Chapter 2

Background

This chapter includes background information on 3-D displays, optical and computational holography, past attempts at holographic information reduction, and useful computational techniques. These concepts are central to the development of diffraction-specific computation and holographic encoding.

2.1 Human Visual System

The goal of a holographic display is to produce 3-D images for viewing by the human visual system (HVS). Optically produced holographic images generally exhibit resolution far beyond the abilities of the HVS. This is essentially why so much of holographic bandwidth is wasted: the HVS simply cannot make use of the information contained in an image with micron-scale resolution. Diffraction-specific computation and the holographic encoding schemes developed from it exploit the limited capabilities of the HVS. Holovideo should provide only visual information that is useful to the HVS and no more. An information-efficient display system must stimulate the viewer in the most efficient way possible[8]. It is important to begin the development of these novel holographic computation techniques with a discussion of the performance of the human visual system.

2.1.1 Acuity

The lateral and depth spatial resolutions (acuities) of the HVS vary with viewing conditions such as brightness and motion. The values listed here are typical values that were derived empirically[8]. The lateral acuity is approximately one minute of arc (1/60 degree of arc) or 290 microradians. A typical viewing situation places the viewer at about 600 mm from the display. At this distance, the HVS can resolve
600 x 0.000290 = 0.175 mm. A point of light that is angularly smaller looks identical

to a 0.175-mm spot. Two small spots confined to less than 0.175 mm appear as a single point of light. Holography is capable of producing spots that are 100 times smaller than 0.175 mm. Therefore, a holographic encoding scheme is allowed to add small amounts of blur to image points, as long as the total blur is imperceptible to the HVS, i.e., less than 0.175 mm.

Depth acuity at a viewing distance of 600 mm is approximately 0.75 mm. Holographic imaging can produce depth resolutions that are 100 times smaller than this. Again, this is wasted bandwidth, and an opportunity for bandwidth compression.

2.1.2 Pupil Size

As a viewer moves throughout the viewing zone, the eyes see different views of the image or object. The number of distinguishably different views (called the *parallax resolution*) depends on the size of the pupil of the eye. Typically the diameter of the pupil is 3 mm. At a viewing distance of 600 mm, the useful angular resolution of an image is arctan(3/600) = 5 milliradians = 0.3 degrees. An optically produced hologram provides a continuously varying parallax, i.e., the perspective views of the image vary continuously as a function of viewing angle as does a real object. This is another way in which a holographic image provides too much information for the HVS.

2.1.3 Depth Cues

The HVS uses a number of depth cues to resolve the depth, shape, texture, and relative distance of scenes. The oculomotor cues are physiological cues that derive depth through physical movements in the muscles of the eyes. Oculomotor depth cues include accommodation (focusing cues) and convergence (triangulation between the two eyes). Stereopsis is the highly acute sensation of depth that results from the depth cue of binocular disparity, in which each eye sees a slightly different view of the scene. Motion parallax is a depth cue sensed from the apparent change in the lateral displacements among objects in a scene as the viewer moves. Motion parallax allows the viewer to move around the object scene, as is the case when viewing real scenes.

The remaining depth cues are pictorial depth cues: they can be found in two-dimensional (2-D) images and pictures. These are monocular depth cues (as is motion parallax) because they can be sensed with a single eye. The strongest pictorial depth cue - perhaps the strongest depth cue - is occlusion (also called overlap). The HVS derives relative depth information when one part of an image is obstructed by another overlapping part. The HVS uses the kinetic depth cue to derive shape and depth information by observing the relative motions of parts of a scene. The remaining pictorial depth cues (linear perspective, texture gradient, aerial perspective, shading, relative sizes, etc.) are discussed in the book by Okoshi[8].

Understanding the visual depth cues is particularly important when comparing holography to other types of 3-D display technologies. Holographic displays require a large amount of computation and display bandwidth. The value of holovideo must be compared to the relatively easier existing 3-D displays.

2.2 Three-Dimensional Displays

Three-dimensional displays[8,9] are generally electronic devices that provide binocular depth cues, particularly convergence and binocular disparity. They also provide some or all of the pictorial depth cues. Some 3-D displays provide additional depth cues such as motion parallax and ocular accommodation. (The reference by McKenna[9] contains a good discussion of the visual depth cues and a detailed evaluation of 3-D display techniques.) A 3-D display enables the viewer to more efficiently and accurately sense both the 3-D shapes of objects and their relative spatial locations, particularly when monocular depth cues are not prevalent in a scene. When viewing complex or unfamiliar object scenes, the viewer can more quickly and accurately identify the content of a scene. Therefore, 3-D displays are important in any application involving the visualization of complex 3-D data, including telepresence, education, medical imaging, computer-aided design, and scientific visualization.

The two types of 3-D displays are *stereoscopic* and *autostereoscopic*. A stereoscopic display presents different views of the imaged scene to the left and right eyes. Examples include boom-mounted, head-mounted, and displays using polarizing glasses. Other than pictorial depth cues, these displays fundamentally add only binocular disparity. Motion parallax can be simulated with the addition of head-tracking. An autostereoscopic display is a 3-D display that does not require special viewing aids. Examples include lenticular, parallax barrier, slice stacking, and holography, each supplying different depth cues to the human visual system. A good autostereoscopic display provides motion parallax by presenting more than two views of the imaged scene.

The merit of a 3-D display depends primarily on its ability to provide depth cues and high resolutions. The inclusion of depth cues - particularly binocular disparity, motion parallax, and occlusion - increases the realism of an image. Holography is the only imaging technique that can provide all the depth cues[9]. All other 3-D display devices lack one or more of the visual depth cues. For example, stereoscopic displays do not provide ocular accommodation and volume displays cannot provide occlusion. Image resolution and parallax resolution are also important considerations when displaying 3-D images. Although most 3-D displays fail to provide acceptable image and parallax resolutions, holography can produce images with virtually unlimited resolutions.

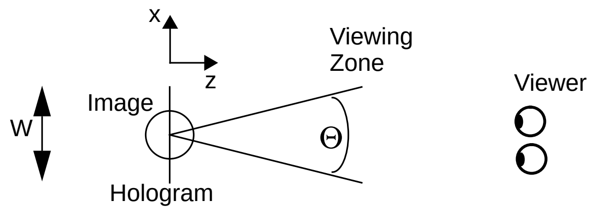
2.3 Holography

Optical holography[4] uses the physical phenomena of interference and diffraction to record and reconstruct a 3-D image. Holographic imaging became practical in the 1960’s with the advent of coherent monochromatic laser light. To produce a hologram, light is scattered from the object to be recorded. A photosensitive medium records the intensity pattern (fringe) that results when the light scattered from an object interferes with a spatially clean mutually coherent reference beam. The reference beam allows the medium to record both the magnitude and phase of the incident object wavefront, in essence recording variations in both the intensity and the direction of the light. The recording medium must have sufficient resolution to record spatial frequencies that are typically 1500 linepairs/millimeter (lp/mm) or more.

Diffraction-specific computation is developed for computation of all types of holograms. For simplicity of discussion, the focus in this dissertation is on off-axis transmission holograms possessing horizontal parallax only (HPO). It is possible to represent an HPO hologram with a vertically stacked array of one-dimensional (1-D) holographic lines[34,36]. Consider an HPO hologram made optically using a reference beam with a horizontal angle of incidence. Spatial frequencies are large in the horizontal direction (~1000 lp/mm) and increase with the reference beam angle. In the horizontal dimension, the sampling rate (or pitch) must be high to accurately represent the holographic information. During reconstruction of this hologram, diffraction occurs predominantly in the horizontal direction. Each hololine (a single horizontal line of the fringe) diffracts light to a single horizontal plane to form image points describing a horizontal slice of the image. Therefore, one hololine should contain contributions only from points that lie on a single horizontal slice of the object. Essentially, the 2-D holographic pattern representing an HPO 3-D image can be thought of as a stack of 1-D holograms or hololines.

A computer-generated hologram (CGH) represents a fringe pattern as an array of discrete samples. Given a fixed (HPO) hologram size, the sample count is simply the width times number of samples per unit length - the pitch, p. The relationship between (minimum) sampling pitch p and angle of diffraction Θ is

$$p = \frac{4}{\lambda} \sin\frac{\Theta}{2}. \quad \textbf{(1)}$$

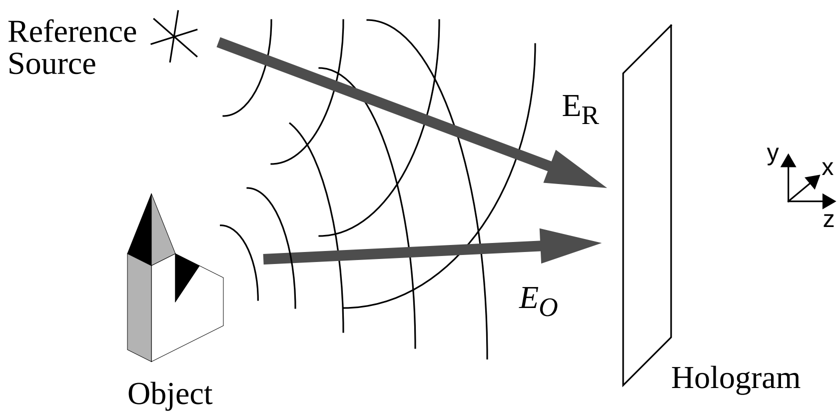


For a hologram located in the image plane (i.e., an *image-plane* hologram), the diffraction angle is equal to the viewing angle. (The image plane is the plane in the middle of the image volume.) Therefore, the image-plane requirement on sample count is determined by the size of the viewing zone. For a hololine of width W, the sample count becomes:

$$\text{number of samples} = Wp = \frac{4W}{\lambda} \sin\frac{\Theta}{2}. \quad \textbf{(2)}$$

Since an HPO CGH contains only a single vertical perspective (i.e., the viewing zone is vertically limited to a single location), spatial frequencies are low (~10 lp/mm) in the vertical dimension. The vertical image resolution is the number of hololines. Eliminating vertical parallax reduces CGH information content by at least a factor of 100 by reducing the vertical spatial frequency content from roughly 1000 to roughly 10 lp/mm.

Consider the typical optical holographic set-up in the following illustration. Light scattered from the object, E_O, interferes with reference light, E_R. (Note: In this analysis, optical wavefronts are represented by spatially varying complex time-harmonic electric field scalars. All wavefronts are assumed to be mutually coherent sources of monochromatic light. The units of an electric field amplitude are normalized so that the square of magnitude equals optical intensity. The polarizations are assumed to be identical and for simplicity are not specified.) The total time-harmonic electric field incident on the hologram is the interference of the light from the entire object and the reference light, $E_O + E_R$.



The total interference fringe intensity is I_T:

$$I_T = |E_O + E_R|^2 \tag{3}$$

This expression for total intensity expands to:

$$I_T = \underbrace{|E_O|^2}_{\text{Object Self-Interference}} + \underbrace{|E_R|^2}_{\text{Reference Bias}} + \underbrace{2Re\{E_O \cdot E_R^*\}}_{\text{Useful Fringes}} \tag{4}$$

The total intensity is a real physical light distribution comprising three components.

- *Object self-interference*: The first term, called object self-interference, is a spatially varying pattern that is generated when interference occurs between light scattered from different object points. During image reconstruction, this component of the holographic pattern is unnecessary and often produces unwanted image artifacts. In optical holography, a common solution is to spatially separate the object self-interference artifacts from the reconstructed image by increasing the reference beam angle to at least three times the angle subtended by the object. However, in computational holography, a large reference beam angle is a luxury that one does not have.

- *Reference bias*: The reference beam intensity represents a spatially nearly invariant (dc) bias that increases the value of the intensity throughout the hologram. This is also an unwanted term because it wastes available dynamic range in the holographic medium.
- *Useful fringes*: The third term is the interference between the object wavefront and the reference beam. This fringe pattern contains all of the holographic information that is necessary for image reconstruction.

To reconstruct an image, the recorded interference pattern modulates an illuminating beam of light. The modulated light diffracts (bends and focuses) and reconstructs a 3-D replica of the wavefront that was scattered from the object scene. Optical wavefront reconstruction makes the image appear to be physically present and tangible. The image possesses all of the depth cues exhibited by the original object, including continuous parallax (vertical and horizontal) and ocular accommodation. Both the image resolution and parallax resolution of an optical holographic image are virtually unlimited.

In a horizontal-parallax-only (HPO) image, the usable depth is limited by the tolerance of the human visual system to astigmatism. HPO holograms focus light only in the horizontal direction and not in the vertical direction. In most cases, the eye sees light vertically scattered from the image plane (the plane in the middle of the image volume). Light horizontally focuses to a range of depths within the image volume. The human visual system cannot tolerate astigmatism beyond its normal range of depth of focus[8] which has been empirically measured to be approximately 0.34 diopters (m^{-1}). Let the distance from the viewer to the image plane be D_V, and the distance to an imaged point of extreme depth be D_H. These parameters are related to the astigmatism constraint as

$$\left|\frac{1}{D_H} - \frac{1}{D_V}\right| = 0.34\ m^{-1}. \tag{5}$$

Typically, the viewing distance is 600 mm, making the maximum deviation of (D_V-D_H) between the image plane and the viewer be 100 mm. (In this thesis, the practical depth of points imaged with the MIT holovideo system was closer to 80 mm.)

Full-parallax image depth is often limited by vignetting, i.e., the windowing effect. The usable depth is a function of the location and lateral dimensions of the image plane and of the viewing angle. Typically, this limitation is about as strict as the astigmatism limit imposed on HPO images.

2.4 Computational Holography

To produce dynamic holographic images, researchers have computed holographic fringes and used them to modulate beams of light. Computational holography generally begins with a 3-D numerical description of the object or scene to be reproduced. Light is numerically scattered from the object scene and propagated to the plane of the hologram. The object wavefront is calculated and a reference beam wavefront added, imitating optical interference. The resulting total intensity - the fringe pattern - is used by a holographic display to produce the 3-D image. Such a display spatially modulates a beam of light with the fringe pattern, mimicking the reconstruction step in optical holography.

Traditionally, computational holography[21,22] was slow due to two fundamental properties of fringe patterns: (1) the enormous number of samples required to represent microscopic features, and (2) the computational complexity associated with the physical simulation of light propagation and interference. A typical full-parallax hologram 100 mm $\times$ 100 mm in size has a sample count (also called space-bandwidth product or SBWP or simply “bandwidth”) of over 100 gigasamples of information. A larger image requires a proportionally larger number of samples. Several techniques have been used to reduce information content to a manageable size. The elimination of vertical parallax[5] provided great savings in display complexity and computational requirements[19] without greatly compromising the overall display performance. Other less desirable sacrifices include reducing the size of the object scene or the size of the viewing zone.

Traditional holographic computation imitates the interference that occurs when a hologram is produced optically using coherent light. In this dissertation, these traditional methods are classified as “interference-based computation.” Early methods made use of the Fourier transform to calculate the phase and amplitude of the object wavefront at the plane of the hologram[22,23,24]. A plane of object points is propagated to the plane of the hologram using a Fourier transform to compute E_O. At each sample point in the hologram plane, this object light wavefront was added to a reference beam wavefront, and the magnitude squared became the desired fringe pattern. This approach was used to create planar images. Multiple image planes were separately Fourier transformed and combined to produce 3-D images[34,35]. This multiplicity makes the Fourier-transform approach slow and inefficient for computing 3-D images.

A more straightforward approach to the computation of holographic fringes resembled 3-D computer graphics ray-tracing. Light from a given point or element of an object contributed a complex wavefront at the hologram plane[26]. Each of these complex wavefronts was summed to calculate the total object wavefront, which was subsequently added to a reference wavefront. Black (non-scattering) regions of the image volume were ignored, enabling very rapid computations of simple object scenes. Again, for more complex images, computation was prohibitively slow.

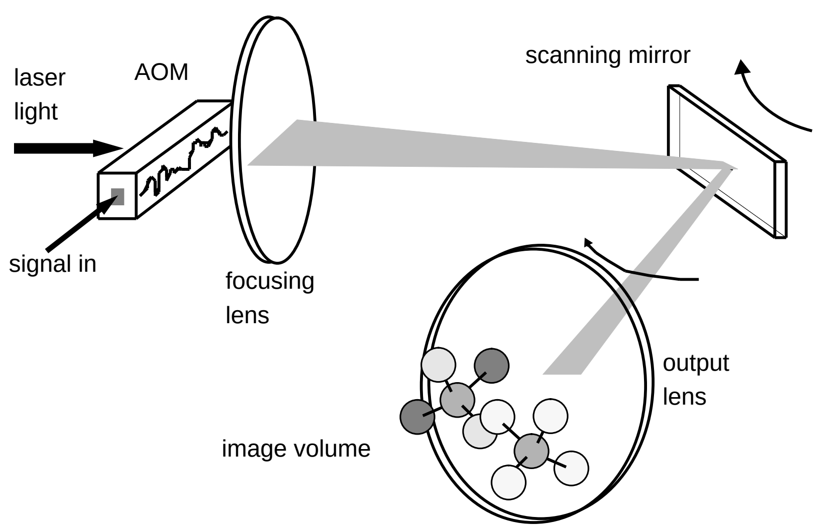
2.5 Holographic Displays

A holographic display is an electro-optical apparatus that modulates light with a holographic fringe pattern. In the earliest work in computational holography, the computed fringe pattern was recorded (permanently) in a piece of film[22]. The film provided the SBWP sufficient to represent the fringe pattern. In some cases, the film also provided grayscale (e.g., research involving the kinoform[29,32], and the ROACH[22]). Light passing through this film created a static holographic image. Some researchers used binary printers as output devices, and photographically reduced the printed binary patterns to achieve the required holographic resolutions. Most employed the detour-phase technique[25,27,28,30] for representing both the amplitude and phase of the computed object wavefront.

To create dynamic holographic images, a dynamic spatial light modulator (SLM) must be used. The SBWP of a holographic SLM must be as high as that of holographic film[74]. Ideally, a holographic SLM must display over 100 gigasamples. Current SLMs, however, can provide a maximum of 10 megasamples. Examples of SLMs include the flat-panel liquid-crystal display (LCD) and the magneto-optic SLM. These SLMs are capable of displaying a very small CGH pattern in real time. Early researchers employed a magneto-optic SLM with a SBWP of 16384 elements[61] or a LCD SLM with 10,000 elements[62] to produce tiny planar images. One group produced a small planar image using a deformable mirror device (DMD) with 16384 binary elements[63]. More recent work employed LCDs with higher pixel counts[64,65], but the images were still very small and essentially two-dimensional. (Note: To be considered a 3-D display, a real-time holographic display must provide at least binocular disparity, the one depth cue that all other 3-D displays provide. Therefore, the viewing zone must be a minimum of about 100 mm to allow for both eyes to see the diffracted light. Otherwise, binocular disparity cannot be supported, and the display is only 2-D.)

An ideal holographic SLM does not yet exist, but time-multiplexing of a very fast SLM provides a suitable substitute[50,51,57]. The display system that we used in this thesis research was the second generation holovideo display developed by the Spatial Imaging Group at the MIT Media Laboratory[58]. This display used the combination of an acousto-optic modulator (AOM) and a series of lenses and scanning mirrors to assemble a softball-sized 3-D holographic image at video frame rates. This time-multiplexed SLM approach is sometimes called the ‘Scophony geometry” after the early contender for television displays[82,83]. A partial schematic is shown in the figure below. A general description follows, and a full description can be found in the references[57,58]. It is important to note that the diffraction-specific fringe computation described in this dissertation is not limited to use with the MIT display. By incorporating the proper physical parameters, wavelengths and sample size, a hologram generated using this method can be viewed using other holographic displays.

MIT Holovideo Display: Partial Schematic



This figure shows a partial schematic diagram (top view) of the MIT holovideo display. The scanning mirror angularly multiplexes the image of the modulated light. A vertical scanning mirror (not shown) positions each hololine vertically. Electronic control circuits synchronize the scanners to the incoming holographic signal.

In the MIT holovideo display, as each line (*hololine*) of the fringe pattern was read out of a high-speed framebuffer, it passed through a radio-frequency (RF) signal processing circuit and into one channel of the AOM. This display used 18 parallel channels from an 18-channel framebuffer and feeding into an 18-channel AOM that modulated 18 separate beams of red laser light. At any instant, as 18 lines of the holographic pattern traversed the aperture of the AOM in the form of acoustic waves (at a speed of 617 m/s), a portion equal to roughly 1000 samples (in each channel) modulated the phase of the wavefront of laser light that passed through each channel of the AOM. Two lenses imaged the diffracted light at a plane in front of the viewer. By reflecting the light off of a synchronized horizontally scanning mirror, the apparent motion of the holographic pattern was cancelled. The scanning mirror also acted to angularly multiplex the image of the acoustic wave. It extended the apparent width of the imaged holographic pattern to 256 Ksamples, with each sample representing a physical spacing of 0.6 μm.

The viewer saw a real 3-D image located just in front of the output lens of the system. The image occupied a volume that was 150 mm wide, 75 mm high and 160 mm deep. The size of the viewing zone - i.e., the range of eye locations from which the viewer can see the image - was 30 degrees horizontal. The viewer experienced the depth cue of horizontal motion parallax. This was a horizontal-parallax-only (HPO) image. Vertical parallax was sacrificed to simplify the display. (This restriction does not limit the display's usefulness in most applications.) Because the holographic image possessed no vertical parallax, there was no need for diffraction in the vertical dimension. The vertical resolution of 144 lines over 75 mm was equivalent to that of a common 2-D display.

2.6 Bandwidth Compression in Holography

Holographic fringe patterns contain more information than can be utilized by the human visual system[73,74,76]. For example, consider a 1-Msample hololine with one byte per sample. For a typical display system, this hololine may represent an array of up to 1000 points, with ~20 bytes each (x, y, z, brightness, and directionally dependent information), for a total of ~20 KB of actual usable information. Most of the bandwidth is unused: only 20 KB of visual information is communicated through a hologram with a 1-MB channel capacity.

Several researchers have attempted optically to reduce bandwidth in holographic imaging. Haines and Brumm[75,78,81] attempted to use a hologram of reduced size to generate an image with that was not reduced in size or in viewing zone. During optical recording, a dispersion (scattering) plate was positioned between the hologram plane and the object to be imaged. A reduced-size hologram was recorded and repositioned so that diffracted light passed through the dispersion plate. Although the hologram was of reduced size, the viewer saw an image that was as large as an image produced by a full-size hologram in a standard holographic imaging system. However, image quality suffered. Either image resolution or signal-to-noise ratio (SNR) was reduced. The choice of dispersion plate determined which of these degradations was traded off against holographic bandwidth. Bandwidth reduction was as much as 60 times in each lateral dimension. As bandwidth reduction increased, the increase in artifacts overwhelmed the image. Hildebrand[80] generalized the dispersion-plate approach by including time-varying scattering functions. The advantage was increased flexibility in alignment and a decrease (in some cases) in image artifacts. In all of these dispersion-plate methods, the intermediate scattering plate was used as an angular multiplier. The optically produced hologram records light propagating at a continuous range of directions. However, the human visual system does not need such precision. Essentially, the dispersion plate encodes angular information during optical recording by dispersing this information throughout the entire hologram. A reduced-size portion of this hologram contains information about light traveling in all directions. (This is not generally the case in standard optical holography.) During reconstruction, light passing in the opposite direction through the same dispersion plate is “decoded” to from a larger image. Light from the small hologram appears to be emerging from a wider range of directions because of the angular multiplication effect of the dispersion plate.

Burckhardt and Enloe[76,79] reduced the information recorded in a hologram by exposing only an array of small regularly spaced areas. As the proportion of the hologram that was exposed decreased, the amount of information recorded in the hologram decreased. This work was equivalent to spatially sampling the hologram. The reconstructed image had an annoying “screen” artifact; the image appeared as if on the other side of a picket fence. Techniques to eliminate this artifact reduced the resolution of the image. Good images were reconstructed with information reduction factors of 6 in each lateral dimension.

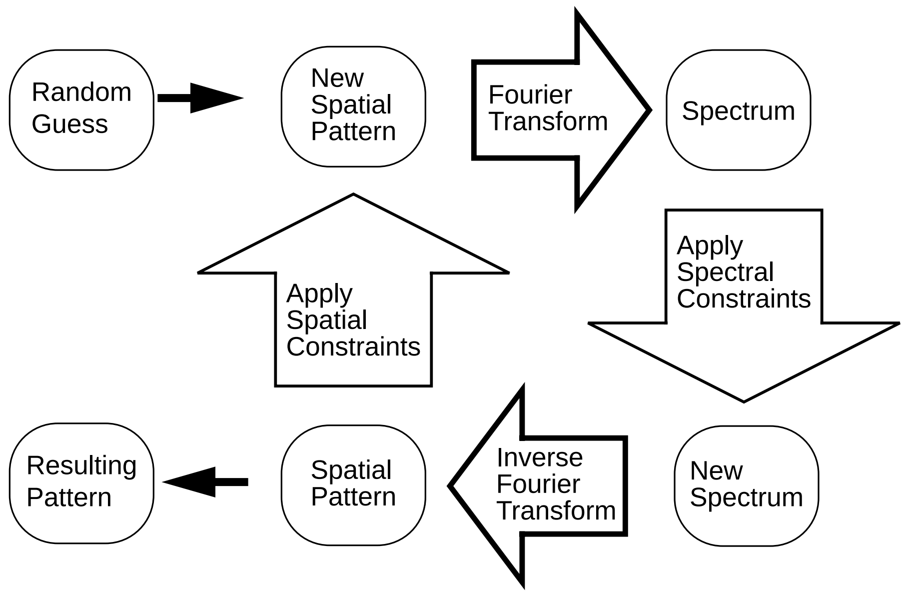
Lin[77] also used a spatial subsampling to reduce holographic bandwidth. In this work, a Fourier transform the hologram was recorded by placing the hologram plate at the focus of a lens that essentially performed a Fourier transform of the image light. Spatially sampling a Fourier transform hologram was equivalent to spectrally sampling the image light. Similar to the work by Burckhardt and Enloe, the subsampled hologram was exposed through a mask (an array of regularly spaced small apertures). Multiple exposures, each preceded by small lateral translations, formed a mosaic hologram in which a given region contained replicas of the sample recorded at that location in the hologram plane. In this way, the spectrum was subsampled, and the samples replicated to recover an approximation of the original hologram. (This work was inspirational for the computational holographic encoding schemes developed later in this dissertation.) A bandwidth reduction of 10 in each lateral dimension was achieved. (More information reduction was achieved through eliminating vertical parallax.) As in all of these optical holographic bandwidth reduction experiments, image fidelity suffered due to decreased image resolution and the presence of artifacts, e.g., graininess and moire-like stripes.

All of these experiments in optical holographic information reduction exploited the redundancy inherent in holographic fringes. Generally, these researchers subsampled (spatially and/or spectrally) to reduce information content. Image quality suffered due primarily to the reconstruction process. Dispersion plates caused graininess and noise. Periodic replication caused moire-like stripes. All methods decreased image resolution, though this was acceptable within a certain range. What these methods lacked was the ability to directly manipulate the recorded holographic fringe information to reduce reconstruction artifacts. However, if the fringes had been computed, direct manipulation would have been possible. For example, Wyrowski, Hauck, and Bryngdahl[45,46] used phase manipulation to reduce speckle-like artifacts in computer-generated hologram produced by replicating smaller subholograms. The advantage of computational holography is the ability - indeed, the necessity - to specify more precisely the nature of the fringes. Therefore, this thesis translates some of these optical information-reduction concepts into computational holography, where they are more useful and more realizable.

2.7 Iterative Hologram Computation Methods

Several numerical computational algorithms were developed for the purposes of phase retrieval[44]. These methods were capable of deriving an unknown phase pattern given the amplitude of a signal and its spectrum[42]. Certain iterative methods were applied to the computation of phase patterns (diffusers[41,49], etc.) for optical systems and in some cases actual holographic fringes[43,48]. These computation methods iteratively applied spatial and spectral constraints while sequentially transforming the signal from the spatial domain to the Fourier (spectral) domain. The constraints in either the spatial or spectral domains were imposed upon the amplitude and the phase of the pattern. The algorithm is as follows:

1. Begin with a random pattern.
2. Transform into the spatial frequency domain.
3. Apply the (spectral) constraints.
4. Inverse transform back to the spatial domain.
5. Apply the (spatial) constraints.
6. Iterate starting at step 2, until pattern converges.



After several iterations, the process converges to an acceptable solution to the constraints.

In the case of holographic fringes, the diffraction integral describing the propagation of light is essentially a Fourier transform integral[12]. Therefore, holographic fringes were computed by using the (2-D) object as a spectral amplitude constraint. The spectral phase was left unconstrained. Spatially, the pattern was left unconstrained, or in some cases was constrained to have a uniform amplitude[48]. After a sufficient number of iterations (~100), the spatial pattern converged to a useful fringe pattern. These holograms were Fourier transform holograms, i.e., the square of the amplitude of the fringe spectrum (the spectral energy) was the desired image.

The method of iterative constraints is used in this thesis to generate synthetic "basis fringes" for diffraction-specific computation. The adaptation of this algorithm is discussed in Appendix C, "Computation of Synthetic Basis Fringes" on page 159.