8.1. Computer generated display holograms based 3D communication paradigm

The idea of producing artificial optical elements capable when illuminated of creating virtual images can be traced back to middle-ages Chinese mirrors (Fig. 8.1). The inventions in holography by E. Leith and Yu. Denisyuk ([1, 2], Fig. 8.2 and 8.3) were primarily motivated by the desire to create efficient means for visualizing 3D images.

Fig. 8.1. Chinese mirrors – middle ages prototype of computer-generated holography

Fig. 8.2. Emmett Leith and Juris Upatnieks at the University of Michigan producing their "Train and Bird " hologram, the first hologram ever made with a laser using the off-axis technique.

Fig. 8.3. Yu. N. Denisyuk with his holographic portrait
There are no doubts that holographic imaging is an ultimate solution for 3-D visualization. This is the only method that is capable of reproducing, in the most natural viewing conditions, 3-D images that have all the visual properties of the original objects including full parallax, and are visually separated from the display device. 3-D visual communication and display can be achieved through generating, at the viewer side, of holograms out of data that contain all relevant information regarding the scene to be viewed. Digital computers are ideal means for converting data on 3-D scenes into optical holograms for visual perception ([3, 4]).

The core of the 3D digital holographic visual communication paradigm is the understanding that, for generating synthetic holograms at the viewer side, one does not need to produce, at the scene side, the hologram of the scene and to transmit it to the viewer’s site. Neither does one need to necessarily imitate, at the viewer site, the full optical holograms of the scene. What one does need is to collect, at the scene side, a set of data that will be sufficient to generate, at the viewer site, a synthetic hologram of the scene for viewing.

What features of optical holograms computer-generated display holograms for 3D visualization should have? The major requirement to computer-generated display holograms is that they should provide natural viewing conditions for the human visual system and, in particular, separation of reconstructed images from the display device.

A crucial issue in transmitting data needed for the synthesis, at the viewer site, of display holograms is the volume of data to be collected and transmitted and the computational complexity of the hologram synthesis. The upper bound of the amount of data needed to be collected at the scene side and transmitted to the viewer site is, in principle, the full volumetric description of the scene geometry and optical properties. However, a realistic estimation of the amount of data needed for generating a display hologram of the scene is by orders of magnitude lower then the upper bound due to the limitations of the human visual system. This also has a direct impact on the computational complexity of the hologram synthesis.

8.2. Properties of human visual system relevant to holographic 3D visualization

The human visual system has quite a number of natural limitations that must be taken into account in the synthesis of computer-generated display holograms:

- Human vision works in incoherent light;
- At every particular moment, each of two eyes perceives only a small fraction of the incoming wave front limited by the size of the pupil (about 3x3 mm);
- 3D perception is achieved through several mechanisms complementing each other:
  - Eye accommodation;
  - Eye convergence, or the inward rotation of the eyes to converge on objects as they move closer to the observer;
  - Binocular disparity, or the difference in the images projected on the left and right eye retinas in the viewing of a 3D scene. Binocular disparity is the salient depth cue used by the visual system to produce the sensation of depth, or stereopsis.
  - Shading, shadowing and play of highlights on diffuse surfaces that do not have edges or textures capable of producing perception of binocular disparity;
Interposition, or occluding, hiding or overlapping one object by another;
Retinal image size;
Linear perspective;
Aerial perspective;
Motion parallax, which provides different views of a scene in response to movement of the scene or the viewer.

8.3. Binocular vision and its limitations

8.3.1. Redundancy of stereoscopic images: qualitative evaluation

It is well known that stereoscopic images are very redundant. One can make an estimation of the redundancy of stereoscopic images using the following rationale.
From the informational point of view, two images of the same scene that form a stereo pair are equivalent to one of the images and a depth map of the scene. Indeed, from two images of the stereo pair, one can build the depth map, and, vice-versa, one can build a stereo pair from one image and the depth map. Therefore, the increase of the signal volume the second image of the stereo pair adds to that of one image is equal to the signal volume that corresponds to the depth map. The number of depth gradations resolved by vision is of the same order of magnitude as the number of resolved image gray levels. Therefore, signal volume increment due to the depth map is basically determined by the number of depth map independent samples.

Every sample of the depth map can be found by localizing corresponding fragments of two images of the stereo pair and measuring parallax between them. All technical devices that measure depth from stereo work in this way, and it is only natural to assume the same mechanism for stereoscopic vision. The number of independent measurements of the depth map is obviously the ratio of the image area to the minimal area of the fragments of one image that can be reliably localized on another image. It is also obvious, that it is, generally, not possible to reliably localize one pixel of one image on another image. For reliable localization, image fragments should contain several pixels. Therefore, the number of independent samples of the depth map will be, correspondingly, several times lower than the number of image pixels, and the increment of the signal volume that corresponds to the depth map will be several times lower than the signal volume of one image. For instance, if the reliable size of the localized fragment is 2x2 pixels, it will be four times lower, for 3x3 fragments, it will be 9 times lower, and so on. Practical experience tells that, for reliable localization of fragments of one image on another image, the fragment size should usually exceed the area of 8x8 to 10x10 pixels.

Therefore one can hypothesize that the signal volume increment associated with the depth map may amount to percents or even fractions of a percent of the signal volume of one image and that this limitation of the resolution of depth maps acquired from stereo images is true for any device for extraction of depth information from stereo images, including human stereoscopic vision.
8.3.2. Tolerance of 3D visual perception to image blur: quantitative data

Perhaps one of the first observations of the phenomenon of the tolerance of stereoscopic vision to blur of one of stereo images was made by Bela Julesz [5], who however did not mention any quantitative measures of this phenomenon. To the best of our knowledge, the first experiments reported in the literature of decimation and subsequent interpolation of one of two stereoscopic images were described in [6]). Results obtained there are summarized in Fig. 8.4 that shows root mean squared error of parallax measurements on a training stereo air photograph analyzed by a professional human operator using stereo-comparator when one of the images was decimated with different decimation factors and then bilinearly interpolated back to the initial size. The data plotted in Fig. 4 were obtained by averaging of the parallax measuring errors over 31 randomly selected image fragments. Decimation and interpolation were carried out in a computer and the operator was working with computer-generated images of a good photographic quality. Decimation factor zero corresponds to data obtained for initial photos (not computer printouts).

One can see from these data that 4x- and even 5x-decimation/interpolation of one of two images of a stereo pair do not dramatically increase the measurement error. With the increase of the decimation order, RMS error grows according to a parabolic law. These experiments showed also that, after 7x-decimation/interpolations, localization failures appear, and the probability of failures grows with further increase of the decimation factor very rapidly. All this is in a good correspondence with the theory of localization accuracy of image correlators ([7]), although the data were obtained for a human operator.

An extended series of experiments on human stereo vision tolerance to image blur is described in [8]. In these experiments, measured were stereopsis threshold and parallax measurement accuracy as functions of the degree of blur of one of two stereo images. As test images, As test images, the following images were used:
- Grayscale random dots images (see an example in Fig. 8.5(a))
- Grayscale random patches images (Fig. 8.5 (b))
- Grayscale texture images (Figs. 8.5 (c), 8.5 (d))
- Real-life stereoscopic images (see an example in Figs. 8.5 (g), 8.5 (h), right and left, respectively, for crossed eyes viewing)
For each of the test images except the real-life stereo ones, its artificial stereo pair was artificially generated using, as a depth map, impulses of square or round shapes of different dimensions (Figs. 8.6 (a), and (b)). Fig. 8.7 shows examples of stereo images generated in this way.
For each particular experiment, the center of the impulse was randomly placed within the image area so as to exclude the viewer’s adaptation. In the experiments on the stereopsis threshold, the tested values of the depth map scale parameter that determines image parallax were randomly reshuffled for the same purpose. In addition, before displaying each next image, the display was kept blank for several seconds to secure the absence of a cross-talk between individual experiments.

All experiments were carried out with several viewers. For each viewer and for each of the measured parameters, statistics were accumulated over several tens of realizations of random depth map impulse positions. The obtained data were
statistically averaged using arithmetic mean and median averaging and standard deviation of the measured data was determined. In addition, time delay between the prompt to viewer to act and the viewer’s response was measured in all experiments in order to obtain supplementary data to the basic ones.

In the experiments on the stereopsis threshold, synthetic stereo images with different parallax were displayed for viewers who were requested to detect a 3D target by indicating its position using computer generated cross cursor. Computer program registered the correct detection if the cursor was placed within the target object. For each degree of blur, from no blur to maximal blur, the target in form of a circle was randomly placed within the image and its parallax was changed from low to higher values until viewer detects the 3D target after which a new session with another blur started. Both random dot and texture test images were used in the experiments.

In the experiments on the accuracy of parallax measurement, a test target with a moderate parallax was randomly placed within the image and viewers were requested to place a computer generated cross mark to the center of the target as soon as it is detected. Coordinates of the mark were compared by the computer program with actual coordinates of the target and the localization error was calculated. Each measurement was repeated, for the same image blur, several tens of times for statistical estimation of the localization error mean value and standard deviation that were used to characterize the parallax measurement accuracy. Note that parallax values in these experiments were taken to be higher then the threshold values found in the previous experiments. This was done because in localization accuracy experiments it is necessary to see the 3D stimulus clearly.

For image blur, approximated low pass filters were implemented in the domain of Discrete Fourier Transform. The filters were specified by the fraction, from 1 to 0.1 with a step of 0.1, of image bandwidth it passes through. The degree of blur specified by filter bandwidth can be directly translated into the reduction of the number of samples required to generate the blurred image.

Typical results of the experiments are summarized in Figs. 8.8-8.9. Both detection threshold and standard deviation of localization error in these figures as well as depth values indicated in Figs. 8.9 and 8.10 are given in terms of the parallax values introduced for test object, the values being measured in units of inter-pixel distance, such that threshold or error equal to 1 correspond to the parallax in one sampling interval.

Figs. 8.8(a), 8.8(b), represent data obtained in experiments on stereopsis threshold for two test images, random dots and color random patches carried out with image blurring using approximated low pass filter. Blur factor in these graphs indicates fraction of the image base band left by the low pass filter: blur factor one corresponds to no low pass filtering, blur factor two corresponds to low pass filtering to half of the base band, three - low pass filtering to one third of the base band, and so on. For random dots test image, results of 10 experiments for one viewer and their mean and median are shown to illustrate typical data spread in experiments with one viewer. For random patches image, average data for 3 viewers are shown along with their mean and median to illustrate spread of averaged data for several viewers. Figs. 8.8 (c), 8.8 (d) represent stereopsis threshold data obtained for random dot test image using for image blur Haar and Daubechis-1 wavelet low pass filters, correspondingly. In these figures, “Scale index” specifies the scale in wavelet multi-resolution decomposition: scale index one corresponds to the initial resolution, scale index two corresponds to half of the initial resolution, scale index three corresponds to quarter of the initial resolution, four – to one eight of the initial resolution, and so on.
Figure 8.9 represents results of evaluating 3D target localization accuracy obtained, using approximated low pass filter, for different test images, different target depth and stereoscope and anaglyph methods of image display, data obtained for stereoscope and for anaglyph as the viewing device are shown in the left column and in the right column of the figure, correspondingly.

Fig.8.8. Stereopsis threshold as a function of image blur: a) test image “random dots”, case “ideal low pass filter, 10 experiments (dots) with one viewer and data average (red) and median (blue); b) - test image “Color random patches”, case “low pass filter, 3 viewers (dots) and data average (red) and median (blue).

Fig. 8.9. Standard deviation of 3D target localization error as a function of image blur using approximated low pass filter

The results of these experiments clearly show that both threshold and target localization accuracy do not substantially suffer from blur of one of two images up to the blur that corresponds to 5x5 to 7x7 times reduction in the number of pixels of one
of two stereo images. For blur factor larger than 1/8-1/10, rapid growth of the stereopsis threshold and of the localization accuracy was observed.

Refs. [10] reports similar experiments with similar results demonstrating tolerance of human visual system to blur depth map. Experiments with quantization of depth maps of stereo images reported in Ref. [11] evidence also that coarse quantization of depth maps of stereo images to about 20-30 levels does not noticeably affects 3D perception.

8.4. Computer generated display holograms

Several solutions that are computationally inexpensive and at the same time are quite sufficient for creating 3D visual sensation with synthetic display holograms have been suggested ([2]):

- Multiple view compound macro-holograms.

In this method, the scene to be viewed is described by means of multiple view images taken from different directions in the required view angle, and, for each image, a hologram is synthesized separately with an account of its position in the viewing angle (see Fig. 8.10, a). Each hologram has to be, approximately, of the size of the viewer’s eye pupil. These elementary holograms will reconstruct different aspects of scenes from different directions, which are determined by their position in the view angle. The set of such holograms is then used to build a composite, or mosaic, macro-hologram.

![Fig. 10. The principle of synthesis of composite holograms](image)

It is essential that, for scenes given by their mathematical models, well-developed methods and software/hardware instrumentation tools of the modern computer graphics can be used for fast generating multiple view images needed for computing elementary holograms. As for the computational complexity of the synthesis of composite holograms, it can be estimated as following. If individual images are of \( N \times N \) -pixels resolution, the complexity of the synthesis of elementary holograms is of the order of \( N \log N \) operation. Therefore, the computational complexity of synthesis of the composite hologram composed of \( M \times M \) elementary holograms, is of the order \( M^2 N \log N \) operations. Note that the computational complexity of
generating a single hologram of the same size is \( O(M^2 N \log M^2 N) \) which is \( \log M^2 N / \log N \) times higher.

In Refs. [10], an experiment on the synthesis of such a composite macro-hologram composed of 900 elementary holograms of 256x256 pixels was reported. The hologram contained 30x30 views, in spatial angle \(-90^\circ \div +90^\circ\), of an object in the form of a cube. The synthesized holograms were encoded as kinoforms and recorded with sample size 12.5 mcm. The physical size of elementary holograms was 3.2x3.2 mm. Each elementary hologram was repeated, in the process of recording, 7x7 times to the size 22.4x22.4 mm. The size of the entire hologram was 672x672 mm². Being properly illuminated, the hologram can be used for viewing the reconstructed scene from different angles, as, for instance, through a window (Fig. 8.11, left). Looking through the hologram with two eyes, viewers are able to see 3D image of a cube (Fig. 8.11, right) floating in the air.

![Fig. 8.11 Viewing compound computer-generated hologram (left) and one of the views reconstructed from the hologram (right). The entire macro-hologram was composed of 900 elementary kinoform holograms of 256x256 pixels, which corresponded to 30x30 views in spatial angle \(-90^\circ \div +90^\circ\). The holograms were recorded with pixel size 12.5 mcm. Size of the elementary hologram was 3.2x3.2 mm. Each elementary hologram was repeated 7x7 times to the size 22.4x22.4 mm.](image)

- **Composite stereo-holograms.**

  A special case of multiple view mosaic macro-holograms is composite stereo-holograms. Composite stereo holograms are synthetic Fourier holograms that reproduce only horizontal parallax ([4, 11, 12]). When viewed with two eyes as through a window, these are capable of creating 3D sensation thanks to stereoscopic vision. With such holograms arranged in a circular composite hologram, full 360 degrees view of the scene can be achieved. Fig. 3 shows such a hologram and examples of images from which it was synthesized ([4,11]). The entire hologram was composed of 1152 fragmentary kinoform holograms of 1024x1024 pixels recorded with pixel size 12.5 mcm. The total size of the hologram was 240 cm. In stationary state of the hologram, viewer looking through the hologram from different positions was able to see a 3D image of an object in a form of a “molecule” of six “atoms” differently arranged in space. When the hologram rotated, the viewer was able to see “atoms” continuously rotating in space and easily recognize the rotation direction.
• “Programmed diffuser” holograms.

The “Programmed diffuser” method for synthesis of Fourier display hologram was suggested for generating digital holograms capable of reconstructing different views of 3-D objects whose surfaces scatter light diffusely ([3, 10]). This method assumes that objects are specified in the object coordinate system \((x, y, z)\) by their “macro” shape \(z(x, y)\), by the magnitude of the object reflectivity distribution \(A(x, y)\) in the object plane \((x, y)\) and by the directivity pattern of the diffuse component of its surface. The diffuse light scattering from the object surface is simulated by assigning to the object a pseudo-random phase component (a “programmable diffuser”), whose correlation function corresponds to the given directivity pattern of the object surface. This pseudo-random phase component is combined with the deterministic phase component defined by the object shape to form the distribution of the phase of the object wavefront. A flow diagram of synthesis of “programmed diffuser” holograms is presented in Fig. 8.13.

Fig. 8.13 Block-diagram of synthesis of programmed diffuser holograms
Holograms synthesized with this method exhibit spatial inhomogeneity that is directly determined by the geometrical shape and diffuse properties of the object surface. This allows imitation of viewing the object from different direction by means of reconstruction of different fragments of its “programmed diffuser” hologram as it is illustrated in Fig. 8.14 on an example of a cone shaped object.

Fig. 8.14 Programmed diffuser hologram of a cone-shaped object and results of its reconstruction from its different fragments

Fig. 8.15 presents frames of a movie that show reconstruction of programmed diffuser hologram of a cone-shaped (left) and hemisphere-shaped uniformly painted object on which one can clearly see a play of light patches on the objects in different viewing positions associated with different fragments of holograms. Fig. 8.16 presents one more simulation example of reconstruction of programmed diffuser hologram of a hemisphere with 9 digits painted on its surface, and Fig. 8.17 illustrate optical reconstruction of computer generated programmed diffuser hologram of a pyramid and a hemisphere. Note in conclusion that for recording such holograms all encoding methods described in Lect. 5 can be used including kinoform.

Fig. 8.15 Frames of a movie demonstrating computer simulated reconstruction of programmed diffuser holograms of cone-shaped and hemisphere-shaped objects
Fig. 8.16. Object’s image (upper left), object’s shape (center left) and its programmed diffuser hologram (bottom left; circles on the hologram outline different reconstruction windows) and nine images reconstructed from northwest, north, northeast, west, center, east, southwest, south and southeast fragments of the hologram (right)

Fig. 8.17. Examples of results of optical reconstruction of programmed diffuser holograms of objects in form of a pyramid (left) and hemisphere (right)
8.5. Digital holographic displays

One of the major obstacles, if not the major one, for implementing 3D digital holographic visual communication is the lack of digital holographic displays, devices for converting data on the distribution of the amplitude and phase in the hologram plane into optical holograms for visual reconstruction of 3D images. One can distinguish two types the holographic displays: static displays for reproducing holograms of static scenes similarly to printers in conventional imaging and dynamic ones for reproducing holograms of dynamic scenes similarly to television monitors. While creation of dynamic digital holographic displays still remains quite problematic, static digital holographic displays such as digital holographic printers are becoming feasible.

Important practical issues in the design of devices for displaying computer generated display holograms are using low coherent and white light for reconstruction of the holograms and encoding of holograms for displaying color 3D images. In what follows we discuss possible solutions.

8.5.1. Low coherent light and white light computer-generated holograms. Hybrid optical-digital holograms.

As computer-generated holograms reproduce distribution of the object wave front amplitude and phase, computed for a certain light wavelength, it is a common belief that, for displaying computer-generated holograms, one needs to use a laser. However, this is not the case. The incoherence of the light beam used for reconstruction of holograms influences the accuracy in reconstruction of the phase of the hologram and may cause image blur due to different scaling of reconstructed images for different wavelengths of the light. In order to keep the image blur due to the imperfect coherence of the light on an acceptable low level, it is sufficient to use a reconstruction source of light with the degree of coherence, measured as the ratio of the light frequency to its spectrum spread, of the order of the largest, in horizontal or vertical dimension, number of pixels in the scene images (say, 1000). This is by 3-4 orders of magnitude lower than that needed for recording optical holograms. Obviously, for such incoherence, the accuracy in hologram phase reproduction will be of the order of 0.1%, which is also quite sufficient.

The use, for reconstruction of computer-generated holograms of such sources of low coherence quasi-monochromatic light can simplify the design of the digital holographic displays and, in addition, is beneficial in terms of reducing reconstruction speckle noise.

Ideally, the use, for reconstruction of hologram, of the natural white light would be the most appropriate. The great advantage of Denisyuk-type optical holograms is that that they do not need, for reconstruction, a coherent light of any coherence and are reconstructed in white light. This property will also be very beneficial for computer-generated holograms, especially for static holographic displays. At least three methods can be suggested for producing computer-generated holograms capable of reconstruction in white light:

1. Recording computer-generated hologram onto optical media pre-exposed by a Denisyuk-type hologram of the reconstruction source of light.
2. Making “sandwich-holograms” out of computer-generated holograms of the object to be displayed and a Denisyuk-type hologram of the reconstruction source of light.
3. Holographic copying onto a Denisyuk-type hologram of an image reconstructed, in coherent light, from the computer-generated hologram

8.5.2. Computer generated color display holograms

Of course, computer-generated display holograms must ultimately be capable of reproducing color image. Color CG-holograms can be generated as color-separated RGB holograms and recorded by interlacing their corresponding samples covered by the respective RGB-filters. Fig. 8.18 illustrates one possible hologram encoding scheme for recording color computer generated holograms on a phase media. The scheme is an extension of the so-called double-phase encoding method ([3, 4, 14], see Lect. 5). The use of phase media for recording computer-generated holograms is advantageous in terms of hologram light efficiency.

For encoding the amplitude and phase of RGB samples of the computed object wavefront in the hologram plane, groups of three cells arranged in the hexagonal sampling grid are used per each RGB component, each group for one of R, G and B sample (see Fig. 8.18, a).

Each of RGB cell groups is covered by corresponding R, G or B color filters. Each \( k \)-th cell in the \( n \)-th R-, G- or B- group modulates phase of the illuminating beam by introducing phase shift \( \varphi_{R,n}^{(k)} \), \( \varphi_{G,n}^{(k)} \) and \( \varphi_{B,n}^{(k)} \), correspondingly. The phase shifts \( \varphi_{R,n}^{(k)} \), \( \varphi_{G,n}^{(k)} \) and \( \varphi_{B,n}^{(k)} \) for encoding of \( n \)-th sample of the corresponding hologram color component with amplitude \( A_{c,n} \) and phase \( \Phi_{c,n} \) are found, according to the drawing in Fig. 6, b), from the relationships:

\[
A_{c,n} \exp(i\Phi_{c,n}) = A_0 \left[ \cos(\Phi_{c,n} - \varphi_{c,n}^{(1)}) + 1 + \cos(\varphi_{c,n}^{(3)} - \Phi_{c,n}) \right]
\]

where \( C = R, G, B \).

Note, that such color holograms can be viewed in white light provided RGB-filters are appropriately narrow-band.
References


[hologr.displ_mask,OUTIMG1,OUTIMG2,OUTIMG3,OUTIMG4,OUTIMG5,OUTIMG6,OUTIMG7, OUTIMG8,OUTIMG9]=progrm_diff_tst(ones(512),20,120,50);