Chapter 11: Non-Laser Illumination of Holograms

Introduction

Laser-lit off-axis transmission holograms remain the "holographer's holograms," even today. The images are razor sharp, and can reach from the tip of your nose to the horizon (we will not dwell on the drawbacks of "laser speckle" at this point). But laser illumination of holograms presents some serious practical problems in many image display environments. Laser light is still very expensive, in terms of dollars per lumen, and very uncommon—compared to sunlight, for example. Laser light is unreliable: most big lasers have a specific startup procedure that must be followed, one that involves cooling water, etc., and many subsystems that may go sour. Also, the beam itself is typically expanded and cleaned with a spatial filter that needs routine cleaning and tweaking for best performance. And, the Bureau of Radiological Health has deemed almost all lasers to be *death rays.* The amount of paperwork required to provide laser illumination for large-scale holograms is incredible! The State of New York requires a "mobile laser operator's license" for everyone who would plug in and turn on a laser in a public place. The desire to bring holograms out of darkened basement laboratories and into the public's awareness has motivated several approaches to white-light viewable holograms, resulting in the development of the white-light transmission or "rainbow" hologram, and the white-light reflection hologram. We emphasize that these holograms must still be made with lasers; they are specially designed to be viewable in white light. Even that light must have spatial coherence, coming from something approximating a point source (a line source, in some cases). In this chapter, we will look at the problems of trying to view ordinary holograms with coherent and incoherent sources, and examine a few preliminary solutions to the problems involved.

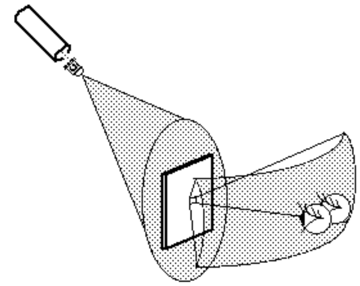
Problems with laser illumination

Let's look again at the laser illumination situation. The most common lasers, by far, are helium-neon gas lasers, our beloved He-Ne type, which are readily available at modest cost for powers up to 5 milliwatts or so. We will see below that 5 mW is just enough to illuminate a 4"x5" hologram under ordinary conditions, which is why we speak of laser light as expensive. Low powers are adequate for darkened laboratory conditions, but we have to be thinking of bringing holograms out into the real world at every opportunity, and then the stakes go up considerably. For large holograms, on the order of a meter square, ion lasers are the only choices as powers on the order of a watt are required. These lasers are expensive to own and to operate (10 kW 3-ø input power, plus cooling water). Within the next few years, diode lasers will offer a reasonable alternative to He-Ne lasers. 10 mW at 635 nm is already available for several hundred dollars. The output beams are markedly elliptical, which requires special optics for shaping into more useful round beams, and the temperature has to be controlled. Diode-pumped frequency-doubled YAG lasers offer a beautiful 532 nm green output with powers of a few hundred milliwatts, and are extremely convenient and reliable to use, but are still very costly compared even to ion lasers. And none of these systems address the chronic problems of laser speckle and safety registration that are bound to arise.

<u>photometry: to get 20 ft-L:</u>

Let's take a moment to examine the question of "how much laser illumination power is needed to provide a reasonably bright image?" First, our conventional method of measuring laser power, in milliwatts, tells us how much heat such a beam can deliver, but doesn't tell us how bright the beam will appear to be (it could be invisible, if it were infra-red or ultra-violet light, for example). The transition takes us from *radiometry* to *photometry*, and both are fairly arcane subjects in their own right. We will try to keep it simple! The conversion from milliwatts to milli-lumens (or "visually apparent power") is via the CIE visibility curve, established in the 1930's and regularly refined ever since. At the peak of the eye's sensitivity, in the green area of the spectrum, one watt of radiation produces 685 lumens. That is our central calibration point. At 633 nm, the eye is about 1/4 as sensitive as at its peak (at 550 nm "green"), and drops off quickly as the wavelength continues to increase. Second, we need to know what a "reasonably bright image" is supposed to be. Here, we can simply rely on previous experience with color television sets, where a peak white is expected to have a luminance of 20 foot-Lamberts (in the old notation), or 70 lumens per square meter per steradian (expressed in SI units, also 70 "nits" or candles/m^2). This is the measure that describes how bright a diffusely-illuminated surface will seem to be; as strange as its units may be, it works pretty well. We would have to take a longer detour through radiometry and photometry than seems justified to explain much more, but at this point we are simply ready to plug in some numbers and look at the results.

4"x5" hologram:



Let’s start with a 100 mm x 125 mm (4" x 5") hologram; for simplicity, lets assume we have 1 mW of He-Ne illumination, or 0.17 lumens. Let’s then assume that the illumination overfills the hologram by a factor of two (to provide more uniform lighting, or because the aspect ratio at 45° is unfavorable; 2X is optimistic, by the way), so that the illuminating intensity averages to be 6.85 lumens/m^2. Let the average diffraction efficiency be 20%, which is plausible for a well-bleached hologram, and let the viewing zone be 60° wide and 30° high (or 0.55 steradians, generous but plausible for a laser transmission hologram). Now we can obtain an overall luminance of 2.5 nits. However, the peak extended-white luminance of an “average” scene is typically five times the average, so we could expect a white surface in the scene to have a luminance of 12.5 nits. To bring this up to the hoped-for 70 nits, we obviously need to increase the illumination five-fold, to about 5.6 mW—not a small He-Ne laser!.

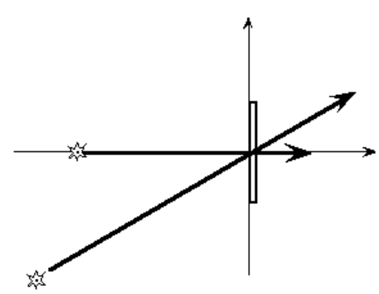
This may come as a surprise to those of us who are used to peering into dim holograms with one milliwatt lasers, but that usually happens in a darkened basement lab, and we are talking here about images bright enough to be seen in ordinary room light situations, alongside televisions and other everyday imaging devices. The power scales with the area of the hologram, so that a 12” x 16” hologram would require about 60 mW, which is getting close to the largest He-Ne lasers made (the venerable Spectra-Physics 125). What are we to do? Some folks have made an art form of limiting the view-zone angle to the minimum acceptable, perhaps 15° across and 10° high, which provides an “antenna gain” of brightness of more than 10X, but limits “look-around,” one of the most charming features of holographic 3-D (limited vertical viewing zones are, however, one of the keys to the brightness of rainbow holograms). Note too that as the image approaches the hologram plane, the brightness of an extended-white area is limited by the diffraction efficiency at that area, and we lose the 5X advantage due to averaging a peak luminance over the entire hologram area. Thus the prospects for bright laser-illuminated holograms are pretty dim indeed!

Of course, higher-power lasers are readily available, but only at substantial expense, and the other problems mentioned above. Practical holographic display has turned instead to non-laser, or “thermal” sources of illumination. The problem is that the output power of such sources increases as their source area and spectrum width increase, both of which cause a degradation in the sharpness of the resulting image. Thus there is an inevitable tradeoff between image brightness and sharpness that is determined by the quality of the illumination. This chapter will examine first the sensitivity of holograms to color blur and source size blur, and then re-examine the qualities of candidate light sources.

Sources of image blur

Our discussion of holography, including image reconstruction, has presumed the use of sources of arbitrarily high spatial and temporal coherence. Which is to say that the sources were perfectly point-like, and monochromatic (indeed, single-frequency). That will typically still be the case for image recording (some relaxations are possible, which we will not have time to describe here), but the fact is that hologram viewing is possible with highly incoherent sources with only mild blurring of the resulting image, under some circumstances. The two dimensions of coherence, temporal and spatial, correspond to independent contributions to image blur, and we will discuss them one at a time.

<u>color blur:</u>

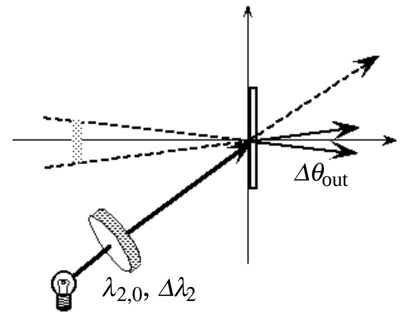


As in many of our hologram analyses, we will consider the angle and distance issues in sequence, corresponding to the grating and Gabor zone plate components of the holographic fringes. As a simplification, we will assume at the end that the object beam angle and central output beam angles are perpendicular to the plate, as is generally the case in display applications. Then, we can expand the output angle as a function of illumination wavelength as

$$\frac{d(\sin\theta_{out})}{d\lambda_2} = \cos\theta_{out}\frac{d\theta_{out}}{d\lambda_2} = m\frac{1}{\lambda_1}\left(\sin\theta_{obj} - \sin\theta_{ref}\right),$$

$$\frac{\sin\theta_{ref}}{\lambda_1} = \frac{\sin\theta_{ill}}{\lambda_2}, \text{ if } \theta_{out,central} = \theta_{obj} = 0, \text{ so} \qquad (1)$$

$$\Delta\theta_{out} = -\sin\theta_{ill}\frac{\Delta\lambda_2}{\lambda_{2,0}},$$



where $\lambda_{2,0}$ is the central wavelength of the illumination, or the filter over the white light source in the sketch. If we consider the angular resolution of the unaided eye to be about one minute of arc (290 microradians), then a hologram illuminated at 45° with 540 nm light would need a spectrum width of less than 0.22 nm to avoid noticeable color blur, a very narrow spectrum indeed! Note that on-axis holograms, with very small θ_{ill}, have much lower color blur than off-axis holograms for the same source bandwidths, as observed earlier.

This analysis gives us the angular subtend of the color blur as seen from the hologram plane. But to gauge the blur at other viewing distances, we will need to know the location of the color blur too. In fact, the color-blurred image of a point source will be tipped at an angle of special interest, which we will come to know as the "achromatic angle," designated by α.

The central location of the image is given by the same "$\cos^2\theta$" equation as before (note that we are concerned with the vertically-focused blur in this case — there is no horizontal component to color blur). Thus

$$\frac{1}{R_{out}} = (1)\frac{\lambda_{2,0}}{\lambda_1}\left(\frac{1}{R_{obj}} - \frac{\cos^2\theta_{ref}}{R_{ref}}\right) + \frac{\cos^2\theta_{ill}}{R_{ill}}, \qquad (2)$$

and the vertical extent of the color blur of the image, or its height, is given by

$$h_{color\ blur} = R_{out}\cdot\Delta\theta_{out,color} = R_{out}\cdot\sin\theta_{ill}\cdot\frac{\Delta\lambda_2}{\lambda_{2,0}}. \qquad (3)$$

the "achromatic angle":

The detailed variation of the image distance with wavelength is given by

$$\frac{d}{d\lambda_2}\left(\frac{1}{R_{out}}\right) = -\frac{1}{R_{out}^2}\frac{dR_{out}}{d\lambda_2} = \frac{1}{\lambda_1}\left(\frac{1}{R_{obj}} - \frac{\cos^2\theta_{ref}}{R_{ref}}\right),$$

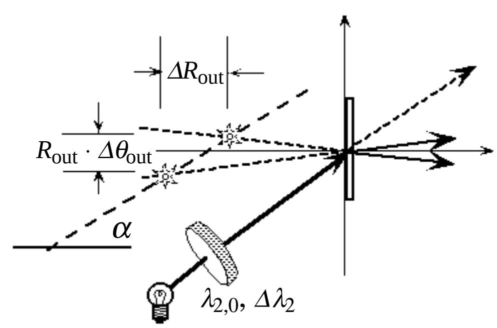
$$\frac{dR_{out}}{d\lambda_2} = -\frac{R_{out}^2}{\lambda_2}\left(\frac{1}{R_{out}} - \frac{\cos^2\theta_{ill}}{R_{ill}}\right), \qquad (4)$$

$$\Delta R_{out} = -\left(R_{out} - R_{out}^2\frac{\cos^2\theta_{ill}}{R_{ill}}\right)\frac{\Delta\lambda_2}{\lambda_2}.$$

The tangent of the angle of the blur image is then given by

$$\tan\alpha = \frac{R_{out}\cdot\Delta\theta_{out}}{\Delta R_{out}}$$

$$= \sin\theta_{ill}\frac{1}{1 - \dfrac{R_{out}\cdot\cos^2\theta_{ill}}{R_{ill}}}. \qquad (5)$$



As a rule, the illumination distance is much greater than the image distance, so that the second term can be ignored and the generally useful relationship becomes,

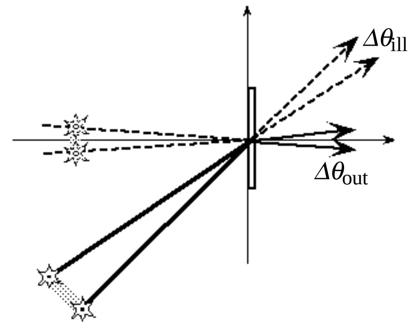
$$\tan\alpha = \sin\theta_{ill}. \qquad (6)$$

A simple off-center zone plate model of an off-axis hologram would predict, for collimated illumination, that the achromatic angle would be exactly equal to the illumination angle. However, this chapter's more careful consideration shows that the color blur is actually tilted significantly more toward the *z*-axis. The difference is large enough to markedly decrease the color blurring in holograms that properly use the achromatic angle concept to help correct color blur (discussed in subsequent chapters).

θ_{ill}	α
30°	26.6°
45°	35.3°
60°	40.9°

<u>source size blur</u>

We have seen before, for on-axis holograms, that motion of the illumination source away from the axis moves the image off of the axis too. For off-axis holograms, the relationship is only slightly different: a vertical motion of the light source produces a vertical motion of the image through an angle that is proportional to the cosine of the illumination angle:



$$\frac{d(\sin\theta_{\text{out}})}{d\theta_{\text{ill}}} = \cos\theta_{\text{out}}\frac{d\theta_{\text{out}}}{d\theta_{\text{ill}}} = \frac{d(\sin\theta_{\text{ill}})}{d\theta_{\text{ill}}} = \cos\theta_{\text{ill}}\ , \qquad (7)$$
$$\Delta\theta_{\text{out}} = \cos\theta_{\text{ill}}\Delta\theta_{\text{ill}}\ .$$

There is also a small variation of the vertically-focused image distance, which we shall consider to be negligible:

$$\frac{d}{d\theta_{\text{ill}}}\left(\frac{1}{R_{\text{out}}}\right) = -\frac{1}{R_{\text{out}}^2}\frac{dR_{\text{out}}}{d\theta_{\text{ill}}} = \frac{1}{R_{\text{ill}}}\frac{d}{d\theta_{\text{ill}}}\left(\cos^2\theta_{\text{ill}}\right), \qquad (8)$$
$$\Delta R_{\text{out}} = \frac{R_{\text{out}}^2}{R_{\text{ill}}}\sin 2\theta_{\text{ill}}\Delta\theta_{\text{ill}} \approx 0\ .$$

To consider the effect of a finite source size, we simply imagine that, instead of moving a single source point over the angle $\Delta\theta_{\text{ill}}$, all those points are present simultaneously, and that they are incoherent with each other so that the images they produce all add in intensity. The result is an enlarged blur spot, with the <u>height</u> of the source-size blur given by

$$h_{\text{source-size blur}} = R_{\text{out}}\cdot\Delta\theta_{\text{out, source}} = R_{\text{out}}\cdot\cos\theta_{\text{ill}}\cdot\frac{\Phi_{\text{ill}}}{R_{\text{ill}}}\ , \qquad (9)$$

where we signify the diameter of the light source by Φ_{ill}.

Motion of the source from side to side gives a variation of output angle equal to the variation of illumination angle (no cosine effect), so that the <u>width</u> of the source-size blur is simply

$$w_{\text{source-size blur}} = R_{\text{out}}\cdot\Delta\theta_{\text{ill}} = R_{\text{out}}\cdot\frac{\Phi_{\text{ill}}}{R_{\text{ill}}}\ . \qquad (10)$$

Thus, to keep the blur due to source size below the perceptible limit of the human eye, the source must subtend an angle of less than one minute in width, and a little over a minute in height. This is about the angular subtend of a quarter-dollar coin at a distance of 82 meters (270 feet)!

The angular sensitivity of the eye puts stringent limitations on the spectrum width and area of conventional "thermal" or non-laser sources of illumination. Virtually monochromatic and point-like sources are needed in the general case! We will go on to consider some of the ways around this apparent dilemma.

Narrow-band illumination

<u>mercury arc</u>

There are a wide range of gases that, under stimulation by electrical currents, give rise to one or more relatively pure spectral lines. Neon is widely seen in advertising signs, and sodium makes an appearance in street lights. However, most of these sources are large-area or large-volume tubes, operating at relatively low pressures. Mercury-vapor lamps, operating at high pressure, are probably the best-known point-like sources of single-color light. The higher the pressure, the smaller the glowing part of the discharge becomes, but the wider the output spectrum becomes also, so there is some room for tradeoffs. The most prominent mercury output spectral lines are “green” at 546 nm, with a line width of about 5 nm, and a “yellow” pair of lines at 577 & 579 nm, which are generally widened to a total of 7 nm. A one-hundred watt lamp radiates about 150 mW/steradian (sr) in the 546 nm line, which comes from a bright spot near the cathode that is about a millimeter across.

filtered white light:
An incandescent tungsten wire emits radiation over a wide spectral band, peaking in the infra-red, but with plenty of energy in the deep blue. The higher the temperature of the tungsten, the greater the number of watts of light per square millimeter (and the more blue light, relative to red) are emitted. However, the lifetime of the wire drops quickly due to evaporation of tungsten from the hot surface. The addition of halogen gases (iodine, especially) and the use of a very hot lamp envelope (quartz, to stand the heat) recycles some of the evaporated tungsten back onto the wire, extending the lifetime of such lamps, especially at high temperatures. Even so, temperatures are limited to about 3400°K for practical lifetimes (a thousand hours or so).

Narrow-band interference filters can be put over the beam to select out a fairly narrow spectral band, but obviously the luminous flux decreases linearly with spectral width (or even faster, as the peak transmittance of such filters decreases for narrow filters; it is only 50% in the best of cases). Filter bandwidths 15 or 20 nm wide are typical of the narrower pass bands.

reflection holograms: __
Rather than put the narrow-band filter in the light source (which is difficult with sunlight, for example), it is possible to put it into the hologram. Very thick transmission holograms can have considerable angle and wavelength selectivity (the Bragg selectivity effects of volume holograms), but it is reflection holograms that offer the highest wavelength selectivity. We will see that their volume structure includes multiple layers of alternating high and low refractive indices, much like a vacuum-coated interference filter, and that they can self-select a narrow (approx. 15 nm wide) portion of the visible spectrum for their reconstruction. Of course, the narrower the reflection spectrum, the slimmer the “photon catch” and the dimmer the image becomes, even if sharper. Thus some reflection holograms are deliberately processed to widen their reflection spectrum to increase brightness at the expense of sharpness.

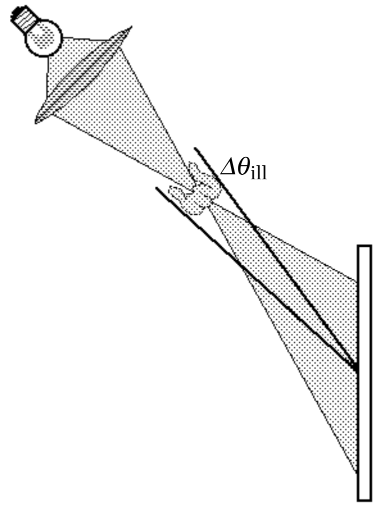
Point-source illumination

The ideal point source of white light simply doesn’t exist. So-called “white” lasers put out three or so wavelengths, not a continuous spectrum (as the sun does). Arc lamps are the next best thing, and have a high enough luminance to be virtual point sources. After that come incandescent lamps. Far behind are fluorescent lamps, and then the holographer’s nightmare, an overcast hazy day!

the sun:
The sun is just another incandescent source, although it has a higher temperature and a higher brightness (about 4000 lumens/m^2-sr). It subtends an angle of about 0.5°, or a solid angle of 76 microsteradians (almost exactly the same as the moon). The luminance of the surface of the sun, as seen directly overhead from the earth’s surface on a clear day, is about 1.6 x 10^9 nits.

high-pressure arc (xenon): __
The discharge near the cathode of these lamps reaches a luminance of 1.8 x 10^8 nits.

zirconium arc:
The tip of the tungsten rod contains zirconium oxide, which can reach a luminance of 4.5 x 0^7 nits, with a particularly small and well-defined area. However, these lamps are getting hard to find.

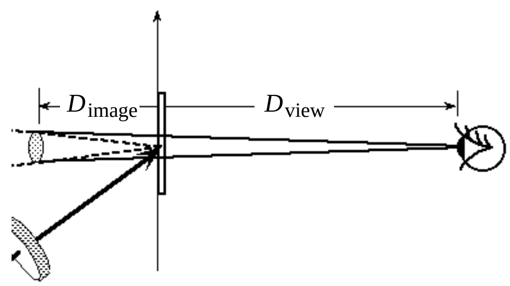


quartz-halogen lamps:

Rather than use a large filament to illuminate a hologram, we can use a small filament and focus the light onto the hologram with a lens or concave mirror. This produces a large-area image of the filament near the output aperture of the illuminator, and the same formulas apply if we use the area of this source image instead of the lamp itself. The “optical brightness” theorem says that the luminance of an extended surface stays constant as that surface is successively imaged through an optical system (assuming that the aperture on the measuring device is the limiting aperture (for example, the pupil of the human eye), which is usually the case). That is, the brightness of the image of the filament (in lumens/m^2-sr) is the same as the brightness of the filament itself. The hologram gets more light because the filament image fills a larger solid angle as seen from the hologram than the filament would at the same distance. There is no other way to get more flux to the hologram surface! The luminance of a tungsten filament reaches about 2.4 10^7 nits.

Image depth effects

So far, we have been thinking of the hologram as a window that the viewer looks through, almost with his or her nose pressed up to it. The height (and width) of the blur image increases linearly with image depth because the image blur subtends a constant angle, $\Delta\theta_{out}$, as seen from the plane of the hologram.



But if we let the viewer back away from the hologram, the angle subtended at the viewer’s eye by the image blur decreases, and it is the blur angle at the eye that must be below one minute of arc for the image to appear sharp. If we designate the viewer’s distance from the hologram as D_{view} and the depth of the image as D_{image} ($= R_{out}$), then the blur angle at the eye (call it ω_{blur}) becomes:

$$\omega_{blur} = \frac{\left(h_{blur} = R_{out} \cdot \Delta\theta_{out}\right)}{D_{view} + D_{image}} = \frac{1}{\left(1 + \frac{D_{view}}{D_{image}}\right)} \Delta\theta_{out} \ . \qquad (11)$$

Thus, if the viewer’s distance is much greater than the image depth, the visual blur decreases significantly. Or, considered differently, much larger sources may be used, allowing more illuminating flux to reach the hologram. For example, if a hologram is viewed at arms length (500 mm) under sunlight (θ_{ill}=45°, $\Delta\theta_{ill}$=0.5°), then the image may be 25 mm deep (behind the hologram, or 23 mm in front) before any visual blurring is perceptible.

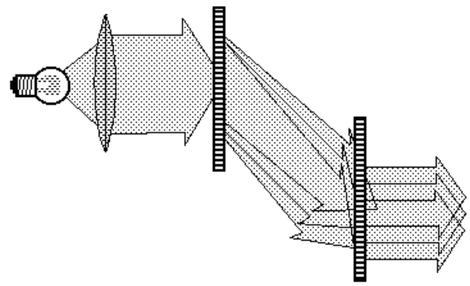
Ordinary television offers a pixel size that is at least double the magic one minute of arc under good conditions. In practice, much greater amounts of blurring are tolerable in holograms too, as long as the visual center of interest is reasonably sharp. But the general trend has become clear: holograms have become things one looks AT rather than THROUGH. They are held at arms length, or viewed on a wall, considered more as a photograph than a porthole into another spatial world.

Other approaches

There are other things we can do besides simply trying to emulate the coherence of laser sources by brute force; we can try to design holograms and systems that are simply designed to work within the limitations of white-light illumination, or even to take advantage of some of its characteristics. Two approaches worth mentioning here are dispersion compensation and parallax limiting.

dispersion compensation: ___

One way of partially overcoming the color blur of a wide-band illumination source is to pre-disperse the various colors from the illuminator so that they are incident on the hologram at angles that result in them all being diffracted at equal angles toward the viewer. This is generally done by diffraction with a pre-grating that has about the same spatial frequency as the hologram, so that rays that start parallel from the illuminator wind up parallel again after diffraction by both gratings. This means that R, G, & B images of a distant point would be superimposed, producing an image free of color blur! Different focal situations can achromatize on-axis images at any chosen depth, or from a viewpoint at a chosen on-axis distance. However, as the parts of the image move from the central location (or the viewer moves from that central location) the blurring (color fringing) will increase—it is not possible to achromatize a large volume of image space in this way, only specific points.



parallax limiting: ___

One way to think about vertical color fringing, which is the main effect of color blur, is that differing vertical perspectives are becoming mixed in different wavelengths. Objects at different depths shift by different amounts with respect to the hologram plane as the wavelength changes; they rotate and move up or down. If we see more than one wavelength at a time, we see more than one image at a time, and the difference of their locations causes a blurring (we will revisit this point of view in two more chapters). So, one way of eliminating this source of blur is to eliminate all but one perspective view in the vertical direction so that no rotation or shift is possible, and hence no blurring. This is one way of looking at the principle of white-light transmission "rainbow" holograms, which we will be talking about in detail fairly soon.

Conclusions

We may never look at a hologram with a laser again in class. If we do, we should remind ourselves what a rare pleasure it is, one that few civilians will ever enjoy. Laser-illuminated holographic images can extend from the tips of our noses to the far horizon, with exquisite sharpness at every depth. But from here, we will be moving toward white-light viewed images, for which we must give up those extravagant vistas for a "space in a box" that we can hold in our hands. It will be impressively and realistically deep, and perhaps provoke new kinds of spatial thinking, but don't forget to drag out your laser from time to time for a look at "the real thing!"