaviour and digital image predistortion, 394
 - three-dimensional theory of, 381–384
 - blurring in, 383–384
 - Bragg condition, 382
 - Snell’s law at the air–hologram boundary, 382–383
 - Thin transmission hologram
 - non-paraxial behaviour and digital image predistortion, 394
 - three-dimensional paraxial theory of, 370–380, 371f
 - chromatic blurring, 379–380, 380f
 - collimated reference and object beams, 373–376, 374f, 375f
 - real image, 373
 - source-size blurring, 376–379, 377f, 378f
 - virtual image, 373
 - 3D display systems
 - HoloVizio from Holografika, 618, 619f
 - SeeReal Technologies holographic 3D display, 618–621, 620f
 - Three-dimensional theory
 - of thick transmission hologram, 381–384
 - blurring in, 383–384
 - Snell’s law at the air–hologram boundary, 382–383
 - of thin transmission hologram, 370–380, 371f
 - chromatic blurring, 379–380, 380f
 - collimated reference and object beams, 373–376, 374f, 375f
 - source-size blurring, 376–379, 377f, 378f
 - 3D Print UAB, 550
 - 3D structured-light scanners
 - full-parallax image data acquisition from physical scene, 346–347, 347f, 348f
 - 3D StudioMax, 317, 348, 349
 - creating 20 cm × 30 cm landscape hologram, 349–357, 350f–356f
 - Through the Looking Glass*, 537
 - Titanium sapphire (Ti:Al₂O₃) laser, 187–189
 - Cr:LiSAF, 187, 188, 189f
 - Czochralski method, 187
 - energy level diagram, 187–188, 188f
 - To Absent Friends*, 8
 - Tonka Toys hologram, 20, 20f
 - Transmission coefficients, 426
 - Transmission colour holograms, 143
 - Transmission holograms, 450. *See also* Thick transmission hologram; Thin transmission hologram
 - boundary conditions, 397–398
 - diffraction efficiency for, 398–399, 399f
 - lossless, behaviour of, 401–405, 403f, 404f, 406f, 407f
 - Transmission holography, 311–312
 - vs. reflection holography, 312
 - Trayner, David, 24
 - Tribillon, Louis, 7
 - Triceratops, 36, 37f
 - Triethanolamine (TEA), 100
 - water solution, 100
 - water treatment, 100
 - Trinne, Mark, 547
 - Triple-beam DWDH printers, 205–215
 - hogel-writing sequence, 205–206, 206f, 207f
 - RGB-pulsed laser, 206–207, 208f, 209f
 - alignment, 213
 - commercial DWDH printers based on 2001 Geola printer, 215, 215f
 - conjugate and non-conjugate operation geometries, 213–214, 214f
 - control and video image stream system, 207–208, 209f
 - energy requirements, 213
 - laser and optical system, 210, 211f
 - laser stability issues, 214–215
 - mechanical plate displacement system, 208–209
 - object beam subsystem, 210–211, 211f
 - reference beam subsystem, 211–212
 - system architecture, 209–210
 - Tristimulus values, 136
 - Tsujiuchi, Jumpei, 34
 - Two-dimensional tracking camera system, 499
- ## U
- Ultimate Holography, 105
 - Ultrafine-grain holographic emulsions, 104
 - Ultragram, 198
 - Ultra-realistic DWDH printers, 226–228
 - 3N-objective printers, 228, 228f
 - Ultra-thick transmission hologram
 - blurring in, 383–384
 - University of Hawaii, 14
 - Unslanted multiple-colour gratings at normal incidence, 430–435
 - Bragg resonance of multiplexed gratings, 430–431
 - N*-chromatic grating, 431
 - polychromatic unslanted index profile, 430
 - region of *j*th Bragg resonance, 431
 - for three-colour, 433–435
 - for two-colour, 431–433
 - Unslanted reflection grating
 - comparison with Helmholtz equation, 429–430
 - at normal incidence, 425–429
 - at oblique incidence, 435–445
 - rotating to slanted grating in primed system, 445
 - Upatnieks, Juris, 3, 4f
 - Updateable holographic display, 521–522
 - Ultimate Holography North America (UHN), 516
- ## V
- Vanin, Valery, 27
 - Variable-angle reference beam
 - DWDH printers incorporating, 217–219, 222f, 223f, 224f
 - Venus de Milo*, 7, 8f
 - Verdi 532 nm laser, 78, 79f
 - Vergnes, Florian, 544, 545, 547
 - Vidmer, Nataly, 548
 - Viewing plane, digital holograms, 242–243, 243f
 - Virtual computer model, images derived from. *See* Image data, from virtual computer model

Virtual holocam program, MAXScript, 561–563
Virtual image, in thin transmission hologram, 373
Virtual space displays, 518
VISIO 20 real-time 3D holographic display system, 609
Volume gratings, 58, 418
 solving Helmholtz equation in. *See* Helmholtz equation
Von Goethe, Johann Wolfgang, 41
Vorobiov, Sergei, 545–546
Voxgram, 25, 25f
Vulcan, 192

W

Wavelength selection, colour holograms, 137–142
Webster, John, 8, 10
Wenyon, Michael, 540
Wiener, Otto, 42
Wolff, John, 537
Woodd, Peter, 539
World Book Encyclopaedia, 7
Wuerker, Ralph, 10

X

Xacti HD electronic camera, 331
Xenon lamp, 191
XYZ Imaging Inc., 199, 201, 215, 216, 507–511, 549
 first commercial RGB pulsed laser digital holographic printer, 550f
 sale of, 552
 start of, 547–548

Y

Yellow emission DPSS lasers, 80

Z

Zacharovas, Stanislovas, 548, 549, 550, 551
Zebra Imaging Inc., 118, 158, 199, 311, 507, 549, 609
 progress at, 610, 611f–612f
 Zscape™ Preview software, 610
Zellerbach, Gary, 539
Zenker, Wilhelm, 41–42
Zscape™ print technology, 610