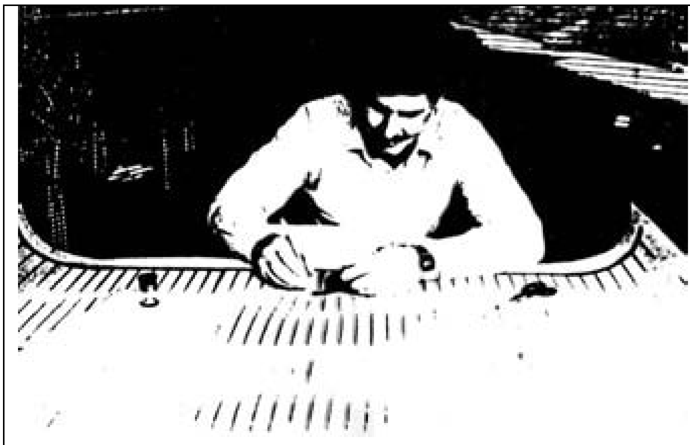
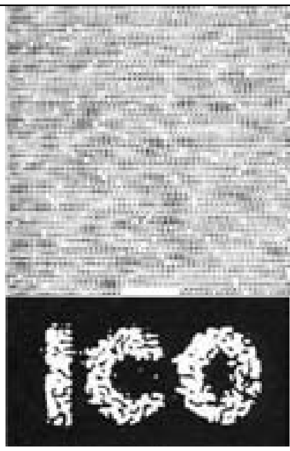
Lect. 4. Computer Generated Holograms (CGH): principles

4.1. Introduction

The origins of digital holography date back to 1960-70th ([1-4]). First computer generated holograms were binary holograms invented by A. Lohmann ([1-2]). They were printed on a computer line printer, then they were optically reduced and reconstructed in optical set ups using coherent laser illumination. Fig. 4.1 shows one of the first computer-generated hologram and its reconstruction. Then the use of more sophisticated devices capable of recording computer generated grey scale images was suggested for recording computer generated holograms and computer generated holograms that can reconstruction good quality grey scale images were produced ([3,4]). Fig. 4.2. represents an example of grey-scale computer-generated hologram and an image reconstructed from a grey scale hologram. The grey scale holograms were recorded using electro-mechanical recording device (a modified faximile machine) with resolution 200 μm and then optically reduced with reduction factor 20 to final resolution 10 μm for optical reconstruction with a laser beam.



Byron Brown completing one of the first computer-generated holograms (adopted from Ref. [5])



An early CGH and its reconstruction (adopted from Ref. [5])

Figure 4.1. One of the first computer-generated holograms and its reconstruction

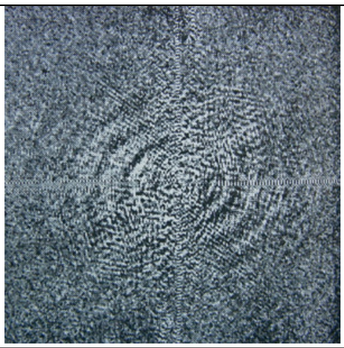


Figure 4.2. One of the first gray scale computer-generated hologram (left) and an image reconstructed from a gray scale computer-generated hologram (right, adopted from [4])

In Sect. 4.2 mathematical models for synthesis of computer-generated holograms are formulated for CGH intended for information display. In Sect. 4.3, the problem of recording CGHs on physical holographic media is discussed and a cllassification of spatial light modulators for recording CGHs is provided. Methods of coding of computer-generated hologram for recording on physical media will be given in Lect. 5.

4.2. Mathematical models

Basic stages in the synthesis of computer generated holograms are (Fig. 4.3):

(i) Formulating mathematical models of the object and of the usage of the hologram;

(ii) Computing the ***mathematical hologram***, an array of complex numbers that represent amplitudes and phases of hologram samples in the hologram plane;

(iii) Encoding samples of the mathematical hologram for recording them on the physical medium. At this stage, which we refer as to ***hologram encoding***, mathematical holograms are converted into an arrays of numbers that control optical properties of the physical recording medium used for recording the hologram.

(iv) Fabrication of the computer generated hologram.

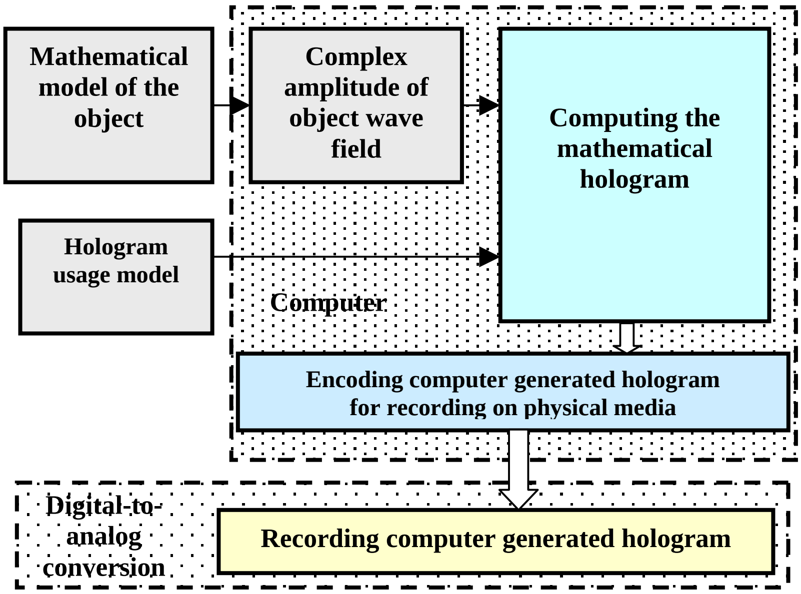


Figure 4.3. Basic stages in synthesis of computer generated holograms

Computer-generated holograms can be used as spatial filters for optical information processing (see Lect. 7), as computer generated diffractive optical elements, such as beam forming elements (for instance, for laser tweezers), laser focusers, deflectors, beam splitters and multiplicators, and for information display. Mathematical models of the usage of holograms to be generated describe geometry of wave propagation from objects to the hologram plane and specify criteria for evaluation hologram performance.

Mathematical model of the object is intended to specify the amplitude and phase distribution of the object wave front. Three types of the models can be distinguished:

- analytical models,
- “geometrical”, or “vector graphics” models that represent objects as compositions of elementary diffracting elements such as point scatters, segments of 2-D or 3-D curves, slits, etc, such as, for instance, polygonal mesh or wire-frame models in 3D computer graphics
- “raster graphics”, or “bitmap” models that represent objects as 2-D or 3-D arrays of points (object samples, pixels or voxels).

Analytical models are used mainly for generating diffractive optical elements. With analytically specified desired distribution of the object wave front synthetic diffractive optical element is computed using methods of analytical or numerical integration with appropriate oversampling to avoid transformation sampling artifacts and then sampled to generate their sampled version for recording on physical media.

When objects are represented using “geometrical models”, mathematical hologram is computed as a superposition of elemental holograms that correspond to the elements of the model (point scatters, edges of wire frame models, faces of polygonal mesh models, etc.) that can frequently can be pre computed and stored in a look up table. This allows to substantially reduce the computational complexity of calculation of mathematical holograms. Graphic models and raster models are illustrated in Fig. 4.4.

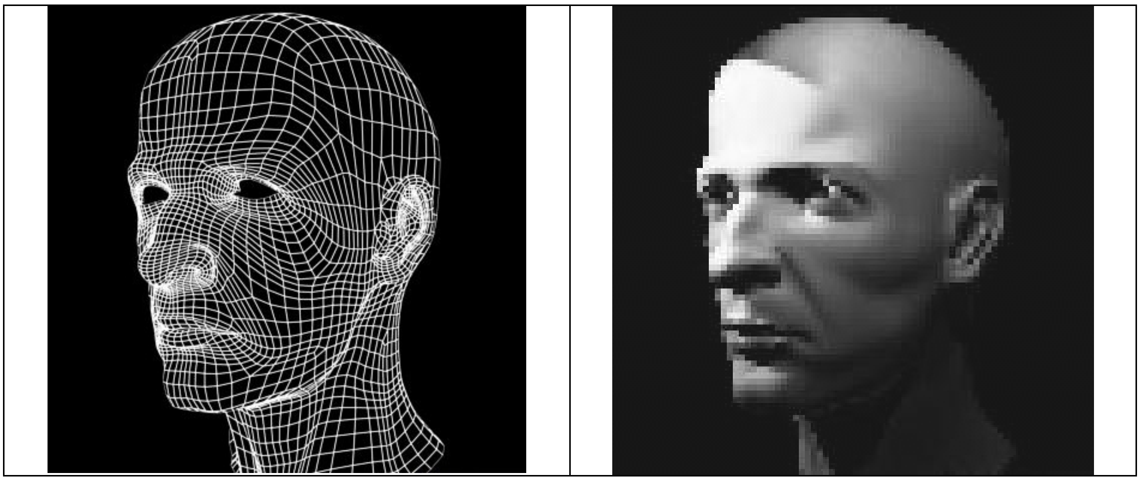


Figure 4.4. Human head: “vector graphic” model (left) and “raster graphic” model (right)

“Raster graphic” models are the most general and better suited to object data provided by physical sensors, such as image and 3D scanners. In what follows, we will describe mathematical models for computer generated display holograms of objects specified by “raster graphic” models.

Consider a schematic diagram for the visual observation of objects shown in Fig.4.5. The observer's position with respect to the observed object is defined by the observation surface where the observer's eyes are situated. The set of observation positions is defined by the object observation angle. In order to allow the observer to see the object at the given observation angle, it suffices to reproduce, by means of a hologram, the distribution of intensities and phases of the light waves scattered by the object on the observation surface.

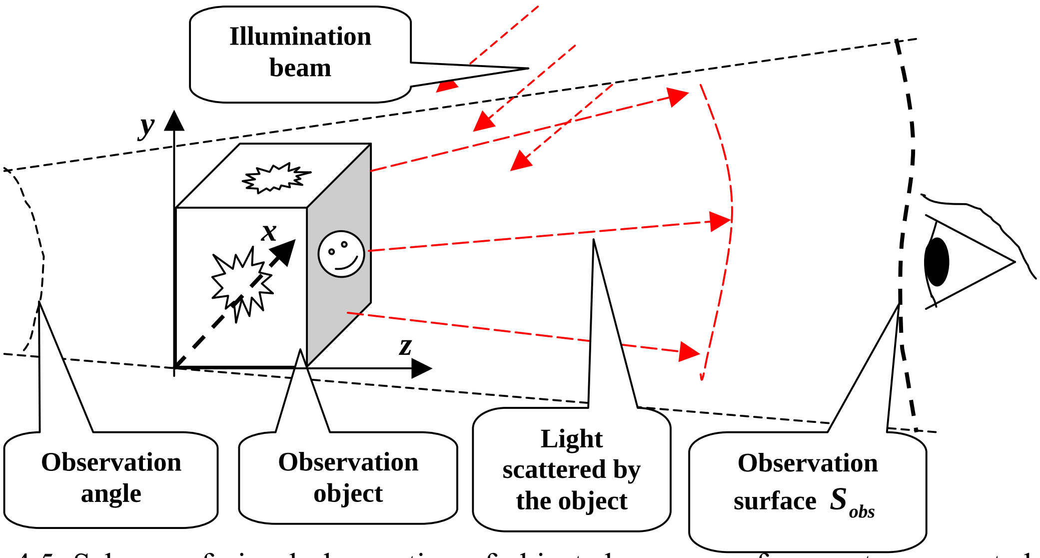


Figure 4.5. Scheme of visual observation of objects by means of computer generated holograms

We will assume monochromatic illumination of objects, which enables one to describe light wave transformations in terms of the wave amplitude and phase. The case of polychromatic illumination may be treated as superposition of monochromatic illumination with different wavelengths. For instance, generating color holograms assumes generating three color separated (red-greed-blue) holograms.

In monochromatic illumination, object characteristics that define its ability to reflect, transmit or scatter incident radiation may be described for our purposes by a radiation (transmission) reflection factor with respect to the light complex amplitude $\boldsymbol{o}(x,y,z)$ or intensity $\boldsymbol{O}(x,y,z)=|\boldsymbol{o}(x,y,z)|^2$ as functions of the object's surface coordinates (x,y,z). Then the intensity of the scattered wave $\boldsymbol{I}_o(x,y,z)$ and its complex amplitude $\boldsymbol{A}_o(x,y,z)$ at the point (x,y,z) are related to the intensity $\boldsymbol{I}(x,y,z)$ and complex amplitude $\boldsymbol{A}(x,y,z)$ of the incident wave as follows:

$$\begin{aligned} \boldsymbol{I}_o(x,y,z) &= \boldsymbol{O}(x,y,z)\boldsymbol{I}(x,y,z) \\ \boldsymbol{A}_o(x,y,z) &= \boldsymbol{o}(x,y,z)\boldsymbol{A}(x,y,z) \end{aligned} . \quad (4.2.1)$$

The amplitude reflection factor $\boldsymbol{o}(x,y,z)$ may be regarded as a complex function of the spatial coordinates represented as

$$\boldsymbol{o}(x,y,z) = |\boldsymbol{o}(x,y,z)| \exp[i\theta_o(x,y,z)]. \quad (4.2.2)$$

Its modulus $|\boldsymbol{o}(x,y,z)|$ and phase $\theta_o(x,y,z)$ show how the modulus $|\boldsymbol{A}(x,y,z)|$ and the phase $\omega(x,y,z)$ of the complex amplitude $|\boldsymbol{A}(x,y,z)|\exp[i\omega(x,y,z)]$ of incident light are changed after reflection by the object surface or transmission through the object at the point (x,y,z):

$$\begin{aligned}\left|A_{obj}(x,y,z)\right| &= \left|A(x,y,z)\right|\left|o(x,y,z)\right| \\ \omega_{obj}(x,y,z) &= \omega(x,y,z)+\theta_{obj}(x,y,z)\end{aligned}, \quad (4.2.3)$$

where $\left|A_{obj}(x,y,z)\right|$ and $\omega_{obj}(x,y,z)$ are, respectively, the modulus and the phase of the object wave front.

The relation between the complex amplitude $\Gamma(\xi,\eta,\zeta)$ of the light wave field over an arbitrary surface defined by its coordinates (ξ,η,ζ) and the complex amplitude $A_o(x,y,z)$ of the wave front at the object surface can by described by a wave propagation integral over the object surface or volume S_{obj}

$$\Gamma(\xi,\eta,\zeta)=\iiint\limits_{S_{obj}} A_{obj}(x,y,z)T(x,y,z;\xi,\eta,\zeta)dxdydz \ , \quad (4.2.4)$$

The kernel $T(x,y,z;\xi,\eta,\zeta)$ of the wave propagation integral depends on the spatial disposition of the object and the observation surface. Reconstruction of the object can be described by a back propagation integral over the observation surface S_{obs}:

$$\tilde{A}_{obj}(x,y,z)=\iiint\limits_{S_{obs}} \Gamma(\xi,\eta,\zeta)\tilde{T}(\xi,\eta,\zeta;x,y,z)d\xi d\eta d\zeta, \quad (4.2.5)$$

where $\tilde{A}_{obj}(x,y,z)$ is complex amplitude of the object as it is seen from the observation surface and $\tilde{T}(\xi,\eta,\zeta;x,y,z)$ is a kernel reciprocal to $T(x,y,z;\xi,\eta,\zeta)$.

Thus, the hologram synthesis requires computation of $\Gamma(\xi,\eta,\zeta)$ through $A_{obj}(x,y,z)$, which is to be defined by the object description and illumination conditions, and subsequent recording the computation results on a physical medium in a form that enables interaction with radiation for visualizing or reconstructing $\tilde{A}_o(x,y,z)$ according to Eq. (4.2.5).

3-D integration required by Eq. 4.2.4 can be reduced to much more simple 2-D integration by taking into consideration the following natural limitations of visual observation:

- The pupil of the observer's eye is usually much smaller than the distance to the observation surface.
- The area of the observation surface is large compared to the inter-pupil distance.
- The depth of objects situated at a distance convenient for visual observation usually is small compared to the distance to the observation surface.

In view of these limitations, one can consider the observation surface as consisted of relatively small sub-areas approximated by planes. Additionally one can, using the laws of geometrical optics, replace the wave amplitude and phase distributions over the object surface by those of the wave on a plane tangent to the object (or sufficiently close to it, so as to make diffraction effects negligible) and parallel to the given plane sub-area of the observation surface. These approximations are illustrated in Fig. 4.6. In this way the problem of hologram synthesis over the whole observation surface is reduced to the synthesis of fragmentary holograms for the plane areas of this surface, with the complete hologram being composed as a mosaic of fragmentary ones. In such assumptions, one then obtains from Eq.(4.2.4) the following equations for fragmentary holograms:

$$\overline{\Gamma}(\xi,\eta)=\iint_{S_{obj}}\overline{A}_o(x,y)T_Z(x,y;\xi,\eta)dxdy\ , \qquad (4.2.6)$$

where $\overline{A}_o(x,y)=\left|\overline{A}_o(x,y)\right|\exp\left|i\overline{\omega}_o(x,y)\right|$ is a complex function resulting from recalculation of the amplitude and phase of the field reflected or transmitted by the object onto the plane (x,y) tangent to it and parallel to the observation plane (ξ,η) and $T_D(x,y;\xi,\eta)$ is a reduced 2-D transformation kernel for the distance between these planes Z .

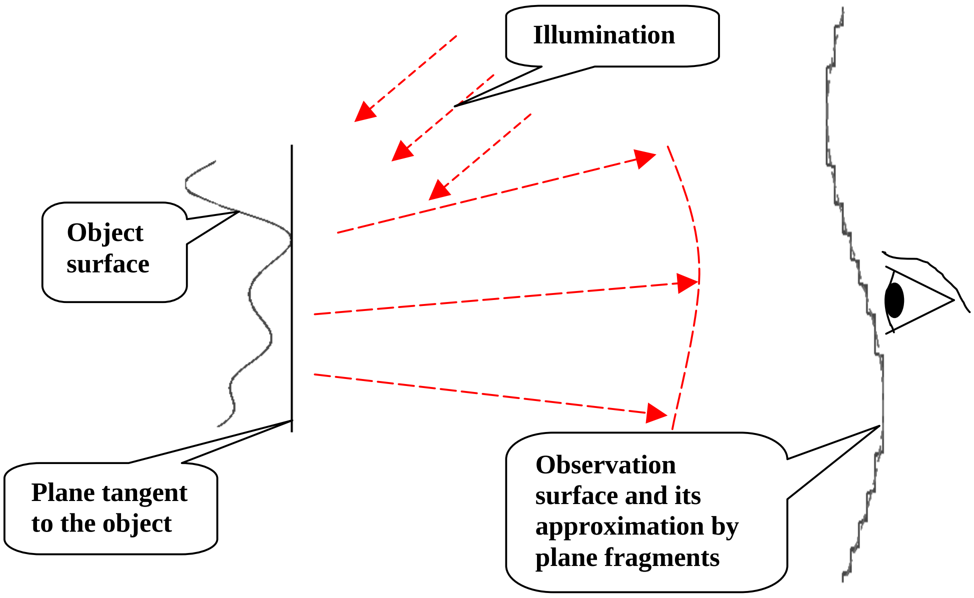


Figure 4.6. Approximations of the scene and observation geometry that lead to Eq. 4.2.6.

If the geometrical dimensions of the observation object and of the observation surface fragment are small compared to the distance Z from the object to the observation plane, the integral relationship of Eq. 4.2.6 is, as it was shown in Lect. 2, reduced to the integral Fresnel transform:

$$\overline{\Gamma}_{Fr}(\xi,\eta)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\overline{A}_o(x,y)\exp\left\{i\pi\left[(x-\xi)^2+(y-\eta)^2\right]/\lambda Z\right\}dxdy, \qquad (4.2.7)$$

where λ is the illumination wave length. Discrete representation of the integral Fresnel transform suitable for its numerical computing is discussed in Lect. 3. The holograms synthesized by means of this relationship will be referred to as synthesized ***Fresnel holograms***.

Further simplification is possible if

$$\pi\left(x^2+y^2\right)/\lambda Z \ll 1 \qquad (4.2.8)$$

in which case the integral of Eq.(4.2.8) can be approximated as

$$\overline{\Gamma}_{Fr}(\xi,\eta) \cong \exp\left[i\pi\left(\xi^2+\eta^2\right)\right]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\overline{A}_o(x,y)\exp\left[i2\pi(x\xi+y\eta)/\lambda Z\right]dxdy \qquad (4.2.9)$$

through the integral Fourier transform:

$$\overline{\Gamma}_F(\xi,\eta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\overline{A}_o(x,y)\exp\left[i2\pi(x\xi+y\eta)/\lambda Z\right]dxdy \ . \qquad (4.2.10)$$

The holograms synthesized according to Eq. 4.2.10 by means of Fourier transformation will be referred to as synthesized ***Fourier holograms***. Fourier and Fresnel holograms can be computed through Discrete Fourier transforms discussed in Lect. 2.

Schemes of using computer generated display Fresnel and Fourier holograms for visual observation of virtual objects are shown in Fig. 4.7, a) and b), respectively.

From the standpoint of object wave front reconstruction, Fresnel holograms differ from Fourier holograms in that they have focusing properties and are capable of reproducing the finite distance from the observation surface to the object. For reconstruction of Fourier holograms, spherical reconstruction light beam is needed and reconstructed image is observed in the plane of the reconstructing beam point source.

Mathematically, reconstructing an object by Fresnel and Fourier holograms is described by the inverse Fresnel and Fourier transformations, respectively. When holograms are observed visually, these transformations are performed by the optical system of the eye.

Having passed from the spatial 3D-problem to the problem of treating its 2D approximation, we, strictly speaking, lose the possibility of taking into account the precise effect of object depth relief on the wave front. Even Fresnel holograms use only a single distance value from the object to the observation plane rather than object relief depth. Nevertheless, it is still possible to synthesize holograms, which are capable of reconstructing 3-D images for visual observation and retain the most important property of holographic visualization - naturalness of object visual observation. At least two options exist for this: composite computer generated stereo- and macro-holograms and programmed diffuser holograms.

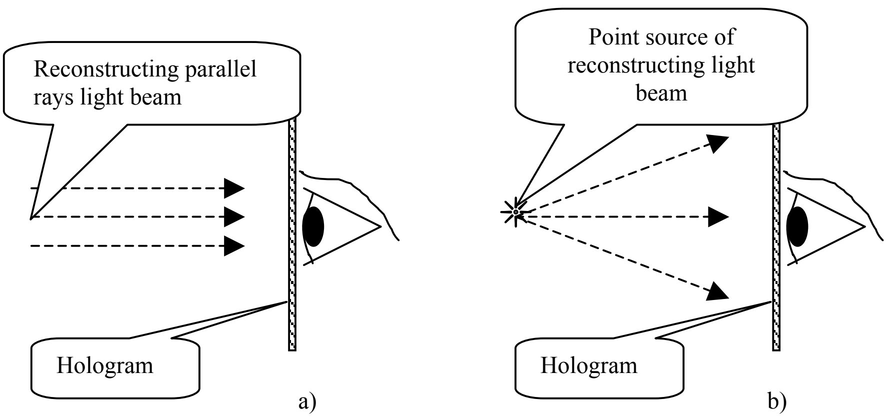


Figure 4.7. Schematic diagram of visual observation of virtual objects using computer generated Fresnel (a) and Fourier (b) holograms

Stereo holograms are holograms synthesized for viewing by left and right eye and to reconstruct the corresponding stereo views of the object. Composite stereo holograms are composed of multiple holograms that reproduce different aspects of the object as observed from different positions. Macro-holograms are large composite stereo-holograms that cover wide observation angle and can be used as a wide observation window.

Programmed diffuser holograms are computer generated Fourier holograms that imitate properties of diffuse surfaces to scatter irradiation non-uniformly in the space and in this way provide visual clue about the object surface shape. Synthesis of stereo holograms and programmed diffuser hologram is discussed in Lect. 8.

4.3. Digital-to-analog conversion problem of recording CGHs and spatial light modulators for recording CGHs

Optical media for recording computer generated holograms may be classified into three categories: ***amplitude-only media***, ***phase-only media***, and ***combined amplitude/phase media*** (Fig. 4.8.).

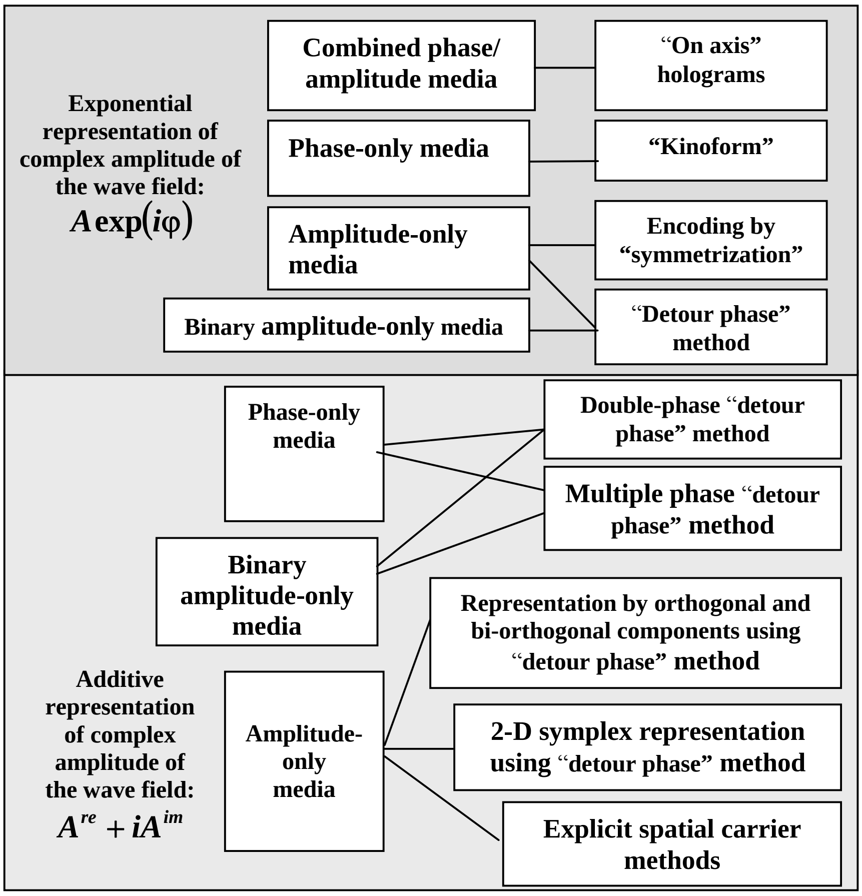


Figure 4.8. Classification of methods for recording computer generated holograms

In amplitude-only media, the controlled optical parameter is the light intensity transmission or reflection factor. This is the most common and available class, whose typical representatives are the standard silver galid photographic emulsions used in photography and optical holography ([6]).

In phase-only media, the optical thickness of the medium can be controllable, for example, by varying the medium refractive index, or physical thickness, or both. Phase-only media include thermoplastic materials, photo resists, bleached photographic materials, media based on photo polymers, etc. Recently, micro-lens array and micro-mirror array technology has emerged ([7,8]) that offer new opportunities for recording computer generated holograms.

Combined media allow independent control of both the light intensity transmission factor and of the optical thickness. Currently, these are photographic materials with two or more layers sensitive to radiation of different wavelengths. This permits the user to control of the transparency of certain layers and the optical thickness of others by exposing each layer to its wavelength independently.

Special digitally controlled hologram recording devices are required for modulating optical parameters of these media according to the mathematical hologram. No such special purpose devices are available as yet, and computer printer/plotters, display devices and other improvised means such as photo- and e-beam (electron) lithography are used instead. The distinctive feature of many of such devices is that they perform only binary or two level modulation of the medium's optical parameters. Amplitude-only and phase-only media whose controlled optical parameters may assume only two values will be referred to as ***binary media***. The use of amplitude and phase media in the binary mode is quite inefficient in terms of the medium information capacity because the possibility of writing information on them is defined only by their spatial degrees of freedom (their resolution power), whereas in amplitude and phase media, in principle, the degrees of freedom related to a transmission (reflection, refraction) factor may be used as well. The major advantage of the binary mode is simplicity of media exposure and copying recorded holograms.

Yet another classification feature for optical media for recording computer generated display holograms is their capability to be used as rewriteable devices.

The most important numerical characteristics of recorders are the total number of cells, which may be exposed, and their sampling interval, that is, the distances $\Delta\xi$ and $\Delta\eta$ between neighboring, separately and independently exposed resolution cells. The sampling interval defines the angular dimensions (θ_x, θ_y) of the reconstructed image according to the known diffraction relationships:

$$\theta_x = 2\pi\lambda/\Delta\xi\,;\; \theta_y = 2\pi\lambda/\Delta\eta\,, \tag{4.2.11}$$

For instance, in order to make the image's angular dimensions about 20^o or more with a reconstruction light wavelength of about 0.5 μm, $\Delta\xi$ and $\Delta\eta$ should be less then 10 μm. The existing spatial light modulators and printing devices that can be used for recording computer generated holograms have a sampling step of about 1 through 10 μm, and the total number of resolution cells, by the order of magnitude, from $10^3 \times 10^3$ to $10^4 \times 10^4$. As it was mentioned, at the early stages of digital holography, holograms were recorded by means of low resolution computer printer/plotters and image recording devices. Recorded hologram hard copies were then photographically reduced in order to achieve acceptable values of the sampling interval.

In conclusion, Table 1 presents, as an example, a short specification of one of modern spatial light modulators “Reflective Phase only display panel (LCoS - Liquid Crystal on Silicon)” produced by HOLOEYE Photonics AG (www.holoeye.com) (Fig. 4.9). Fig. 4.10 shows amplitude and phase modulation properties of this SLM.

Tab. 1: Quick reference data for HDTV phase-only LCOS

Active area dimensions	15.36 x 8.64 (0.7” diagonal) mm
Screen aspect ratio	16 (H) : 9 (V)
Display resolution	1920 (H) x 1080 (V) pixels
Pixel pitch	8.0 μm
Pixel configuration	Orthogonal
Phase Levels	256 (8 bit)
Optical Mode	Reflective
Liquid crystal type	Nematic (ECB-mode)

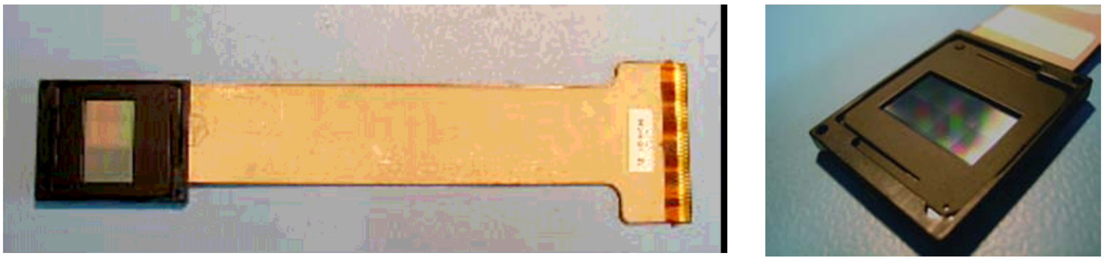


Figure 4.9. Reflective Phase only (LCoS - Liquid Crystal on Silicon)” display panel (HOLOEYE Photonics AG)

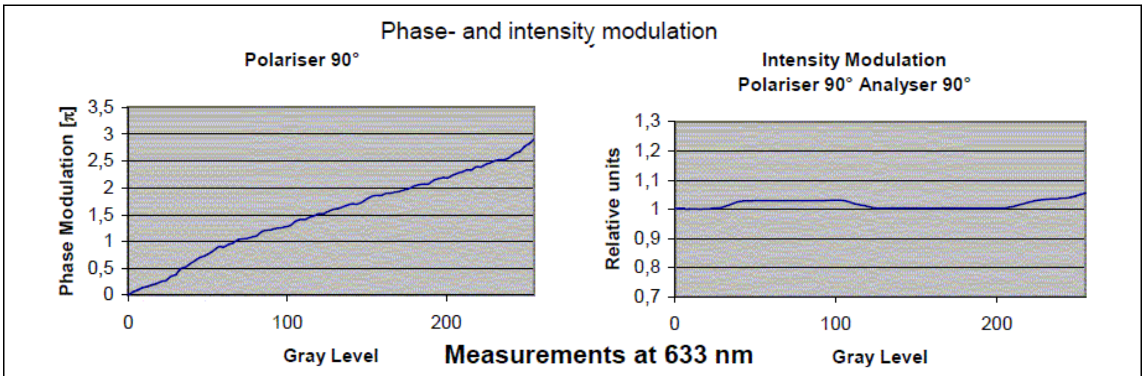


Figure 4.10. Phase and intensity modulation functions of the Holoeye display panel

References

1. B. R. Brown, A. Lohmann, Complex spatial filtering with binary masks, *Appl. Optics*, **5**, No. 6, 967-969 (1966)
2. B. R. Brown, A. Lohmann, Computer generated binary holograms, IMB J. Res. Dev., v. 13, No. 2, 160-168 (1969
3. M.A. Kronrod, N.S. Merzlyakov, L.P. Yaroslavsky, Computer Synthesis of Transparency Holograms, Soviet Physics-Technical Physics, v. 13, 1972, p. 414 - 418.
4. L. Yaroslavskii, N. Merzlyakov, Methods of Digital Holography, Consultance Bureau, N.Y., 1980
5. A. Lohmann, A Prehistory of Computer Generated Holography, OPN, February, 2008
6. H. I. Bjelkhagen, Silver Halid Recording Materials, Springer Verlag, Heidelberg, 1993
7. R.S. Nesbitt, S. L. Smith, R.A. Molnar, S. A. Benton, Holographic recording using a digital micromirror device, Conf. on Practical Holography, Proc. SPIE, 3637, 12-20, 1999
8. T. Kreis, P. Aswendt, R. Hoeffling, Holographic reconstruction using a digital micromirror device, Optical Engineering, v. 40 (6), pp. 926-933, June 2001