L. Yaroslavsky
School of Electrical Engineering,
Tel Aviv University,
Tel Aviv, Israel

Computer-Generated Holograms and 3-D Visual Communication: a Tutorial

3DTV-CONFERENCE 2008
28-30 MAY 2008, ISTANBUL, TURKEY

Chinese “magic mirrors” (copper-plate engraving)
Invention of Holography

Dennis Gabor, inventor of holography (Nobel Prize in Physics, 1971), and his holographic portrait.

From D. Gabor’s Nobel Lecture

From D. Gabor’s Nobel Lecture: “...The response of the optical industry to this was so disappointing that we did not publish a paper on it until 11 years later, in 1966 (5). Around 1955 holography went into a long hibernation.

The revival came suddenly and explosively in 1963, with the publication of the first successful laser” holograms by Emmett N. Leith and Juris Upatnieks of the University of Michigan, Ann Arbor. Their success was due not only to the laser, but to the long theoretical preparation of Emmett Leith, which started in 1955. This was unknown to me and to the world, because Leith applied his ideas first to the problem of the “side-looking radar” which at that time was classified. This was in fact two-dimensional holography with electromagnetic waves, a counterpart of electron holography. When the laser became available, in 1962, Leith and Upatnieks could at once produce results far superior to mine”.

"Train and Bird", the first hologram ever made with a laser using the off-axis technique. This pioneer image was produced in 1964 by Emmett Leith and Juris Upatnieks at the University of Michigan
Yu. Denisyuk’s method of recording “volume” holograms

Computer-Generated Display Holography: Motivation

Holographic imaging is the only method that is capable of reproducing, in the most natural viewing conditions, 3-D images that have all 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 and signal processors are ideal means for converting data on 3-D scenes into synthetic optical holograms for visual perception.
The idea of making synthetic optical devices capable of producing, from light, visual objects dates back to middle-age Chinese magic mirrors. Some science museums have ancient Chinese "magic mirrors" on display. These look just like a flat mirror, but if you bounce the sun off them, you'll see a picture in the projected sunbeam.

A pre-History of Computer-Generated Holography

- B. R. Brown, A. Lohmann, Complex spatial filtering with binary masks, Appl. Optics, 5, No. 6, 967-969, 1966


Basic stages of computer synthesis of holograms

1. Mathematical model of the scene
2. Complex amplitude of the scene wave field
3. Computing the mathematical hologram by means of imitating scene-to-viewer wave propagation
4. Scene observation geometry
5. Computer
6. Encoding the mathematical hologram for recording computer generated hologram
7. “Digital-to-analog conversion”
8. Recording computer generated hologram
Two basic problems in computer synthesis of holograms:

• Computationally efficient digital representation of wave propagation integrals

• Converting numerical data that describe the “mathematical” hologram to a physical optical hologram capable of reproducing a viewable image
Reduced 2D model of scene-to-viewer wave propagation

Scene plane

Observation (hologram) plane

\[ \alpha(\bar{f}) = \int_{\bar{x}} a(\bar{x}) T_z(\bar{x}, \bar{f}) d\bar{x} \]
# Wave propagation diffraction integrals

<table>
<thead>
<tr>
<th>Transform</th>
<th>Integral Formula</th>
</tr>
</thead>
</table>
| **Kirchhoff-Rayleigh-Sommerfeld integral transform** | \[
\alpha(\tilde{\mathbf{r}}) = \int_{-\infty}^{\infty} a(\tilde{x}) \exp\left(i2\pi \frac{Z\sqrt{1+\|\tilde{x} - \tilde{\mathbf{r}}\|^2/Z^2}}{\lambda} \right) d\tilde{x}
\] |
| **Fresnel integral transform**                      | \[
\alpha(\tilde{\mathbf{r}}) = \int_{-\infty}^{\infty} a(\tilde{x}) \exp\left(i\pi \frac{\|\tilde{x} - \tilde{\mathbf{r}}\|^2}{\lambda Z} \right) d\tilde{x}
\] |
| **Angular spectrum propagation transform**           | \[
\alpha(\xi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(x) \exp\left(i2\pi \tilde{x} \xi \right) dx \exp\left(-i\pi \lambda Z \|\xi\|^2 \right) \exp\left(-i2\pi \xi \tilde{\mathbf{r}} \right) d\xi
\] |
| **Fourier integral transform**                      | \[
\alpha(\tilde{\mathbf{r}}) = \int_{-\infty}^{\infty} a(\tilde{x}) \exp\left(-i2\pi \frac{\tilde{x} \tilde{\mathbf{r}}}{\lambda Z} \right) d\tilde{x}
\] |
## Discrete Fourier Transforms

<table>
<thead>
<tr>
<th>Transform Type</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canonical Discrete Fourier Transform</td>
<td>$\alpha_r = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \exp\left(i2\pi \frac{kr}{N}\right)$</td>
</tr>
<tr>
<td>Shifted DFT</td>
<td>$\alpha_{r,v} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \exp\left(i2\pi \frac{(k+u)(r+v)}{N}\right)$</td>
</tr>
<tr>
<td>Discrete Cosine Transform</td>
<td>$\alpha_{r}^{DCT} = \frac{2}{\sqrt{2N}} \sum_{k=0}^{N-1} a_k \cos\left(\pi \frac{k+1/2}{N} - r\right)$</td>
</tr>
<tr>
<td>Discrete Cosine-Sine Transform</td>
<td>$\alpha_{r}^{DcST} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \sin\left(\pi \frac{k+1/2}{N} - r\right)$</td>
</tr>
<tr>
<td>Scaled DFT</td>
<td>$\alpha_{r}^{\sigma} = \frac{1}{\sqrt{\sigma N}} \sum_{k=0}^{N-1} a_k \exp\left(i2\pi \frac{(k+u)(r+v)}{\sigma N}\right)$</td>
</tr>
<tr>
<td>Scaled DFT as a cyclic convolution</td>
<td>$\alpha_{r}^{\sigma} = \frac{1}{\sqrt{\sigma N}} \sum_{k=0}^{N-1} a_k \exp\left(i\pi \frac{\tilde{k}^2}{\sigma N}\right) \exp\left[-i\pi \frac{(k-\tilde{k})^2}{\sigma N}\right]$</td>
</tr>
<tr>
<td>Canonical 2D DFT</td>
<td>$\alpha_{r,s} = \frac{1}{\sqrt{N_1N_2}} \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} a_{k,l} \exp\left[i2\pi \frac{kr}{N_1} + \frac{ls}{N_2}\right]$</td>
</tr>
<tr>
<td>Affine DFT</td>
<td>$\alpha_{r,s} = \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} a_{k,l} \exp\left[i2\pi \left(\frac{rk}{\sigma A N_1} + \frac{sk}{\sigma c N_1} + \frac{rl}{\sigma B N_2} + \frac{sl}{\sigma D N_2}\right)\right]$</td>
</tr>
<tr>
<td>Rotated &amp; Scaled DFT (RotScDFT)</td>
<td>$\alpha_{r,s} = \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} a_{k,l} \exp\left[i2\pi \left(\frac{rk + sl}{\sigma N} \cos \theta - \frac{sk - rl}{\sigma N} \sin \theta\right)\right]$</td>
</tr>
<tr>
<td>Discrete Sinc-function</td>
<td>$\text{sinc}(N, x) = \frac{\sin x}{N \sin(x/N)}$</td>
</tr>
</tbody>
</table>
**FAST TRANSFORMS FOR DIGITAL HOLOGRAPHY:**
Discrete Fresnel and Kirchhoff-Reyleigh-Sommerfeld Transforms

<table>
<thead>
<tr>
<th>Transform Type</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canonical Discrete Fresnel Transform (DFrT)</td>
<td>[ \alpha_r = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \exp \left[ i\pi \left( \frac{k/\mu - r\mu}{N} \right)^2 \right] \quad \mu^2 = \lambda Z / N \Delta f^2 ]</td>
</tr>
<tr>
<td>Shifted DFrT</td>
<td>[ \alpha_{r(\mu,w)} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \exp \left[ -i\pi \left( \frac{k\mu - r / \mu + w}{N} \right)^2 \right] \quad w = u/\mu - v\mu ]</td>
</tr>
<tr>
<td>Fourier Reconstruction algorithm for Fresnel holograms</td>
<td>[ \alpha_{r(\mu,w)} = \frac{1}{\sqrt{N}} \exp \left( -i\pi \frac{r^2}{\mu^2 N} \right) \sum_{k=0}^{N-1} a_k \exp \left[ -i\pi \left( \frac{k\mu + w}{N} \right)^2 \right] \exp \left( i2\pi \frac{k + w / \mu}{N} r \right) ]</td>
</tr>
<tr>
<td>Convolutional Discrete Fresnel Transform (ConvDFrT)</td>
<td>[ \alpha_r = \sum_{k=0}^{N-1} a_k \text{frincd}(N;\mu^2;r + w - k) = \frac{1}{N} \sum_{s=0}^{N-1} \sum_{k=0}^{N-1} a_k \exp \left( i2\pi \frac{k - r - w}{N} s \right) \exp \left( -i\pi \frac{\mu^2 s^2}{N} \right) ]</td>
</tr>
<tr>
<td>Frincd-function</td>
<td>[ \text{frincd}(N;q;x) = \frac{1}{N} \sum_{r=0}^{N-1} \exp \left( i\pi \frac{qr^2}{N} \right) \exp \left( -i2\pi \frac{xr}{N} \right) ]  [ \text{frincd}(N;\pm q;x) = \pm \frac{i}{\sqrt{Nq}} \exp \left( \mp i\pi \frac{x^2}{qN} \right) \text{rect} \left( \frac{x}{q(N-1)} \right) ]</td>
</tr>
<tr>
<td>Discrete Kirchhoff-Rayleigh-Sommerfeld Transform (DKRST)</td>
<td>[ \alpha_r = \sum_{k=0}^{N-1} a_k \frac{\exp \left[ i2\pi \frac{\tilde{z}^2 \sqrt{1 + (\tilde{k} - \tilde{r})^2/\mu^2}}{\mu^2 N} \right]}{1 + (\tilde{k} - \tilde{r})^2/\tilde{z}^2} ]</td>
</tr>
</tbody>
</table>
Spatial light modulators (SLM) for recording CGHs

\[ I_{out} \exp(i \varphi_{out}) = A_{SLM} \exp(i \theta_{SLM}) I_{in} \exp(i \varphi_{in}) \]

- **Amplitude SLM:**
  - Modulated parameter \( A_{SLM} \)

- **Phase SLM:**
  - Modulated parameter \( \theta_{SLM} \)

- **Amplitude/phase SLM:**
  - Modulated parameters \( A_{SLM} \) & \( \theta_{SLM} \)
Encoding methods for recording CG-holograms:

“Phase-detour” method for binary amplitude SLM

Direction to the observation plane

Direction of image reconstruction

Hologram encoding cell

Encoded computer generated hologram

Nodes of the regular sampling grid

Wave propagation transform

Amplitude
Phase

Computer Generated Binary Hologram

Reconstructed image

$\phi = \frac{2\pi \alpha \Delta \xi}{\lambda} \cos \theta$
**Encoding methods for recording CG-holograms:**

“Phase-detour” method for continuous-tone amplitude SLM

Orthogonal and bi-orthogonal representation of complex numbers

Three versions of using SLM resolution cells for orthogonal/biorthogonal encoding of hologram samples

<table>
<thead>
<tr>
<th>Redundancy</th>
<th>Encoding Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>2:1</td>
<td>( \text{Re} + \text{bias} ) (0^\circ) (\text{Im} + \text{bias} ) (90^\circ) (-\text{Re} + \text{bias} ) (180^\circ) (-\text{Im} + \text{bias} ) (270^\circ)</td>
</tr>
<tr>
<td>4:1</td>
<td>( \text{e}<em>{re} ) (0^\circ) (\text{e}</em>{im} ) (90^\circ) (-\text{e}<em>{re} ) (180^\circ) (-\text{e}</em>{im} ) (270^\circ)</td>
</tr>
<tr>
<td>2x2:1</td>
<td>(-\text{e}<em>{re} ) (180^\circ) (-\text{e}</em>{im} ) (270^\circ) (\text{e}<em>{re} ) (0^\circ) (\text{e}</em>{im} ) (90^\circ)</td>
</tr>
</tbody>
</table>
Encoding methods for recording CG-holograms:
“Phase-detour” method for continuous-tone amplitude SLM

Simplex representation of complex numbers

Three versions of using SLM resolution cells for symplex encoding of hologram samples

Redundancy 3:1
Encoding methods for recording CG-holograms:

“Symmetrization” method for continuous tone amplitude SLM

Symmetrization by duplicating

Symmetrization by quadruplicating

Redundancy 2:1

Redundancy 2x2:1
Encoding methods for recording CG-holograms:

“Double-phase” and “Multiple-phase” method for phase SLM

Representation of complex numbers by a sum of equal-length vectors

Phase detour two-phase coding for hologram recording on a binary phase SLM: two separate cells and decomposition of cells into alternative sub-sells

An example of arrangement of SLM resolutions cells in triple-phase encoding of samples of ”mathematical” holograms
Encoding methods for recording CG-holograms: Kinoform: phase-only holograms

Iterative algorithm for generating pseudo-random phase mask for kinoform
Reconstruction of computer generated holograms: illustrative examples

- Phase detour encoding method
- Symmetrization method
- Reconstruction of Fresnel Hologram
- Reconstruction of combined Fresnel/Fourier Hologram
Reconstruction of computer generated holograms: Point Spread Function

\[ I(x, y) = \int WPK(x, y; \xi, \eta) CGH(\xi, \eta) d\xi d\eta \]

\[ CGH(\xi, \eta) = \sum_{k,l} EMH_{r,s} RA(\xi - r\Delta\xi; \eta - s\Delta\eta) \]

\[ EMH_{r,s} = ENCODING \left\{ \sum_{k,l} DWPK(k,l;r,s) Obj(k,l) \right\} \]
Computer-generated display holography and 3D visual communication: Questions to begin with

- For 3D communication, does one need to produce, at the scene side, a hologram of the scene and to transmit it to the viewer’s site?
- Does one need to imitate at the viewer side full optical hologram of the scene? What features of optical holograms computer-generated display holograms should maintain?
- What is the upper bound of the amount of data needed to be transmitted to the viewer side?
- What is a realistic estimation of the amount of data needed for generating, at the viewer site, a hologram of the scene?
- As computer-generated holograms reproduce distribution of the object wave front amplitude and phase, computed for a certain wave length, does one necessarily need to use, for displaying computer-generated holograms, a laser?
Computer-generated display holography for 3D visual communication:
Outline of the approach

• FOR 3D COMMUNICATION, ONE DOES NOT NEED TO PRODUCE, AT THE SCENE SIDE, A HOLOGRAM OF THE SCENE AND TO TRANSMIT IT TO THE VIEWER’S SITE. INSTEAD ONE NEEDS TO COLLECT, AT THE SCENE SIDE, AND SEND, TO THE VIEWER SIDE, A SET OF DATA, WHICH WILL BE SUFFICIENT TO GENERATE, AT THE VIEWER SITE, A SYNTHETIC HOLOGRAM OF THE SCENE FOR VISUAL OBSERVATION
Computer-generated display holography for 3D visual communication: Outline of the approach

For generating synthetic holograms at the viewer side, there is no need to imitate full optical holograms of the scene.

The major requirement to computer-generated display holograms is that they should provide natural viewing conditions for human visual system and visual separation of the reconstructed image from the display device.
Computer-generated display holography for 3D visual communication: Outline of the approach

The upper bound of the amount of data to be collected at the scene side and transmitted to the viewer site is the full volumetric description of the scene geometry and its optical properties.

In the design of computer-generated display holography systems, one should take full advantage of using natural limitations of the human visual system.

Due to these limitations, a realistic estimation of the amount of data needed for generating, at the viewer site, a hologram of the scene is by orders of magnitude lower than the upper bound.
Limitations of the human visual system

- Human vision works in incoherent and non-monochromatic light;
- At every particular moment, each of two eyes perceives only a small fraction of the incoming wavefront limited by the size of the pupil (about 3x3 mm);
- 3D perception is achieved through several mechanisms complementing each other:
  - eye accommodation and eye convergence
  - binocular disparity;
  - linear perspective; aerial perspective; interposition, or occluding, hiding or overlapping one object by another;
  - shading, shadowing and play of highlights on diffuse surfaces that do not have edges or textures capable of producing perception of binocular disparity;
  - retinal image size;
  - motion parallax
Limitations of binocular vision

- The resolving power of binocular vision in measuring scene depth is by the order of magnitude lower that the vision resolving power in terms of the number of resolvable pixels in the scene.
- The number of quantization levels in scene depth map required to secure the absence of visual artifacts does not exceed couple of tens

This means that the increment in the amount of data, which should be added to the scene image data to satisfactorily, for binocular vision, describe the scene depth map, accounts only several percents of the amount of scene image data.
Holographic displays and reconstruction of C-G display holograms: no laser needed

For displaying computer-generated holograms, the ratio of its light frequency to the width of its spectrum spread should be of the order of the largest (vertical or horizontal) size (in pixels) of the scene images (~1000). This is by 3-4 orders of magnitude lower than what is needed for recording optical holograms.

Therefore, sources of quasi-monochromatic light with this ratio of the order $10^3$ are sufficient as computer-generated hologram illumination devices.

In addition to simplifying hologram reconstruction device, this is beneficial also in terms of reducing reconstruction speckle noise.
Compound macro-holograms: a practical solution for computer-generated display holograms

- Compound macro-holograms provide, in a very natural way, binocular disparity, the salient depth cue used by the visual system to produce the sensation of depth; they can also provide eye accommodation.
- They are much easier to compute and fabricate than the corresponding whole holograms. The computational complexity of computing compound hologram is $O(NM\log N)$ ops vs $O(NM\log NM)$ for computing one large hologram of the same size.
- They are very well suited for parallel computations.
- They can be computed using soft-and-hardware tools of 3D computer graphics.

Object

Compound macro-hologram: a 2-D mosaic of $M \times M$ elementary holograms of $N \times N$ samples.

Elementary hologram. The size of elementary holograms should be commensurable with the size of the eye pupil (3-5 mm in diameter); $N = O(1000)$ samples.
Three methods for generating compound display macro-holograms

- Circular stereo Fourier holograms that reproduce horizontal parallax*
- Mosaic Fourier holograms that reproduce both vertical and horizontal parallax*
- “Programmed diffuser” holograms that reproduce highlights on diffuse surfaces of the displayed objects viewed from different directions

*For fast computing different views of objects given by their mathematical description, soft&hardware tools of modern computer graphics can be used
Computer generated circular stereo hologram and movie

Figure shows a circular hologram composed of 1152 fragmentary kinoform (phase-only phase) holograms of 1024x1024 pixels recorded with pixel size 12.5 mcm. Total size of the circular composite hologram was 240 cm.

Viewing CG macro-holograms with white-light illumination source

The entire macro-hologram was composed of 900 elementary holograms of 256x256 pixels, which corresponded to 30x30 views in spatial angle –90° - +90°. The holograms were recorded with pixel size 12.5 μm. Size of the elementary hologram was 3.2x3.2 mm. Each elementary hologram was repeated 7x7 times to the size 22.4x22.4 mm.
“Programmed diffuser” method is a method for generating digital holograms capable of reconstructing different views of 3-D objects whose surfaces scatter light diffusely. The method assumes that objects are specified by their shape \( z(x,y) \), by the magnitude of the object reflectivity distribution \( A(x,y) \) given in the object plane \((x,y)\) and by the spatial 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 phase distribution of the object wave front.

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 imitating viewing the object from different direction by means of reconstruction of different fragments of its “programmed diffuser” hologram.
“Programmed diffuser” method for synthesis of display Fourier holograms

Object reflectivity distribution $A(x,y)$

Object 3-D shape $z(x,y)$

Object’s wave front complex amplitude $A(x,y)\exp[i2\pi(z(x,y)+\text{PrDiff}(x,y))]$

PrDiff($x,y$)

Programmed diffuser generator

Wave propagation transform (e.g. DFT)

Reconstructed images

Hologram

3-D object’s model

Object’s surface directivity pattern

Deterministic component

Random component
Examples of images of a complex scene reconstructed from “programmed-diffuser” holograms

An object and its views from different parts of its PrDiff-hologram
Examples of images reconstructed from programmed diffuser holograms (simulated movies)
Examples of images reconstructed from programmed diffuser holograms (a simulated movie)
Examples of images optically reconstructed from CG “programmed diffuser” holograms

Pyramid

Hemisphere
### White light computer-generated display holograms

AT LEAST THREE METHODS CAN BE SUGGESTED FOR PRODUCING COMPUTER-GENERATED HOLOGRAMS CAPABLE OF RECONSTRUCTION IN WHITE LIGHT:

- Making a hybrid hologram by recording computer-generated hologram onto optical media pre-exposed by a Denisyuk-type hologram of the reconstruction source of light

- Making “sandwich-hologram” out of a computer-generated hologram of the object to be displayed and a Denisyuk-type hologram of the reconstruction source of light

- Making a Denisyuk-type hologram of an image reconstructed, in coherent light, from the computer-generated hologram
COLOR CG-HOLOGRAMS CAN BE GENERATED AS COLOR-SEPARATED RGB HOLOGRAMS AND RECORDED BY INTERLACING THEIR CORRESPONDING PIXELS COVERED BY THE RESPECTIVE RGB-FILTERS.

Such color holograms can be viewed in white light provided RGB filters are appropriately narrow-band.
References

- L. Yaroslavsky, N. Merzlyakov, Methods of Digital Holography, Consultens Bureau, N.Y. 1980
Digital Holography and Digital Image Processing
Principles, Methods, Algorithms

Digital holography and digital image processing are twins born by computer era. They share origin, theoretical basis, methods and algorithms. The book describes these common fundamental principles, methods and algorithms including image and hologram digitalization, data compression, digital transforms and efficient computational algorithms, statistical and Monte-Carlo methods, image restoration and enhancement, image reconstruction in tomography and digital holography, discrete signal resampling and image geometrical transformations, accurate measurements and reliable target localization in images, recording and reconstruction of computer generated holograms, adaptive and nonlinear filters for sensor signal perfecting and image restoration and enhancement.

Digital Holography and Digital Image Processing
Principles, Methods, Algorithms combines theory, heavily illustrated practical methods and efficient computational algorithms, and is written for senior-level undergraduate and graduate students, researchers and engineers in optics, photonics, opto-electronics and electronic engineering.