COMPUTER GENERATED DISPLAY HOLOGRAPHY

Course notes for the tutorial presented at Eurographics 2017

Lyon, France, 24 April 2017

Petr Lobaz

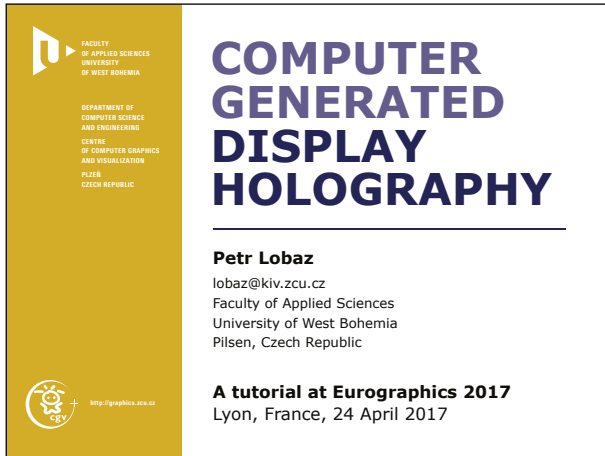lobaz@kiv.zcu.cz
Faculty of Applied Sciences
University of West Bohemia
Pilsen, Czech Republic

Please check **http://holo.zcu.cz** for the last revision
of this document and for sample scripts

COMPUTER GENERATED DISPLAY HOLOGRAPHY

FACULTY
OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

DEPARTMENT OF
COMPUTER SCIENCE
AND ENGINEERING
CENTRE
OF COMPUTER GRAPHICS
AND VISUALIZATION
PLZEŇ
CZECH REPUBLIC

Petr Lobaz
lobaz@kiv.zcu.cz
Faculty of Applied Sciences
University of West Bohemia
Pilsen, Czech Republic

http://graphics.zcu.cz

A tutorial at Eurographics 2017
Lyon, France, 24 April 2017

The name of the tutorial is "Computer generated display holography", which deserves some explanation.

Computer generated display holography (Eurographics 2017)   slide 2 / 126
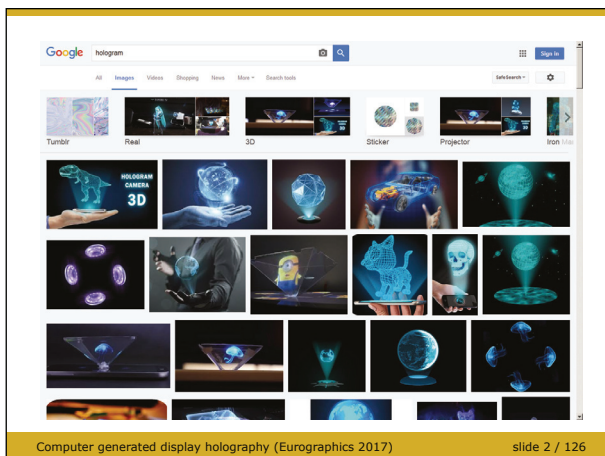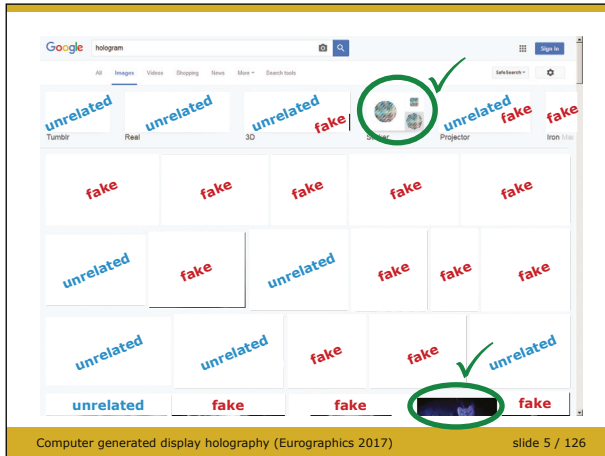
A realistic 3-D illusion is often called "a hologram", stunning displays are often called "holographic". However, if you search for the word "hologram" on the Internet, the results are quite confusing.

Fake images:

Computer generated display holography (Eurographics 2017)   slide 3 / 126
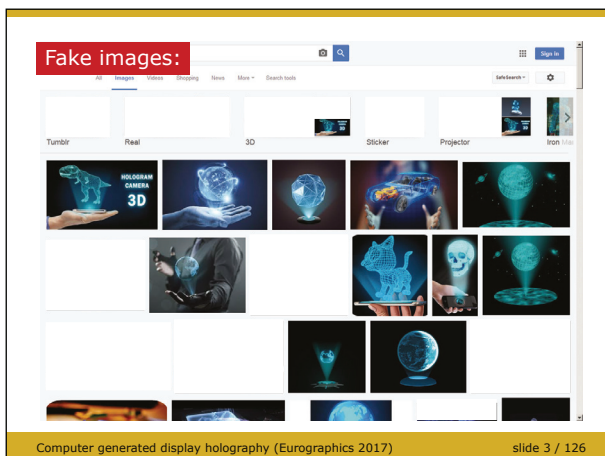
First of all, most found images are fakes, i.e., they are just a result of image retouching. Not only they are not photographs of real devices – current physics just does not have tools to make such displays.

Unrelated to holography:

Computer generated display holography (Eurographics 2017)   slide 4 / 126
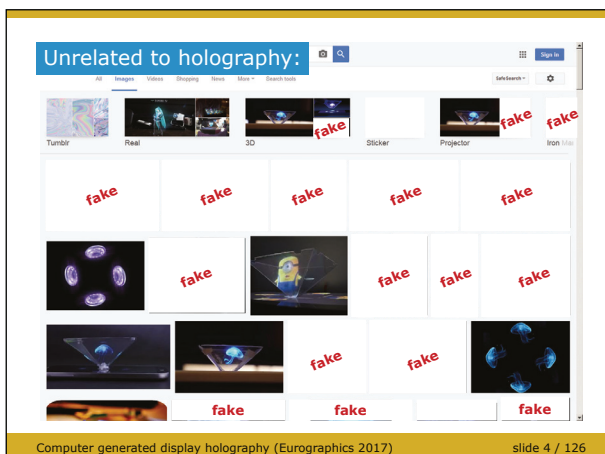
Most remaining images are real photographs of devices that have nothing to do with holography. It just happened that the term "hologram" is so fancy that it is often used in advertising new display technologies even if they utilize completely different principles.
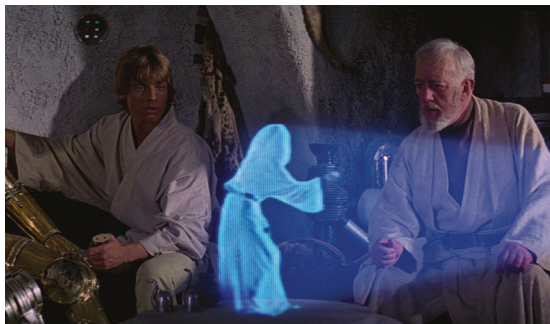
Tutorial contents

Star Wars: A New Hope (directed by G. Lucas, 1977)

Microsoft HoloLens: visualization of augmented reality

In this particular search, the only two remaining images are related to holography. The one in the upper part of the screen is a sticker with a diffractive structure; this particular one is, strictly speaking, not a hologram too. So the only real hologram on this screen is partly displayed in the last row of images.

The first aim of this tutorial is thus to clarify what is a hologram. Besides that, I will show what technologies are often inappropriately called "holographic".
Holography is a technology that offers ultimate viewing experience. However, it is not easy to understand it without good knowledge of optics, especially Fourier optics. Digital holography further requires complex analysis, signal processing, and, indeed, computer graphics. Before we delve into details, I will show how holography works in principle. After that, I will introduce basic algorithms that generate a simple hologram. Then I introduce more advanced algorithmic techniques and true holographic displays and show competing technologies. At the end, I will give some advices how to start with computer generated display holography.

Non-holographic technologies
If you ask "what is a hologram", most people recall sci-fi movies and imagine something like "futuristic display". However, illusions such as the one in this iconic picture cannot be produced with current knowledge of physics.

A lot of manufacturers advertise their products as "holographic", for example Microsoft HoloLens. Campaigns are often accompanied by pictures such as this one. Again, it is just a fake image. In particular, HoloLens is an augmented reality system that requires a user to wear special goggles. Images such as this one just give an impression what the user sees. Indeed, anyone without the goggles does not see the illusion.
Strictly speaking, HoloLens technology uses optical elements (waveguides) that have something to do with holography. As "hologram" sounds better than "waveguide", the name HoloLens emerged.

Kagamine Rin & Len at a Hatsune Miku concert

Then we have a broad variety of stage illusions – the whole stadium sees that a "hologram" appeared on the stage. Common technologies such as rear projection to a transparent screen, or projection to a screen and its reflection from a semi-transparent mirror are utilized. It follows that these illusions are just 2-D images that appear to be present on the stage.
For details, see for example [Maas].
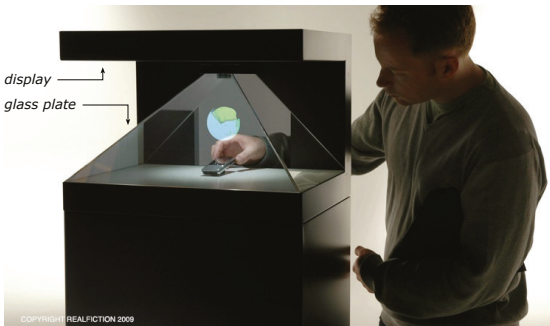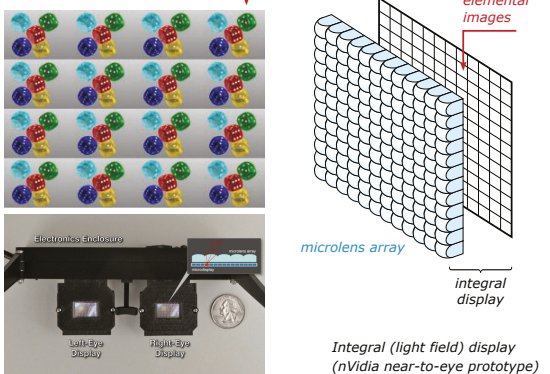

display
glass plate

Cheoptics 360™ by viZoo

Exactly the same principle, often called (a bit imprecisely) "Pepper's ghost", is utilized in small to medium sized displays that show an object seemingly floating in the air. Such displays, often in a shape of a pyramid, just reflect an image produced by an ordinary 2-D display by a semi-transparent mirror. However, if the object displayed rotates, changes, etc., the observer is fooled enough to believe he or she sees a perfect 3-D illusion. See for example [Simonsen]

display

mirror

rotating base

360° Light Field Display
University of Southern California
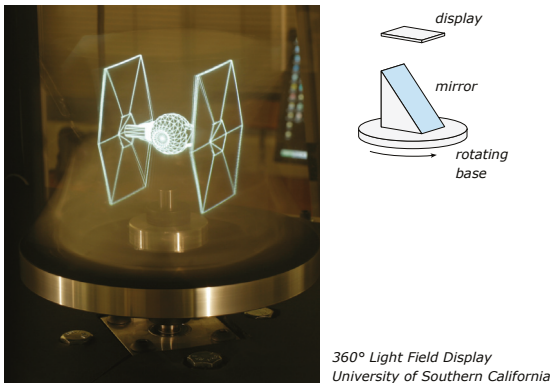
There are, of course, technologies that provide true 3-D illusion, such as the rotating display, see for example [Favalora, Jones]. Despite that authors of such displays often do not claim they are holographic (as they are not), they are coined as such in newspapers, on TV, etc.


elemental images

microlens array

integral display

Integral (light field) display
(nVidia near-to-eye prototype)

One of the most promising technologies for 3-D display is based on G. Lippmann's idea of "integral display"; it is often called "light field display" as well, although light field is just a theoretical concept not related to any display technology.
Similarly to holographic displays, they use a flat display and can provide goggles-free 3-D illusion. Contrary to holographic displays, they do not rely on wave optics principles and it seems that affordable integral displays could be build with current technology.
For details, see [Okoshi, Lueder]
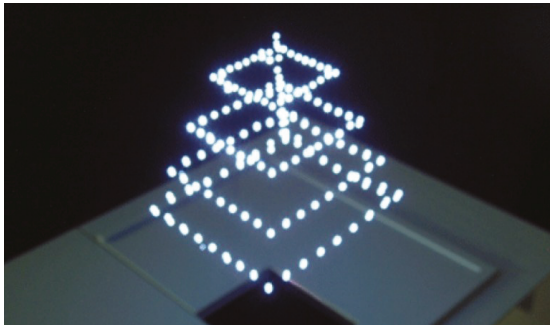
Plasma volumetric display by Burton Inc.

Finally, we should not forget attempts to display a 3-D illusion that really floats in space, can be observed from anywhere and are closest to "sci-fi holograms". There are several attempts how to make such an illusion; one of them uses lasers to make points in the air glow, see for example [Ochiai]. However, as the glowing points can be seen from anywhere, such displays cannot show opaque objects as every surface is always visible.
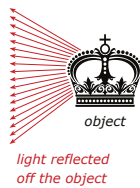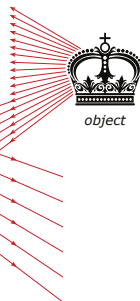
Ultimate 3-D illusion

Let us think for a while how to provide a perfect illusion, visually indistinguishable from the real world.

Recall that our eyes respond just to light that enters their pupils. Thus, in order to see an object, some light must be reflected off its surface, and some light rays must find their way to the retina.

A light ray need not to travel in a straight line. For example a mirror changes its direction abruptly.
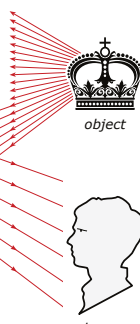
A mirror does not bend light rays arbitrarily. It bends them so that they appear to originate from the "mirror image" of their source. Thus, an observer looking towards the mirror sees the original object behind the mirror surface. Strictly speaking, the observer sees the virtual image of the original object. If the mirror is perfect, it is hard to tell if we are looking at the original object or its virtual image. Anyone who ever visited a mirror maze can confirm that.
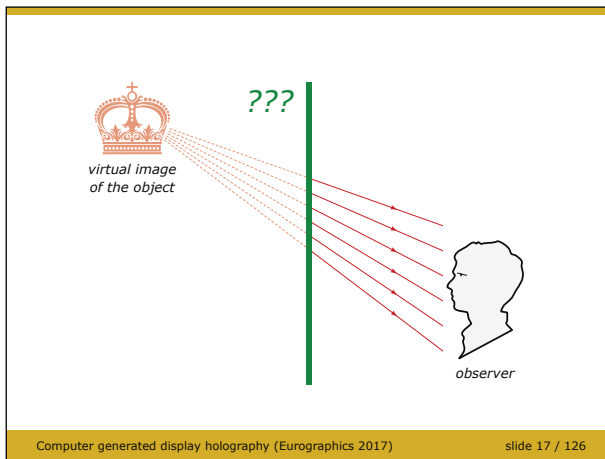
In order to create a perfect illusion, it is necessary to "freeze light" somehow – to make a surface that emits exactly those light rays that were leaving the mirror on the previous slide.

As the light rays are exactly the same, the observer cannot tell if he or she watches the original object or its virtual image.

The points of the hypothetical display could emit light themselves such as CRT or OLED displays. Or, such as in LCD displays, some backlight could be provided. In this case, the task of the hypothetical display is to split and bend light rays from the light source so that light rays produced are the same as the light rays formed by the original object.

In order to make such a hypothetical display, two questions have to be answered: how to capture complete information about light rays leaving the object, and how to replicate them.
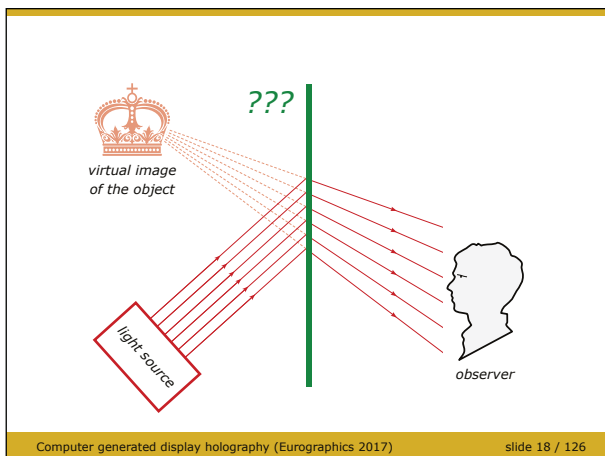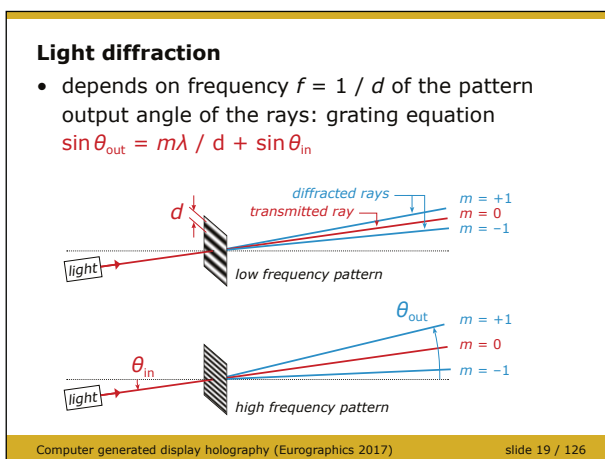
Both questions were answered by Dennis Gabor in his seminal papers [Gabor48, Gabor49]. He proposed a new method for lensless image capture. He called the image formed by the process "the hologram"; the method itself was later named "holography".

Light diffraction
- depends on frequency $f = 1 / d$ of the pattern output angle of the rays: grating equation
$\sin \theta_{out} = m\lambda / d + \sin \theta_{in}$

Recall that the hypothetical display has to bend and split rays coming from the backlight. Light diffraction does exactly this. A light ray passing through a fine locally periodic structure of stripes splits to several new rays called diffracted rays. Their direction depends on stripes' distance d, wavelength of light λ and direction of the original ray. If a structure, called a diffractive structure, forms several rays at once (it depends on stripes' properties), they are numbered by integer m, and they are called diffraction orders. Diffraction order $m = 0$ is just the directly transmitted ray. For now, let us remember the most important fact: fine stripes bend a ray a lot, coarse stripes bend a ray a little.

Image formation by means of diffraction
- grating equation: $\sin\theta_{out} = m\lambda / d + \sin\theta_{in}$
- example: λ = 0,5 μm   d = 10 μm θ_{in} = 0   m = 1
  $\Rightarrow \theta_{out}$ = 2,87°

Thus, to make an illusion of a point floating behind the display surface, the diffraction pattern has to vary its properties. Somewhere it is necessary to bend the backlight a lot, thus the structure has to be fine. Somewhere else the backlight rays are almost in the right direction, therefore the structure can be rough. It is easy to see that the structure has to look like a set of concentric circles and their density grows from their common centre.

Now it is time to substitute some real-world numbers to the grating equation. Wavelength of visible light is between 0.4 and 0.7 μm; let us pick 0.5 μm (green-cyan) as an example. If the stripes are just 10 μm apart (that is, one cycle from opaque through transparent to opaque is 10 μm), then the ray deflection is only 2.87° for diffraction order $m = 1$ (we are usually interested in this diffraction order). This is not too impressive – when watching a common display, the field of view is about 30°. Such a big deflection angle would require stripe width about 1 μm.

Diffraction pattern formation using interference
- an interference pattern can be recorded and subsequently used as a diffraction pattern

Besides the original two questions, a new one appears: how to fabricate such a fine pattern? Luckily, there is a phenomenon that gives an answer: light interference. If two lights are mutually coherent (for example coherence of laser light is quite good), they interfere and create a pattern of light and dark stripes that can be recorded by a fine photographic film. For example, let us imagine a wide beam of laser light illuminating a screen (or a photographic film) and a dust particle in between. Light scatters on the dust particle and the particle starts to behave as a point light source. The screen is thus illuminated by two light sources – from the laser unit and from the dust particle. These two lights interfere and create *exactly the pattern we want*.

Light interference
- two "coherent" light beams "interfere": create a pattern of light and dark stripes

light intensity on the screen: the interference pattern

$$d = \frac{\lambda}{\sin\theta_A - \sin\theta_B}$$

example: λ = 0,5 μm
θ_A = 45°
θ_B = −45°
$\Rightarrow$ d = 0,35 μm

To explain why, we need to understand light interference quantitatively. If two "coherent" lights illuminate a screen, they form an interference pattern that locally looks like a set of straight stripes. Their width depends on angles of the forming light rays and their common wavelength.

It it worth mentioning that it is quite tricky to define coherence. For now, let us just say that if two lights interfere, they are mutually coherent. This is not a very useful definition, indeed, but we will not need rigorous theory of coherence. See for example [Saleh] for more details on this subject.

Principle of holography

- interference equation: $\quad d = \lambda_1 / (\sin\theta_A - \sin\theta_B)$
  grating equation: $\quad \sin\theta_{out} = m\lambda_2 / d + \sin\theta_{in}$
- after substitution of d: the $\sin\theta$ equation

$\sin\theta_{out} = m\dfrac{\lambda_2}{\lambda_1}(\sin\theta_A - \sin\theta_B) + \sin\theta_{in}$

- for m = 1, $\lambda_1 = \lambda_2$, $\sin\theta_B = \sin\theta_{in} \Rightarrow \sin\theta_{out} = \sin\theta_A$

photographic film — *light B* — θ_A — θ_B — *light A* — **recording step**

photographic film — *diffracted light* — θ_{out} — θ_{in} — *transmitted light* — **reconstruction step**

Now it is straightforward to combine principles of interference and diffraction. We know that two lights interfere and make a pattern of certain density described by the interference equation. We also know that light passing through a pattern of certain density deflects; this is described by the grating equation. We can combine both equations together and get "the sine-theta equation" (the name is the same as in [Benton]).

Now imagine a simple experiment. Let us illuminate a photographic film by two beams. Let their angles of incidence be θ_A and θ_B. An interference pattern is created and recorded by the film. Then, let us illuminate the film by a light ray at an angle θ_B. Substitution to the sine-theta equation reveals that the first order diffracted ray leaves the film at an angle θ_A. That is, the pattern encoded the ray direction somehow, and we are able to reconstruct it.

- hologram: the interference pattern of
  – an object wave: $\qquad\qquad \theta_{obj} (= \theta_A)$, $\lambda_1 = \lambda_{ref}$
  – a reference wave: $\qquad\quad \theta_{ref} (= \theta_B)$, $\lambda_1 = \lambda_{ref}$
- hologram observation: illuminate it by
  – an illumination wave: $\qquad \theta_{ill} (= \theta_{in})$, $\lambda_2 = \lambda_{ill}$
- $\sin\theta_{out} = m\dfrac{\lambda_{ill}}{\lambda_{ref}}(\sin\theta_{obj} - \sin\theta_{ref}) + \sin\theta_{ill}$
- example: $\lambda_{ill} = \lambda_{ref}$, $\theta_{ill} = \theta_{ref} = 0$

recording　　**reconstruction (m = +1)**　　**reconstruction (m = −1)**

$\theta_{ref} = 0$ — θ_{obj1} — θ_{obj2} — *light sensitive material*

$\theta_{ill} = 0$ — θ_{out1} — θ_{out2} — *virtual image* — *hologram*

$\theta_{ill} = 0$ — θ_{out1} — θ_{out2} — *hologram* — *real image*

It is now very easy to describe holography; in fact, we have described it right now. Let us call light leaving an object as "the object wave" (the term wave is used because interference and diffraction rely on wave nature of light). Let us put a photographic film somewhere so that it is illuminated by the object wave. Let us illuminate the film with additional light called "the reference wave". If the reference wave and the object wave are coherent, the interference pattern is formed and recorded by the film. This recording is called "the hologram".

To explain how the hologram works, let the object wave be light leaving a point in space. Also for simplicity, let us assume that reference rays are mutually parallel. We can imagine that the hologram is split to tiny elementary areas so that each area is illuminated by a single ray from the object and a single reference ray. Each elementary area thus records a simple interference pattern. Its density changes across the hologram area because the angle between the object and the reference ray varies slightly.

If we illuminate the hologram by the copy of the reference rays ($\theta_{ref} = \theta_{ill}$, $\lambda_{ref} = \lambda_{ill}$), the object rays are perfectly reconstructed for diffraction order $m = 1$ (after substitution to the sine-theta equation, we get $\theta_{out} = \theta_{in}$). That is, if an observer looks towards the hologram, he or she sees the virtual image of the original point. As a general object can be decomposed to many point light sources, the same would apply for a general object.

We can also see an additional detail. If we use the reference light at normal incidence to the hologram ($\theta_{ref} = \theta_{ill} = 0$) and substitute $m = -1$ to the sine-theta equation, we get $\theta_{out} = -\theta_{in}$. That is, if we place a sheet of paper to a certain distance from the hologram, a bright point appears. At this place, a real image of the original point was formed.

We are usually interested in just one of these diffraction orders, either +1 or −1; the other one is then treated as the unwanted diffraction order and we usually look for a way how to avoid it.

Virtual image formation

- illuminate hologram with a light source
- light beams diffract on the interference pattern
- diffracted rays are the same as the rays from the original object
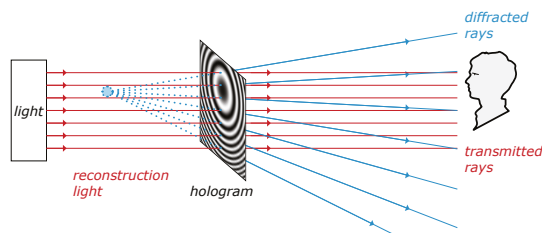
Real image formation

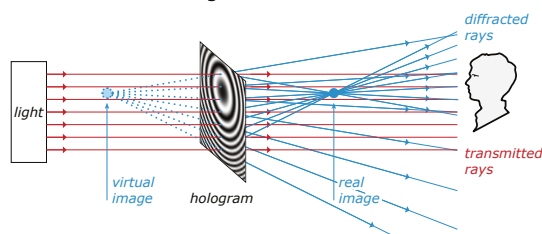- output angle of the rays: $\sin\theta_{out} = m\lambda / d + \sin\theta_{in}$
- for $m = -1$, rays can create real image of the scene
- both rays for $m = +1$ and -1 appear at once
  $\Rightarrow$ no need to distinguish between them

Classical holography

- capturing the interference pattern of laser lights using a photosensitive material
  - requires high quality lasers
  - requires high resolution recording materials (currently up to 10 000 lines/mm)
  - requires vibration-free environment
  - usually requires chemical processing
- reconstructing the hologram using light source
  - custom lighting setup required
- properly recorded and illuminated holograms provide ultra realistic image

Digital holography (DH)

- light sensitive sensor (e.g. CCD or CMOS) instead of photochemical light sensitive material
  - very fast
  - cannot capture high spatial frequencies (currently about 250 lines/mm)
- numerical simulation of the hologram reconstruction
- digital processing of the captured hologram instead of its visual inspection
  - automatic evaluation
  - allows processing hard to achieve in classical holography

Let us summarize our findings. A hologram is a fine pattern that diffracts light. If we illuminate it from behind with suitable light source, some diffracted light diverges and creates a virtual image. It is possible to make a hologram and to illuminate it in such a way that the virtual image is the exact replica of the original object.

Sometimes, some diffracted light creates other images, for example the −1st diffraction order often creates the real image. Recall that the real image can be seen in front of the hologram if the eye is located in the bunch of rays behind it. Moreover, if we place a sheet of paper to the location of the real image, we observe it right there.

If we calculate diffraction of a ray and both ±1st diffraction orders appear, it is usually not clear which ray makes the real image and which one the virtual one. However, it is rather academic question: as they appear at once, there is no need to distinguish between them. Moreover, idea of a ray is not very appropriate here, as light diffraction is better described using wave model of light.

Now it is time to define some terms. By "classical holography", I mean something similar to classical photography: capturing light using a photosensitive material, chemical development of the recording and its proper illumination. Unfortunately, making a hologram is not easy, especially outside a lab. On the other hand, contemporary holography provides ultra realistic imaging, see [Bjelkhagen13].

Digital holography relates to classical holography in the same sense as digital photography to classical (analogue) photography. Some part of the process, namely the light recording, introduces electronic digital sensors. A digital hologram can be "printed" and observed as a classical one, but more frequently this process is simulated computationally. Moreover, it is possible to apply various signal processing procedures to a digital hologram to denoise it, remove unwanted diffraction orders and so on. On the other hand, while electronic image sensors usually outperform classical photography, this is not the case in digital holography.

Computer generated holography (CGH)
- numerical simulation of the
  hologram recording process ("sort of")
- electronic display of a hologram
  – e.g. microdisplays with very fine pixels
  (spatial light modulators),
  currently up to 130 lines/mm
- "printing a hardcopy"
  – laser lithography
  expensive, up to 600 lines/mm
  – electron beam lithography
  very expensive, up to 10 000 lines/mm
- other technologies – let us talk about them later

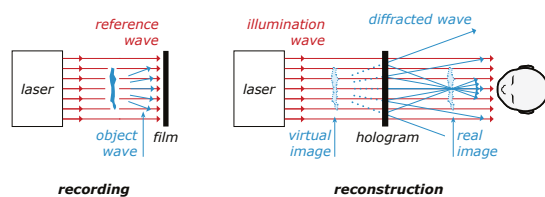**Computer generated display holography (CGDH)**
- computer generated hologram of a 3-D scene
  for **display purposes**
- computer graphics
  – makes a digital image to be displayed
  on a common electronic display
- computer generated display holography
  – makes a pattern to be displayed
  on a holographic display
- combination approaches are common,
  e.g., computer graphics for image rendering,
  subsequent classical holography for making an
  interference pattern

Basic hologram recording setups
- on-axis (Gabor) hologram
  – mostly for transparent objects (restrictive)
  – image damaged by the 0th order,
  ±1st orders overlap (bad)
  – low spatial frequencies (100 lines/mm – good)

Computer generated holography is maybe a part of digital holography, maybe its supplement. Its task is to make a diffractive structure (a hologram) computationally. It can simulate hologram recording process or it can take a completely different approach. Once the digital hologram is calculated, it is usually printed or displayed on a suitable electronic display. Unfortunately, electronic displays are still far worse than holographic recording materials; and "printing" a computer generated hologram is usually very expensive. Please note that while it is definitely possible to simulate the hologram reconstruction process (as in digital holography), it is usually employed just for evaluation purposes; otherwise, it makes little sense.

Computer generated display holography (CGDH) is a part of computer generated holography. It can be thought as an counterpart of computer graphics. While computer graphics can make a digital image of a 3-D scene, CGDH can make its digital (computer generated) hologram.

Hologram recording setups

There were two emphasised phrases in the last two slides: "hologram recording process" and "display purposes". The first one indicates we should understand classical hologram recording process before we delve into its computational simulations. The second one indicates there are other than display applications of holography.

Let us start with basic hologram recording setups. The oldest one is called on-axis or Gabor setup. It is used mostly for fully transparent objects, such as microscopic mounts. The laser light shines through the object that diffracts it a bit. A holographic recording material ("film" for short) captures the interference of the direct and the diffracted light. When the object is removed, the hologram reconstructs its both real and virtual images. Unfortunately, an observer sees both of them simultaneously. Moreover, the observer looks directly into the laser light. It follows that the on-axis setup has little use in display holography. On the other hand, it is often employed in technical holography and especially in technical digital holography. Main reason is that the interference pattern is comparatively coarse, and contemporary electronic sensors cannot handle fine patterns.

As a side note, it should be noted that Gabor used this setup in 1948 because he did not have laser light (laser was invented in 1960's). Actually, this setup works with light whose coherence is far from perfect.

- off-axis transmission (Leith-Upatnieks) hologram
  - for both opaque and transparent objects
  - clear image (good)
  - high spatial frequencies (1000 lines/mm – bad)
 - visible in laser light only (uncomfortable)
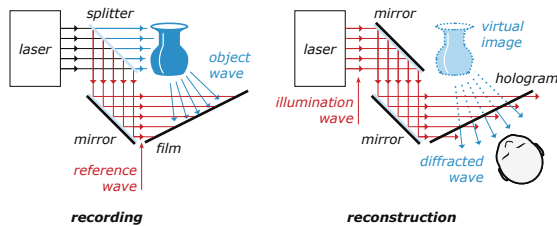
The problem of overlapping images was solved by E. Leith and Y. Upatnieks (U.S.A.) in 1960's. They used just invented laser light which allowed them to separate the object wave from the reference wave. In consequence, the observer does not look into laser light. Moreover, the real image is formed next to the virtual image, which means that the virtual image can be seen unobtruded (and in fact, the real image is not formed at all in some cases). Unfortunately, off-axis holography requires high resolution recording material and it must be observed in laser illumination. If we want to make an off-axis CGH, we should be prepared for high pixel density (up to 2000 pixels/mm).

- reflection (Denisyuk) hologram
  - the simplest setup (good)
  - visible in white light (good)
  - simply allows colour imaging (very good)
  - high spatial frequencies (4000 lines/mm – bad)
  - the diffraction pattern is volumetric, i.e., 3-D, not planar, i.e., 2-D (very bad)
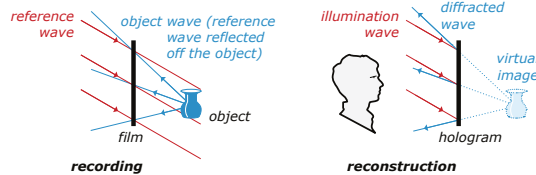
Meanwhile, Y. N. Denisyuk (USSR) independently developed a completely different approach to holography. Here, the object and the reference waves illuminate the recording medium ("film") from opposite sides. The recording medium must be "thick" (about 10 μm is common) because volumetric structure of the interference pattern must be recorded. Moreover, the medium must capture very fine details, e.g. 4000 lines/mm (in volume!), which actually prevents calculation of such pattern. On the other hand, Denisyuk holograms (often called reflection holograms) can be viewed in common white light and easily provide full colour imagery – which is why it is so suitable for any display holography.

Applications of holography
- cultural heritage conservation
  - holograms instead of real exhibits
  - the exhibit too valuable or fragile, multiple exhibitions at once, multiple views of the same exhibit at once
  - almost perfect image of the exhibit, scale 1 : 1

A full colour Denisyuk hologram of the "15th anniversary Fabergé Easter egg", A. Sarakinos, HIH, 2015.

Applications of holography

The most natural application of holography is thus ultra realistic imaging, for example for cultural heritage conservation. The example provided here is a full colour Denisyuk hologram of a very valuable exhibit, see [Bjelkhagen16] for details. When properly illuminated, the observer hardly recognizes if he or she watches the original subject or a perfect illusion.

- microscopy, visual inspection
  1. perfect recording of light
     (from a biological sample, a bubble chamber, …)
  2. hologram examination
     (unlimited time of observation,
     examination in safe environment,
     holograms can be archived, …)

Recall that a hologram records *everything*, which means that a reconstructed hologram provides the same information as the original object. It is thus for example possible to take a hologram of a biological sample and to inspect it under a microscope instead of the original sample. The advantage is clear: the hologram does not move, can be observed indefinitely, can be archived, etc. In the same way, it is possible to take a hologram in a dangerous environment and to observe it in a safe laboratory, and so on.

- digital holographic microscopy
  – acquisition of a digital hologram
  – numerical reconstruction
    ⇒ signal filtering, unwanted diffraction removal, numerical analysis, …
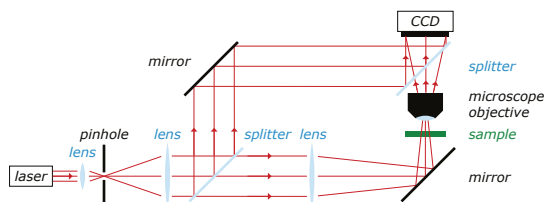
Using classical holography in microscopy is a bit uncomfortable as it is slow (making a classical hologram is about as slow as making a classical photograph). It is thus advantageous to introduce an electronic sensor (e.g. CCD) for light recording. The captured digital hologram can be reconstructed numerically. As the hologram captures everything, it is possible to refocus, change the angle of observation, and so on. Please note that the sequence "laser – lens – pinhole" is used to clean the beam from imperfections. It is called *spatial filtering* and it is used in most holographic setups.

- enhancing electron microscopy
  – original D. Gabor idea behind holography
    (although in fact, it never worked)
  – hologram recording with electron beam
    (λ is 100 000× smaller than for visible light)
  – hologram enlargement, visible light illumination
    ⇒ image 100 000× bigger

in the $\sin\theta$ equation: $\lambda_{ill} / \lambda_{ref} = 100\,000$

$\sin\theta_{out} = m \dfrac{\lambda_{ill}}{\lambda_{ref}} (\sin\theta_{obj} - \sin\theta_{ref}) + \sin\theta_{ill}$

In fact, Gabor invented holography to improve electron microscopy. His idea was simple. Recall that in the sine-theta equation, there is the factor $\lambda_{ill} / \lambda_{ref}$, where λ_{ref} is the wavelength of light used in hologram recording and λ_{ill} is the wavelength used in hologram reconstruction. Until now, we assumed they are the same. If they are not and $\theta_{ref} = \theta_{ill} = 0$, the output angles are magnified – which is exactly the same what a microscope does. Gabor assumed electron waves for a hologram recording and visible light for its reconstruction, which means that the angular magnification would be $\lambda_{ill} / \lambda_{ref} \approx 100\,000\times$, far bigger than in common microscopy. In fact, this idea never worked in practice, but it shows hologram versatility.

- holographic optical elements (HOE)
  - mimicking any optical element
  - cheaper, easier aberration correction, …
  - also called diffractive optical elements (DOE)
    (the difference between HOE and DOE is subtle)
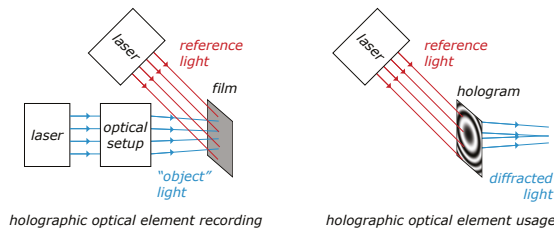
holographic optical element recording

holographic optical element usage

- example: holographic optical element
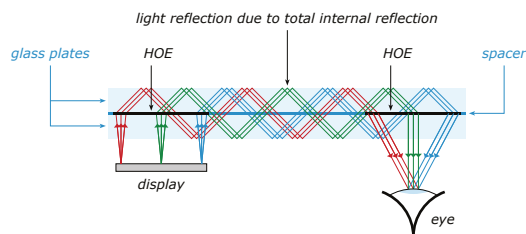 (waveguide coupler) for augmented reality
  head-up displays

- non-destructive testing
  - double object recording on one hologram: shifts
    between recordings smear hologram fringes
  - taking a hologram of a vibrating object:
    vibration causes loss of hologram fringes
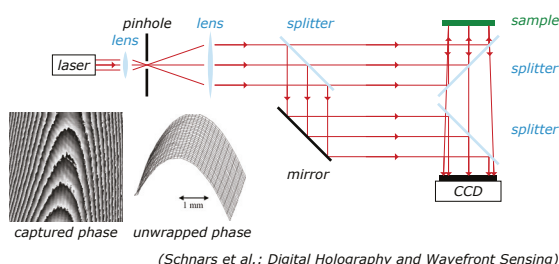 - ⇒ no fringes = no image = black strips on the
    object
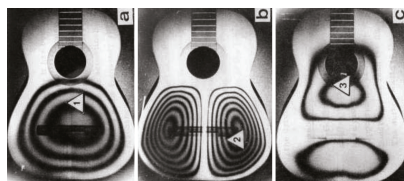
K. Molin, N. Stetson, Institute of Optical Research, Stockholm (1971)

- surface metrology
  - digital hologram of a real object
  - numerical reconstruction of a hologram
  - reconstructed phase ~ surface bumpiness

captured phase   *unwrapped phase*

(Schnars et al.: Digital Holography and Wavefront Sensing)

A hologram, as we know changes direction of light rays. Any optical setup does this. Thus, what about using a hologram as an optical element? We can illuminate an expensive optical setup and record output light as a hologram. This light can be subsequently easily reproduced. Moreover, as making a hologram duplicate is cheap, we can use the cheap flat hologram instead of the original expensive and bulky optical setup. Moreover, it is possible to design a holographic optical element that alters light in a way hardly possible with conventional optics.

For example, holographic optical elements can be used as the input and output elements in a waveguide for an augmented reality system. Here, light leaving a display enters the first HOE that mimics a prism – it bends its direction. Light then travels inside a glass plate due to total internal reflection. Finally, light enters the second HOE. It acts as a prism cemented to a lens – first, it bends light rays so that they can leave the waveguide again, and second, it focuses them to the observer's eye.

I have mentioned that no motion is allowed when making a hologram. It is not surprising – as a hologram is composed of very fine fringes (1000 lines/mm is common), any movement during recording can be significant. Thus, if we need to know if an object moves or vibrates, we can make a hologram – and if it moved, black area appears instead of the object when reconstructing the hologram. More details can be found, e.g., in [Collier].

Holography is also used in surface metrology. Recall that a perfectly flat surface illuminated by a plane wave (i.e., light rays are parallel) reflects a perfect plane wave. Any surface imperfection causes some change in phase of reflected light. In digital holography, it is possible to capture a hologram of the reflected light and analyse its phase on the reflecting surface. Thus, imperfections in order of nanometres can be measured. More details can be found, e.g., in [Schnars]

- remote digital holographic interferometry
  – hologram of a master sample (A)
  – reconstruction of a real image of a master over a tested object B
  – contours ~ objects differences

object B *real image of a master*

observer

splitter

pinhole *lens*
lens
SLM
laser
mirror

(Schnars et al.: Digital Holography and Wavefront Sensing)

master

contours

It is also possible to compare two objects using holography, for example a certified specimen to a just fabricated product. For example, the real image of the certified specimen can be superimposed on the product. If they are not perfectly the same, interference strips appear. Moreover, they can be evaluated and the difference can be measured. It should be emphasised that the certified specimen can be safely stored. If we employ digital holography, the hologram of the specimen can be taken in the laboratory and just transmitted to the production facility. For more details, see again [Schnars]

- embossed holograms
  – bumpy surface diffracts light ⇒ surface relief hologram:

metallic layers

foil

illuminated holographic material

holo. material after special development

master stamp production

relief duplication by embossing

  – making a master stamp expensive
  – making embossed copies cheap
  – can contain hidden features
  ⇒ hard to counterfeit
  ⇒ suitable as a security element

hologram by Optaglio

The most ubiquitous application of holography is, however, a metallic-looking sticker used as a security element on credit cards, passports, banknotes, etc. Here, a master hologram is created in such a way that its relief is bumpy, i.e. it diffracts reflected light. Such a master is metallized and a stamp is created. Thus, any number of copies can be made by embossing it to a plastic foil. Making a master hologram is quite expensive and difficult, but it is used for a million of copies. Thus, each copy is cheap and hard to counterfeit. For more details, see for example [Renesse, Saxby]

Until now, our model of holography was very informal. In order to calculate a hologram, we must understand light in more detail. In this part of the tutorial, I will explain the most important aspects and show actual code that calculates a simple hologram.

Light is just a special electromagnetic force between charged particles. In particular, if a charged particle oscillates, it can cause oscillation of another charged particle in our retina. Some movements of charged particles in the retina cause neural reaction, which means "we see the light source – the original oscillating particle". We can imagine that the original particle moves "up and down" according to the cosine function of certain amplitude, frequency and phase. In the distance r from the particle (let us call it "the light source"), the electromagnetic field can be described by a quantity $u(r, t)$, see the slide. Here we are not interested if u stands for the electric field, for the magnetic field, what is the direction of the field – we are just interested in "how big is the field". We call this approximation "the scalar theory of light"

– see [Goodman] for details. Anyway, we can see on the slide that the optical field $u(r, t)$ is again a cosine function with another amplitude, another phase and the same frequency. We can also see that the amplitude and the phase are actually functions of r.

BASIC TOOLS OF COMPUTER GENERATED DISPLAY HOLOGRAPHY

Nature of light

- force interaction between (oscillating) point charges
- a point source of light:
  movement up and down $\sim A \cos(\varphi - \omega t)$
  (amplitude, phase)
- optical field ($\sim$ electromag. force) at a distance r:

$$u(r, t) = \frac{A}{r} \cos\left[\varphi - \omega\left(t - \frac{r}{c}\right)\right] = A'(r) \cos(\varphi'(r) - \omega t)$$

(amplitude at r, phase at r)

T	period of oscillation	1.7×10^{-15} s
$f = 1/T$	(time) frequency	600 THz
$\omega = 2\pi/T$	angular frequency	
c	speed of light	
$\lambda = cT$	wave length	0.5 μm
$k = 2\pi/\lambda$	wave number	1.2×10^7 m^{-1}

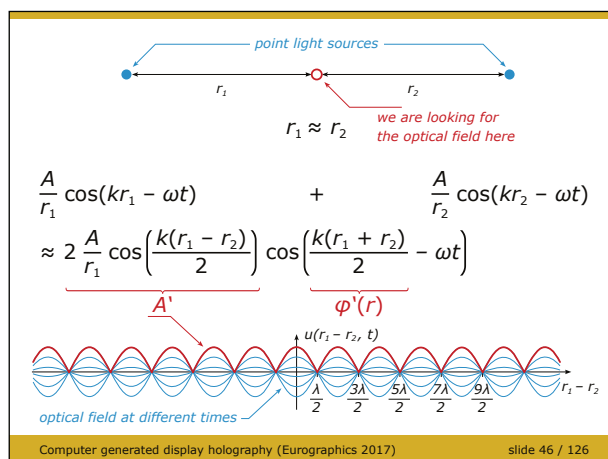First, let us explore interference of light. Imagine there are two point light sources (i.e., two oscillating charged particles), and we want to know optical field somewhere between them. As electromagnetic field is linear, we just need to calculate the optical field of each light source and to sum them. If we assume $r_1 \approx r_2$, we can rewrite the sum to the final form. We see that the final formula is again a time varying cosine function with the same angular frequency ω and some phase φ'. We can also see that the final amplitude A' depends on $r_1 - r_2$. In particular, there are

points in space where oscillation is high, and there are point where amplitude of oscillation is zero.

---

point light sources

$r_1 \approx r_2$

we are looking for the optical field here

$$\frac{A}{r_1} \cos(kr_1 - \omega t) \quad + \quad \frac{A}{r_2} \cos(kr_2 - \omega t)$$

$$\approx 2\frac{A}{r_1} \cos\left(\frac{k(r_1 - r_2)}{2}\right) \cos\left(\frac{k(r_1 + r_2)}{2} - \omega t\right)$$

A' $\varphi'(r)$

$u(r_1 - r_2, t)$

optical field at different times

- photographic film reacts on time average of light intensity $\propto (A')^2$
- $\Rightarrow$ cannot distinguish close "dimmer" light from distant "brighter" light

- constructive
  $\times$
  destructive
  interference

intensity of light

Light intensity is proportional to square of amplitude of oscillation – and a photographic film, a digital image sensor or a retina are sensitive to light intensity. We can thus conclude that there are locations in space where light intensity is high, and places where there is no light. We say that somewhere, light interfered constructively, and somewhere else destructively.

By the way, as photographic film cannot record phase of light, it cannot distinguish between "bright distant light" and "dim close light" – and this is the reason why photography lacks 3-D.

Phasor arithmetic

- $j^2 = -1$
- $e^{jx} = \cos x + j \sin x$

- $u(r, t) = A(r) \cos[\varphi(r) - \omega t] = \mathrm{Re}\{A(r)\, e^{j[\varphi(r) - \omega t]}\}$
  $= \mathrm{Re}\{\underbrace{A(r)\, e^{j\varphi(r)}}_{\text{phasor}} e^{-j\omega t}\}$

- phasor (complex amplitude):
  $U(r) = A(r)\, e^{j\varphi(r)}$

- light amplitude: $A = |U|$
  light phase: $\varphi = \arg(U)$
- light intensity:
  $I = |U|^2 = U U^* = A\, e^{j\varphi}\, A\, e^{-j\varphi} = A^2$

While our analysis was correct, the calculation was a bit uncomfortable. In the end, we were interested in light amplitude (intensity) only, but we had to work with time-dependent cosine function.

The phasor arithmetic helps us to get rid of time dependency. We use a mathematical trick: a cosine function is a real part of a complex exponential, and the exponentials can be easily factored. Thus, we can split time and space dependency. The complex function $U(r)$ dependent on space only is called *the complex amplitude* or *the phasor*. In fact, it encodes both real amplitude and phase into one complex number. Finally, as light intensity is proportional to the square of amplitude, it can be easily calculated as $|U(r)|^2 = U(r)\, U^*(r)$, where * denotes complex conjugate.

Advantage of phasor arithmetic

- optical field – time dependent function:
  $u(r, t) = A(r) \cos(\varphi(r) - \omega t)$
- its phasor (complex amplitude):
  $U(r) = A(r) \exp[\, j\varphi(r)\,]$
- sum of optical fields:
  $A_1(r) \cos(\varphi_1(r) - \omega t) + A_2(r) \cos(\varphi_2(r) - \omega t) + \cdots$
  $= ?$
- in phasor arithmetic:
  $A_1(r) \exp[\, j\varphi_1(r)\,] + A_2(r) \exp[\, j\varphi_2(r)\,] + \cdots$
  $= U_{\text{total}}(r)$
- optical field (if needed):
  $u_{\text{total}}(r, t) = \mathrm{Re}\{U_{\text{total}}(r)\, e^{-j\omega t}\}$

The advantage of phasor arithmetic becomes clear when we sum several optical fields. When working with functions $u(r, t)$, it is difficult to find the amplitude of their sum. On the other hand, it is easy with phasor arithmetic: we just calculate several phasors at a single point in space and sum them. The resulting complex number encodes both amplitude and phase of the total optical field. Please note that a phasor is not a physical quantity! If we wanted the optical field, i.e. the function $u(r, t)$, we should multiply a phasor with $\exp(-j\omega t)$ and get the real part.
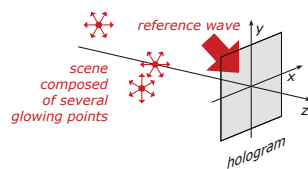
Calculation of a simple hologram

Thus, if we want to calculate a hologram (assume it is in the plane $z = 0$), we should follow these steps.

Hologram recording simulation

- assume hologram in the plane $z = 0$
- calculation of a hologram of a synthetic scene: for every point $(x, y, 0)$ of the hologram:
  - get the complex amplitude U_{obj} of the object wave at $(x, y, 0)$
  - get the complex amplitude U_{ref} of the reference wave at $(x, y, 0)$
  - calculate captured intensity at $(x, y, 0)$
    $I(x, y, 0) = |U_{obj} + U_{ref}|^2$

As an example, let us calculate a hologram of an object composed of three point light sources, see the image. This calculation is the "hello world" algorithm of computer generated display holography. Please note that many advanced algorithms of computer generated display holography just optimize it, so in the and, the algorithm we are going to implement is quite important.

Computer generated hologram of a point cloud

- the simplest algorithm in CGH
- basic building block of advanced algorithms of computer generated display holography
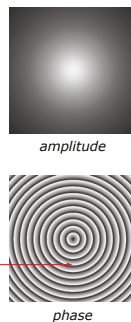

First of all, we will need to calculate the phasor of the optical field created by a point light source. A point light source creates a diverging spherical wave, i.e. the phasor at a point **x** just depends on distance r between **x** and the point light source location. In fact, the function $U(r)$ we have used until now is the phasor of a spherical wave, where A is its amplitude in the unit distance and φ is an arbitrary initial phase. If we are not interested in the initial phase, we can simply set $\varphi = 0$. Please note that surfaces of constant phase, $\arg(U(r)) = const.$, are concentric spheres

- spherical wave
  - light emitted by a point light source
  - r: distance from the light source

  - complex amplitude:
    $$U(r) = \frac{A}{r} \exp(j[kr + \varphi])$$

  - locally resembles a plane in a big distance
  - rays: "directions perpendicular to wavefronts"

amplitude

wavefronts: surfaces of constant phase

phase

separated λ apart. These surfaces are called wavefronts. Is some simple situations such as this one, we can define light rays as lines perpendicular to wavefronts. Also note that when r is large, the wavefronts locally look like planes, which means that rays are almost parallel. Also note that when r is large, the amplitude is locally almost constant.

- plane wave
  - light emitted by a point light source located in a direction −**n** far away, |**n**| = 1
  - **n** is a direction of light propagation and the normal vector of the wavefronts
  - point in space **x** = (x, y, z)
  - wavefront plane equation
    n · **x** = *const.*
  - wavefronts separation λ
  - complex amplitude:
    $U(\mathbf{x}) = A \exp(j[k\,\mathbf{n} \cdot \mathbf{x} + \varphi])$
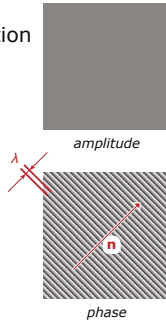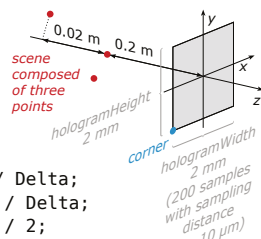
amplitude

phase

We can thus extend this approximation and define the plane wave, i.e. the optical field where wavefronts are perfect planes separated by λ and the amplitude is constant. The plane wave travels along direction **n**. In computer graphics, such light is called *directional*. As we know equation for a wavefront (**n** · **x** = *const.*), we can easily write the phasor definition.
Please note that plane waves are often used as reference or illumination waves.

Really unoptimized Matlab (Octave) code Initialization

```
lambda       = 532e-9;
hologramHeight = 2e-3;
hologramWidth  = 2e-3;
hologramZ = 0;
Delta        = 10e-6;
samplesX     = hologramWidth / Delta;
samplesY     = hologramHeight / Delta;
cornerX      = -hologramWidth / 2;
cornerY      = -hologramHeight / 2;
points       = [        0,           0, -0.2;
        -hologramWidth/4,-hologramHeight/4, -0.2;
         hologramWidth/4, hologramHeight/4, -0.22];
```

scene composed of three points

hologramHeight 2 mm

hologramWidth 2 mm (200 samples with sampling distance Δ = 10 μm)

corner

Following slides show a straightforward, but *really unoptimized code* for Matlab or Octave that calculates a hologram according to the image. You can find its full version (including displaying results, etc.) in the supplementary archive. You will find there an optimized version too that is more suitable for experiments as it is much faster, as well as some additional scripts. Also check http://holo.zcu.cz for more scripts.
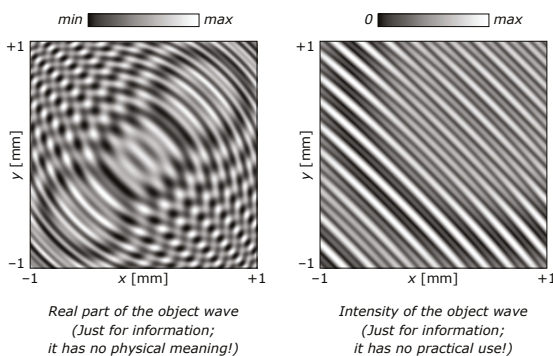Please note that all dimensions are in meters. The parameters were chosen such that the results are easy to understand.

Object wave calculation

```
k = 2*pi/lambda;
objectWave = zeros(samplesY, samplesX);
for s = 1:rows(points)
  for column = 1:samplesX
    for row = 1:samplesY
      x = (column-1) * Delta + cornerX;
      y = (row-1)    * Delta + cornerY;
      r = sqrt((x         - points(s, 1))^2 ...
            + (y          - points(s, 2))^2 ...
            + (hologramZ - points(s, 3))^2);
      objectWave(row,column) += exp(1i*k*r) / r;
    end
  end
end
```

The first step is the object wave calculation. The cycles iterate over all point light sources, rows and columns of the digital hologram. At each point, the phasor of the spherical wave is calculated, and results are accumulated in the 2-D array objectWave.

---

*Real part of the object wave
(Just for information;
it has no physical meaning!)*

*Intensity of the object wave
(Just for information;
it has no practical use!)*

Here is the result. Please note that both images are for information only. The left one, the real part of the phasor, has no physical meaning, as a phasor is not a physical quantity. The right one, the intensity, would appear if we illuminated a film with the object light only. It is physically correct, but it has no practical use – we need to add the reference wave before recording in order to make a hologram.

Computer generated display holography (EG2017 Tutorial Notes)   page 18/42

Reference wave calculation

```
alpha = 90   * pi/180;
beta  = 90.5 * pi/180;
nX = cos(alpha); nY = cos(beta);
nZ = sqrt(1 - nX^2 - nY^2);
refAmplitude = max(max(abs(objectWave)));
referenceWave = zeros(samplesY, samplesX);
for column = 1:samplesX
  for row = 1:samplesY
    x = (column-1) * Delta + cornerX;
    y = (row-1)    * Delta + cornerY;
    referenceWave(row,column) = refAmplitude * ...
        exp(1i*k*(x*nX + y*nY + hologramZ*nZ));
  end
end
```
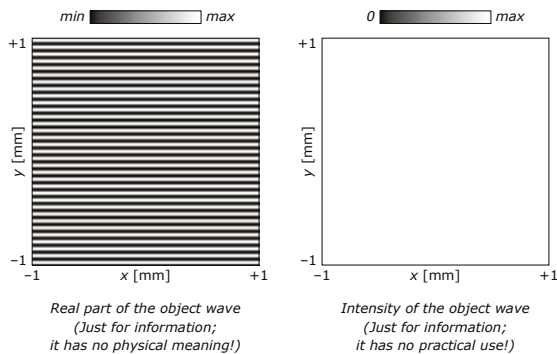
Real part of the object wave
(Just for information;
it has no physical meaning!)

Intensity of the object wave
(Just for information;
it has no practical use!)

Hologram calculation

```
optField = objectWave + referenceWave;
hologram = optField .* conj(optField);
```


The hologram (intensity picture)

Computer generated hologram of a 3-D scene

6144 × 6144 pixels
Size 4,3 × 4,3 cm²
(resolution 3600 dpi
~ pixel size 7 µm)
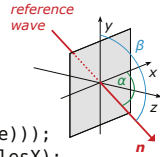
We will use a plane wave for the reference wave, as plane waves have some nice properties, for example they usually provide cleaner holograms. Please note that the direction vector **n** is defined using direction cosines, i.e. cosines of angles of **n** with axes *x* and *y*. We see that the reference wave travels almost in the direction of the axis *z*, i.e. it illuminates the hologram almost at normal incidence. This is necessary in our situation – sampling distance of the hologram plane is quite coarse (Δ = 10 µm) and steeper angles would easily cause aliasing. Also note that amplitude of the reference wave is chosen as the maximum amplitude of the object wave. If the reference wave amplitude is bigger, the hologram will be less noisy, but the image will be dimmer; and vice versa.

Here is the result. Please note that the real part of the phasor is composed of horizontal stripes, which indicates a plane wave with direction vector in the *yz* plane. Also note that the intensity image is completely flat, which indicates that the amplitude is constant.

Finally, we just add the object wave and the reference wave together and calculate the intensity. The result is the hologram.
Easy, isnt't it?

Here is an example of a computer generated display hologram of a 3-D scene with diffuse surfaces. You cannot see anything there – it looks like noise. Thus, we need some methods to verify if the hologram is correct.
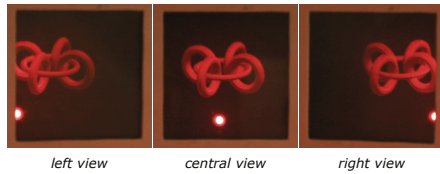
Hologram reconstruction

The most obvious way to test a hologram is to "print" it somehow and to illuminate it with a copy of the reference light. It is indeed possible, but unfortunately, the technologies such as laser lithography or e-beam lithography are quite expensive and difficult to operate. For details on e-beam lithography see [Verheijen], for laser lithography see [Nakahara].

For first experiments, it is possible to use an imagesetter (a device that was very common in printing industry) or even a laser printer. Do not expect any breathtaking results, though.

Please note that "holographic printers" that make an actual hologram (see [Bjelkhagen13, Saxby]) do not print the interference pattern pixel by pixel; they are based on different principle and the diffraction pattern is created in the same way as in classical holography, i.e. using laser light interference. See slide 109.

While "holographic print" is definitely nice, the holy grail of computer generated display holography is an electronic holographic display. Such a display is usually a common microdisplay, such as the one found in a beamer, that has very fine pixels. Such microdisplays are often called spatial light modulators (SLMs). Currently the best one offer diagonal about 40 mm and pixel size about 4 μm, which means that both image and viewing angle are too small. Thus, additional tricks are employed to make the display better. We will cover electronic displays later in the tutorial.

Finally, it is possible to simulate the hologram reconstruction process numerically. This is useful for evaluation purposes, at least until much better spatial light modulators become available. But what is important: the numerical simulation of hologram reconstruction is the second basic tool employed in algorithms of computer generated display holography.

How to "print" a calculated hologram?

- electron beam lithography – very expensive
  - 0.05 μm details ⇒ diffraction up to 90°
  - size up to ~ 5 × 5 cm², recording 1 mm²/min
- laser lithography – expensive
  - 1 μm details ⇒ diffraction up to 20°
  - size up to ~ 20 × 20 cm², recording 4 mm²/min

Hologram by K. Matsushima

left view   central view   right view

- imagesetter
  - 10 μm details
    ⇒ diffraction up to 2°
  - price ~ 5 € per A4
- laser printer
  - 100 μm details
    ⇒ diffraction up to 0.5°

- "holographic printers" do not usually print the calculated pattern

Hologram by I. Hanák, M. Janda

Electronic "holographic display"?

- microdisplays = spatial light modulators
- transmissive or reflective
- size up to 40 mm diagonal
- resolution up to 8K (7680 × 4320 pixels)
- pixel size down to ~ 4 μm
- usually LCD (liquid crystal display), DMD (digital micromirror device) or LCOS (liquid crystal on silicon)

- small size, small diffraction angle (up to 5°)
  - tiling, multiplexing, additional optics to improve performance

---

Numerical simulation of hologram reconstruction

- useful for evaluation purposes
- the second basic building block of advanced algorithms of computer generated display holography
- allows to observe the real image
- to simulate the virtual image, it is necessary to add a lens simulation (easy, but not covered in the tutorial)

Real image calculation
- illuminate the hologram "from behind"– each point of the hologram becomes a light source
  $U(x, y, z_{holo}) = hologram(x, y) \times U_{illum}(x, y, z_{holo})$
- place "a target screen" somewhere
- calculate light from the hologram on the target screen
- constructive interference creates the real image
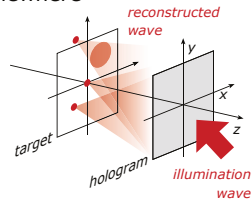
The hologram can be reconstructed in such a way that either virtual or real image is perfect. To reconstruct the virtual image, it is necessary to illuminate the hologram with the exact replica of the reference wave; the observer is located on the opposite side of the hologram than the illumination wave source, see slide 32. The observer has to be equipped with an ordinary imaging system composed of a lens and a photosensitive layer, such as a human eye (it has the eye lens and the retina). Unfortunately, we must simulate this optical system in numerical reconstruction in order to "see" the virtual image. It is not hard, but for simplicity it is not covered in this tutorial. Anyway, check sample scripts in the supplementary material to see how lens simulation works.

Here we are going to simulate the real image formation, see the image. The hologram has to be illuminated with light rays that are exact opposite to the reference light (compare this slide with slides 51 and 57). If we want to avoid propagation in the direction $-z$, we can imagine we flip the hologram, thus $\theta_{ill} = -\theta_{ref}$. If we substitute $m = -1$ to the sine-theta equation, we get

$$\sin \theta_{out} = m\, \lambda_{ill} / \lambda_{ref}\, (\sin \theta_{obj} - \sin \theta_{ref}) + \sin \theta_{ill} = (-1)(\sin \theta_{obj} - \sin \theta_{ref}) - \sin \theta_{ref} = -\sin \theta_{obj}$$

Thus, object rays are reversed as well and the real image is formed at the original location of the object. If we put a screen (e.g., a sheet of paper) there, real image manifests itself as high light intensity. Please note that a point of the real image that lies in the plane of the target screen is sharp, while point out of this plane looks like a blurry spot.

The idea of numerical hologram reconstruction is easy. If we illuminate the hologram, each of its points becomes a point light source with complex amplitude

$$U(x, y, z_{holo}) = hologram(x, y) \times U_{illum}(x, y, z_{holo}),$$

where $hologram(x, y)$ is the transmittance of the hologram at a point (x, y, z_{holo}) and $U_{illum}(x, y, z_{holo})$ is the phasor of the illumination wave at this point.

- sum of spherical waves
  – phasor of a spherical wave
    $K(x, y, z) = A \exp(\mathrm{j}kr) / r$
    $r = [(x - x_C)^2 + (y - y_C)^2 + (z - z_C)^2]^{1/2}$
  – each has its own amplitude A
  – their origins (x_C, y_C, z_C) are in the plane of the hologram, $z_C = z_{holo}$
  – the screen is usually parallel to the hologram in $z = z_{target}$
- it is possible to use convolution
  $U(x, y, z_{target}) = U(x, y, z_{holo}) \otimes K(x, y, z_{holo} - z_{target})$
  kernel

Thus, we just need to take many point light sources and add them together, exactly as we did in the object wave calculation on slide 55. Such a calculation would be, however, very slow. If we take into account that the hologram and the target screen are parallel to each other and the point light sources are uniformly distributed, we can use convolution for the calculation.

Please note that the function K (phasor of a spherical wave) in the convolution is usually called _the convolution kernel_.

- convolution calculation using the Fourier transform
  $$U(x, y, z_{\text{screen}}) = \mathcal{F}^{-1}\{\mathcal{F}\{U(x, y, z_{\text{holo}})\} \times \mathcal{F}\{K(x, y, z_{\text{holo}} - z_{\text{target}})\}\}$$
- efficient numerical algorithm: the fast Fourier transform (FFT)
- convolution calculation using FFT requires some additional steps

The reason to introduce convolution is simple: it can be calculated using the Fourier transform, and there is an algorithm for its fast calculation, see for example [Smith]. Please note that the actual convolution calculation using the fast Fourier transform (FFT) requires some minor additional steps covered on the next slide. If they are not applied, artifacts stemming from cyclic nature of FFT appear.

- to calculate $T = S \otimes K$
  S: source, $M_X \times M_Y$ samples
  T: target, $N_X \times N_Y$ samples
  K: kernel (spherical w. phasor), $C_X \times C_Y$ samples

1. pick any $C_X \geq M_X + N_X - 1$,
             $C_Y \geq M_Y + N_Y - 1$
2. add zero samples to S
   to get size $C_X \times C_Y$ *(zero padding)*
3. calculate $C_X \times C_Y$ samples of K
4. calculate
   `tmp = ifft2(fft2(padS) .* fft2(K))`
5. T = first $N_X \times N_Y$ samples from tmp
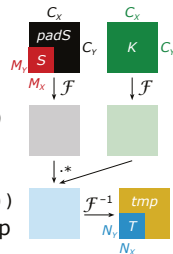
Comment: the numbers C_X, C_Y are usually selected in order to make FFT very fast, for example a power of 2, see [Lobaz11].

Actual code for hologram propagation

```
propagZ = 200e-3;
cX = 2*samplesX-1; cY = 2*samplesY-1;
auxX = ((0:cX-1) - (samplesX-1)) * Delta;
auxY = ((0:cY-1) - (samplesY-1)) * Delta;
[auxXX, auxYY] = meshgrid(auxX, auxY);
padS = zeros(cY, cX);
padS(1:samplesY, 1:samplesX) = ...
          hologram .* conj(referenceWave);
r = sqrt(auxXX.^2 + auxYY.^2 + propagZ^2);
kernel = k*1i*propagZ*exp(1i*k*r) ./ r.^2;
kernel = circshift(kernel, ...
          [-samplesY + 1, -samplesX + 1]);
tmp = ifft2(fft2(padS) .* fft2(kernel));
target = tmp(1:samplesY, 1:samplesX);
```

1. initialization
2. zero padding & illumination
3. kernel calculation
4.
5.

Here you can see the actual implementation. Please note that the reference wave reversal necessary for the real image formation is calculated as the complex conjugate, see the last red line (step 2).
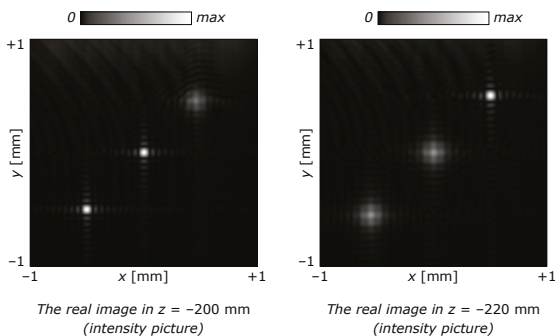
The real image in z = −200 mm (intensity picture)   *The real image in z = −220 mm (intensity picture)*

Result of the real image reconstruction in planes $z = -200$ mm and $z = -220$ mm. Compare them with the setup depicted on slide 65. Please note these images are intensity pictures, i.e. they are calculated as `target .* conj(target)`

- strictly speaking, we should use a proper propagation kernel instead of spherical waves
- Rayleigh-Sommerfeld propagation kernel
  - derived from the Maxwell's equations

$$K_{RS}(x, y, z) = -\frac{1}{2\pi}\left(jk - \frac{1}{r}\right)\frac{z}{r}\underbrace{\frac{\exp(jkr)}{r}}$$

$$\underbrace{\qquad\qquad}_{\substack{\text{additional factor}\\ \text{(can be often approximated or neglected)}}}\quad \underset{\substack{\text{spherical}\\\text{wave}}}{}$$

- Fresnel propagation kernel
  - follows from the Taylor approximation of r

$$K_{FS}(x, y, z) = -\frac{\exp(jkz)}{j\lambda z}\exp\left(jk\frac{x^2 + y^2}{2z}\right)$$

  - nice numerical properties ⇒ often employed

Strictly speaking, using the spherical wave as the convolution kernel for light propagation simulation is not exact. Instead, we should use for example the function K_{RS} that is formally derived from the Maxwell's equations, i.e., it correctly describes electromagnetic field (within some constraints). See [Goodman] for details. On the other hand, the function K_{RS} is in fact the phasor of a spherical wave multiplied by some additional factor. If we are interested just in relative intensity of the resulting light, simplifying or omitting this factor is usually not harmful.

It should be noted that some other functions are sometimes used in light propagation calculation, such as the Fresnel approximation K_{FS}, see [Goodman] for details. There are several reasons why this approximation is preferred sometimes: it is separable (see slide 82) and it has better numerical properties than function K_{RS}, see [Lobaz15].

Are the calculations correct?
- physical hologram observation note various artifacts:
  - blurry ghost images: unwanted diffraction orders
  - speckle noise: property of coherent llumination of diffuse surfaces
  - unrealistic contrast: weakness of the generation algorithm
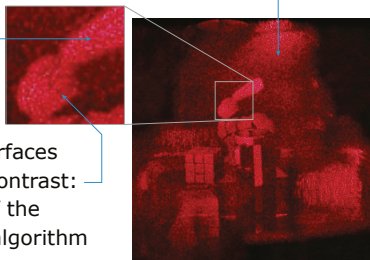
Is the calculation correct?

Before we proceed further, it is worth emphasising that coherent light used in holography has uncommon properties. If we look at a physically "printed" and properly illuminated hologram, we can spot several artifacts. Some of them (ghost image, speckle noise) are consequence of coherent illumination, some of them (scene contrast) are weakness of the calculation algorithm.

- simulated hologram observation: note various artifacts:
  - speckle noise: physically correct
  - 0th diffraction order: weakness of the simulation setup, but physically correct
  - replicas of the image: error of the algorithm
- it is not easy for a beginner to distinguish errors in algorithm from a physical phenomenon
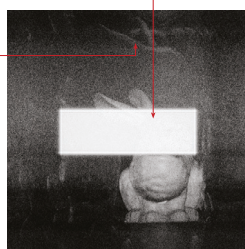
If we look at numerical simulation of the hologram simulation, we can see some artifacts that are physically correct consequence of coherent light illumination. Now, however, we can see an artifact that is actually an error of the calculation algorithm. For a beginner, it is often difficult to tell which artifact is a programming error and which would be replicated in real hologram observation as well.

Aliasing error
- difficult to recognize and fix for a beginner

real part of $K_{RS}(x, y, 200\ mm)$ *real part of $K_{RS}(x, y, 20\ mm)$*
(sampling distance 10 µm, λ = 532 nm)

The most common error that appears in digital holography calculations is aliasing. That is, frequency of the phasor we are working with exceeds the Nyquist-Shannon limit, i.e. one half of the sampling frequency (see [Smith]). As an example, the images show the real part of the function $K_{RS}(x, y, z_0)$ for different values z_0. Recall that the real part of the function $K_{RS}()$ should look like a set of concentric circles (the additional factor depicted on slide 71 does not change this). We can see that for small z_0, aliasing appears. Presence of aliasing in a calculation can be harmless, but often it has some negative impact on the result. Anyway, it is always welcome *to know* if aliasing appears. If we are not sure if it is dangerous, we can temporarily modify some calculation parameters so that it avoids aliasing.

Local frequency
- $\sin(2\pi fx)$ – frequency $f = \partial(fx) / \partial x$
- the most oscillatory part of $K_{RS}(x, y, z)$:
  $\exp(j\ 2\pi r/\lambda) = \exp(j\ 2\pi g)$
  $r = (x^2 + y^2 + z^2)^{1/2}$
- by analogy: local frequency

$$lf_X = \frac{\partial g}{\partial x} = \frac{1}{\lambda}\frac{\partial r}{\partial x} = \frac{x}{\lambda r} \qquad lf_Y = \frac{\partial g}{\partial y} = \frac{1}{\lambda}\frac{\partial r}{\partial y} = \frac{y}{\lambda r}$$

- due to sampling theorem, it must hold
  $lf_X \leq 0.5 \times$ sampling frequency   (the same for lf_Y)
- big x, y or small z leads to high local frequency
- usually safe to set $K_{RS}(x, y, z) = 0$
  at points where the local frequency is too high

The most important tool in aliasing detection is the local frequency analysis. Recall that a frequency f of the function $\sin(2\pi fx)$ can be symbolically written as the derivative of the sine argument (fx). In the same way, we can say that a function $\sin(2\pi g(x))$ has "spatially varying frequency" $\partial g(x) / \partial x$. We call this "the local frequency", see [Goodman] for details. As the most oscillatory part of the function $K_{RS}()$ is the complex exponential, we can easily derive its local frequencies in both x and y directions and compare them with the sampling frequency before calculation.

It is worth noting that *some algorithms* (not all!) work well if they use a simple anti-aliasing technique: they set $K_{RS}(x, y, z_0) = 0$ (or other function involved in the calculation) at points (x, y, z_0) rather then under-sample it, see for example [Matsushima09].

Angular spectrum decomposition
- propagation using convolution:
  $U(x, y, z_{screen}) =$
  $\quad \mathcal{F}^{-1}\big\{\mathcal{F}\{U(x, y, z_{holo})\} \times \mathcal{F}\{K(x, y, z_{holo} - z_{target})\}\big\}$
- it holds
  $\mathcal{F}\{K_{RS}(x, y, z)\} = H_{RS}(f_X, f_Y, z) =$
  $\qquad\qquad \exp[j\ 2\pi z(\lambda^{-2} - f_x^2 - f_y^2)^{1/2}]$
- transfer function H_{RS} has low local frequency for
  small $z \Rightarrow$ allows propagation calculation for small z
- usually safe to set $H_{RS}(f_x, f_y, z) = 0$ if local
  frequency is too high
- also allows to calculate propagation between
  non-parallel planes

Angular spectrum decomposition

The last very important tool of digital holography is *the angular spectrum decomposition*, see [Goodman], that works with optical field in the frequency domain. It can be used for many purposes, for example for light propagation calculation between non-parallel planes (not covered in the tutorial). Another use is described on the slide. Please note that we have used the Fourier transform for the convolution calculation (slides 65–69) as well, but it was just because there is the FFT algorithm. Here we actually calculate the Fourier transform (spectrum) of the optical field and modify it somehow. Thus, we should be very careful because some properties of the discrete Fourier transform calculated by FFT are very different from properties of the continuous Fourier transform that is the core of the angular spectrum decomposition.

SELECTED BUILDING BLOCKS OF ADVANCED CGDH ALGORITHMS

Now I am going to introduce various simple techniques often used in computer generated display holography. Please note I will not describe complete methods – these techniques are often combined together rather than used alone.

Scene replacement with a point cloud
- extraordinary number of points needed
- each point illuminates the whole hologram ⇒ slow
- does not handle surface visibility
- easy parallelization ⇒ fast for thousands of points

hologram

We actually know the first technique: to decompose a 3-D scene to a point cloud. It is used very often as it is very simple and the calculation can be easily parallelized.

Restricted area of contribution
- intermediate plane close to the point cloud
  – point contributes to a small area only ⇒ fast
- propagation of intermediate plane to the hologram ⇒ fast
- works well for shallow scenes only

1 *2*

intermediate plane   *hologram*
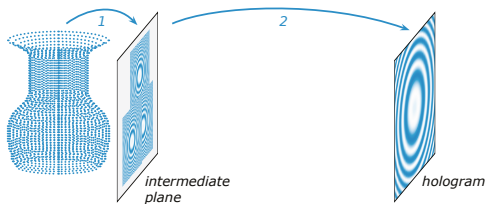
See e.g. [Shimobaba]. Any algorithm of light propagation between parallel planes can be used, for example the convolution one described on slides 65–69.

Multiple intermediate planes
- propagation of a point to the nearest intermediate plane
- propagation of each intermediate plane to the hologram, summation of results

2
1 *3* *4*
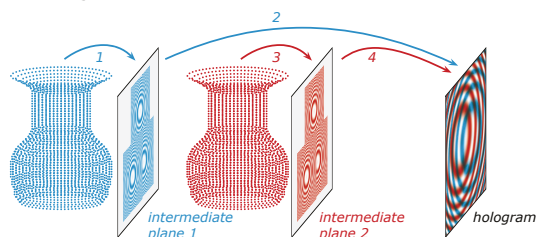
intermediate plane 1   *intermediate plane 2*   *hologram*

See for example [Okada]. As the intermediate planes are often close to each other, an angular spectrum based method of light propagation calculation is usually preferred.

Look-up tables (LUT)

- pre-calculation of propagation from $(0, 0, z_0)$ to the plane $z = 0 \Rightarrow$ base patterns for many z_0
- propagation from (x_0, y_0, z_0) to the plane $z = 0$ is a shifted base pattern
  $\Rightarrow$ "copy & paste" instead of the pattern calculation

base pattern

shifted base pattern from LUT
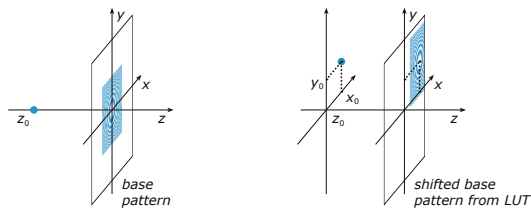
Look-up tables of various kinds are employed very often, for example this one [Kim]. For other look-up tables, see for example [Huang, Lee12, Nishitsuji].

Separable look-up tables

- Fresnel approximation of a spherical wave:

$$K_{FS}(x, y, z) = -\frac{\exp(jkz)}{j\lambda z} \exp\!\left(jk\frac{x^2}{2z}\right) \exp\!\left(jk\frac{y^2}{2z}\right)$$

- separable
  – only 1-D patterns have to be pre-calculated: $\exp(jkx^2/2z)$ and $\exp(jky^2/2z)$
  – look-up tables small (they are just 1-D)
- multiplication on the fly

The only other look-up table technique covered in this tutorial is based on separability of the Fresnel approximation of the convolution kernel $K_{RS}()$, as this property can be used in other ways as well. See e.g. [Ritter].

Other primitive objects

- point $\Rightarrow$ spherical wave
- line $\Rightarrow$ cylindrical wave
- line segment $\Rightarrow$ "clipped cylindrical wave"
- triangle $\Rightarrow$ wave can be also written in the closed form

- problem: flat triangle
  $\Rightarrow$ constant phase on the surface
  $\Rightarrow$ emits light in one direction
  $\Rightarrow$ hard to simulate diffuse surfaces
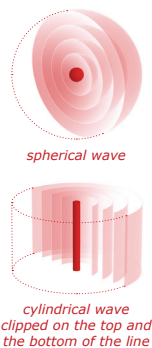
*spherical wave*

cylindrical wave clipped on the top and the bottom of the line

It is indeed possible to use other elementary object than a point light source, provided that their phasor can be written in the closed form. The most popular are lines (e.g. [Frère]) or triangles (e.g. [Koenig]). It should be, however, noted that a benefit of an analytic description of a phasor of a "large" primitive is quite limited. It is difficult to incorporate optical surface properties such as texture or reflectance model.

Single billboard

- object replacement with a flat image (billboard)
- propagation of the billboard to the hologram
- hidden surface elimination solved by computer graphics when rendering the image
- reconstructed scene is not 3-D

image (billboard)
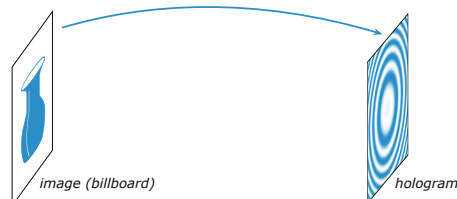
hologram

Maybe the most versatile primitive other than a single point is a rectangular grid of points. I call it "a billboard" here as it is similar to "billboards" or "sprites" used in computer graphics, where a complicated object is replaced by its image.
Here, each pixel of the image can be treated as a point light source, and we know that phasor of the whole grid can be easily calculated using convolution or other method of light propagation.

Multiple billboards
- scene replacement with several billboards
- propagation A → H, B → H, C → H, …, summation
- reconstructed scene is 3-D
- does not handle hidden surface elimination between billboards

Sweep propagation plane
- propagation A → B, masking, B → C, masking, …, C → H
- solves hidden surface elimination
- reconstructed scene is 3-D, but sliced (many billboards – better quality, slower)

Generalized sweep propagation plane
- scene split to rotated billboards
- optical rotation of billboard 1 to sweep plane 1 propagation $S_1 \rightarrow S_2$, rotation $S_2 \rightarrow B_2$, masking, rotation $B_2 \rightarrow S_2$, propagation $S_2 \rightarrow S_3$, …

- natural solution for triangular meshes
  – every textured triangle is a billboard
- good hidden surface elimination
  – however, it has the same problems as the painter's algorithm
- some steps can be simplified for faster calculation
- allows to calculate high resolution holograms in reasonable time (~ 4 Gpixels, several hours)

It follows we can replace a 3-D scene with a set of billboards, propagate each of them to the plane of the hologram and sum the results. Unfortunately, the algorithm cannot handle hidden surface elimination.

A simple modification was suggested by [Lohmann]. The first billboard (*A*) is propagated to the plane of the second billboard (*B*). It is assumed that some pixels of the image *B* are transparent, while others are "opaque", i.e. they represent point light sources. Such pixels replace just calculated phasors in the same way as with alpha blending in computer graphics. The modified phasor is propagated to the plane of the next billboard, and so on.

It is possible to extend the idea to billboards that are not parallel to each other, see e.g. [Leseberg, Matsushima02]. Here, the phasor of the billboard B_1 is first rotated to the plane S_1 that is in its vicinity and is parallel to the hologram plane. Please note that rotation of the optical field (in fact propagation between non-parallel planes) is not covered in this tutorial, details can be found, e.g., in the aforementioned references.
The optical field is propagated from the plane S_1 to the parallel plane S_2 that lies in the vicinity of the billboard B_2. The optical field is rotated to the plane of the billboard B_2 and some phasors are rewritten by pixels of the image in B_2. Now, we are in the same situation as in the beginning of the algorithm, and the procedure can be repeated: rotation to S_2, propagation to S_3, rotation to B_3, masking, …
The algorithm naturally works for any triangular mesh, as each triangle can be treated as a billboard. For further details, see e.g. [Matsushima14].

Basic ray casting

- point cloud rendering enhanced with ray casting for visibility testing
- extremely slow

point invisible ⇒ does not contribute

single hologram pixel

point visible ⇒ contributes
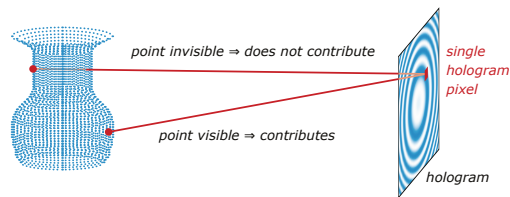
hologram

Other approach to hidden surface elimination is the ray casting approach. At each point of the hologram, a ray is cast to each point light source of the scene. If it does not intersect the scene, the point light source contributes to the point of the hologram.

As the hologram contains many "pixels" and the scene is composed of many points, it follows the algorithm is very slow. But again, it is a good starting point to more elaborate algorithms.

Billboards & ray casting

- scene decomposition to rectangular patches
- hologram decomposition to subholograms
- ray casting between billboards and subholograms only
- billboard → subhologram propagation fast

billboard visible ⇒ contributes

sub-hologram

scene approximation

billboard invisible ⇒ does not contribute
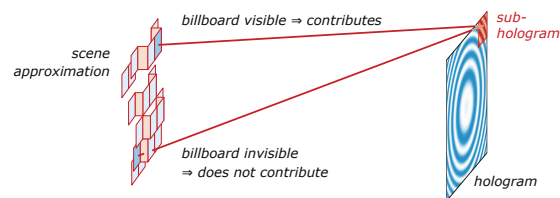
hologram

We can approximate the scene by many billboards composed of, for example, 32×32 point light sources. The hologram can be subdivided to "subholograms", areas composed by, for example, 32×32 samples. If the sampling distances of the billboards and the subholograms are the same, light propagation between them can be calculated fast. Moreover, it is possible to perform ray casting just between subholograms and billboards, see [Hanák].

Holographic stereogram

- for each subhologram
  - take a picture from its centre (i.e. record light intensity coming from various directions)
  - make a diffractive structure that replicates light rays – can be done using one Fourier transform
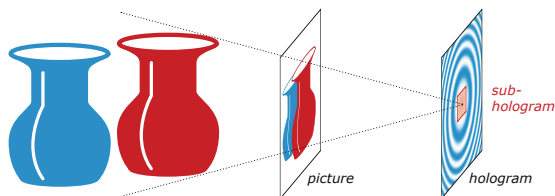
picture

sub-hologram

hologram

A completely different approach is based on images of the scene again. The hologram is decomposed to subholograms. A virtual camera is placed to the subhologram's center and a perspective image of the scene is rendered using common computer graphics techniques. Recall that each pixel of a perspective image contains information what light comes from a particular direction. We can thus imagine that the subhologram is illuminated by many plane waves, each pixel of the image describes one plane wave. Fortunately, summation of many plane waves can be efficiently calculated using single Fourier transform, see [Yatagai].

Horizontal parallax only rainbow hologram
- the image should be visible from "a window" only
- hologram decomposition to subholograms ("slits")
- determine which points of the scene contribute to the subhologram
- calculate the subhologram

these points contribute

subhologram

auxiliary plane

hologram

window for the observer

Most algorithms of computer generated display holography can be further accelerated if we sacrifice vertical parallax, i.e. an observer can watch a hologram from the left or from the right (and the image changes appropriately), but not from the top or from the bottom (the image either disappears or does not change).

As an example, we can demonstrate how to modify the basic algorithm for a hologram of a point cloud. First of all, we design the geometry, i.e. we put the scene and the hologram somewhere in space. Second, we design "a window" for the observer, i.e. a rectangular area where we expect the observer's eyes. For a horizontal parallax only hologram, the window usually degenerates to a thin rectangle or even a line. The hologram is again decomposed to subholograms, usually as wide as the hologram and not very tall. For each subhologram, we make a plane that crosses both the subhologram and the window. Only points of the scene that lie in the close vicinity of this plane contribute to the subhologram. Example algorithm that uses this technique is [Yoshikawa].

point A

hologram

subhologram A

auxiliary planes

window for the observer

point B

subhologram B

reference light

Hologram recording: *subhologram A records light from a point A only.*

virtual image

hologram

diffracted rays

virtual image

observer

virtual window

diffracted rays

hologram illumination

Hologram reconstruction: *both points can be seen in the window only*

To see how the hologram works, imagine a scene composed of two points only. Thus. the point *A* contributes to the subhologram *A* only, the same applies to the point and the subhologram *B*. When the hologram is illuminated by the replica of the reference wave, the subhologram *A* replicates light from the point *A* (the same for *B*). However, the light originates just in the subhologram area. Thus, an observer should be located at a specific position – behind "the window" – in order to see both points *A* and *B* at once.

virtual image

hologram

diffracted rays

virtual window

observer

virtual image

diffracted rays

hologram illumination – wrong wavelength
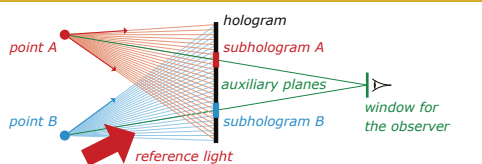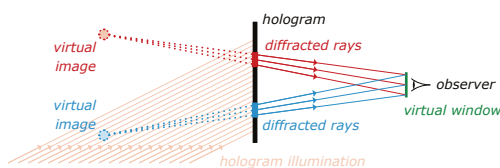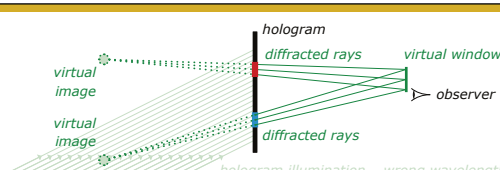
Hologram reconstruction with wrong (shorter) wavelength: *image shifts, virtual window shifts ⇒ observer does not see the image*

virtual images

hologram

diffracted rays

virtual window for green light

observer

virtual window for red light

diffracted rays

hologram illumination – white light
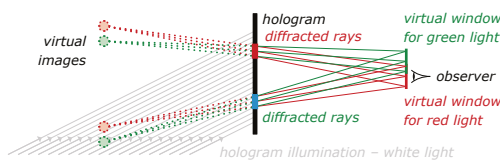
Hologram reconstruction with white light: *many virtual windows*

A horizontal parallax only hologram designed in this way has one striking property. Imagine we have calculated the hologram for red light $\lambda = 630$ nm (i.e. it was the wavelength of both the object and the reference waves). If we illuminate the hologram with the same red light, everything works as expected. What happens if we illuminate the hologram with green or blue illumination wave, e.g. $\lambda = 532$ nm or 460 nm? Recall from the grating equation that shorter wavelength is diffracted less. (Hint: Blue Barely Bends. Red Rotates Radically.) Thus, virtual images of the points are shifted from their original positions. Moreover, "the window" shifts as well, which means that the observer does not see anything. It is thus possible to illuminate the hologram with white light. In this case, many mutually shifted virtual images appear, but each colour is visible just from its virtual window. Thus, the observer sees a sharp monochromatic image. If he or she moves the head up or down, perspective of the image does not change, only colour does.

center left view:
perspective changes

top right view

no change in
perspective

center right view

no change in
perspective

bottom right view

Classical rainbow hologram
by Š. Němcová

Diffraction specific algorithms

- try to make a diffractive structure that creates desired light distribution in the target distance
- for example iterative:
  1. initialize "hologram" **H** to random numbers
  2. propagate to target distance: **H'** = propag(**H**)
  3. compare with desired light distribution **L**:
       if **H'** ≈ **L** then **DONE**
  4. modify **H'** so that it is closer to **L**
  5. propagate back to the plane of the hologram:
     H = backpropag(**H'**)
  6. apply constraints to **H** (e.g. 0≤transmittance≤1)
  7. go to step 2

This is why this type of hologram is often called "a rainbow hologram". The images on the slide show a classical rainbow hologram viewed from the top, bottom, left and right. Please note that this hologram had to be illuminated from the top, not from the bottom as in images on slides 93 and 94. This is why we see a red image when watching the hologram from the top, while the sketch on slide 94 shows "the red window" at the bottom. For more details on rainbow holograms, see [Benton].

Until now, all described algorithms more or less simulated optical setups used in classical holography. But there are algorithms that calculate a hologram in a completely different way. As they often aim to create some light distribution in space by diffraction, they are called *diffraction specific*. An example of such algorithm is shown on the slide, other example can be found in [Lucente]. Please note that similar algorithms are often used to optimize hologram performance or to design diffractive optical elements.

This list of techniques used in computer generated display holography is far from complete. I have tried to select techniques that could be explained during a single tutorial. Also note that I have selected the references that show the technique in a simple way, and I usually preferred original simple methods over more elaborate modern incarnations.

HOLOGRAPHIC DISPLAYS

In the rest of the tutorial, I am going to explain basic techniques for construction of a true holographic display that are currently used. Again, the techniques are often combined together.

Space-bandwidth product
- a hologram (diffraction pattern)
  – spatial extent (width × height)
  – frequency range f_{MIN} to f_{MAX} in X/Y axis
  – frequency f is related to diffraction angle θ:
    $\sin\theta = \lambda f$
- for pixelated holographic displays:
  spatial extent W × frequency range F
  $\qquad\qquad$ = 2 × number of samples N
  $\qquad\qquad$ (space-bandwidth product)
- example: $\lambda = 0.5$ µm, $\theta = 30°$, size 100 × 100 mm^2
  $N = 2W\sin(\theta)/\lambda \;\Rightarrow\; 200\,000 × 200\,000$ samples
- contemporary 8K displays: 8000 × 4000 samples

We know that a hologram is just a fine structure that diffracts light. Thus, any flat display such as LCD can be used to display such a pattern. The biggest obstacle in constructing a holographic (or other 3-D) display is the limited number of available pixels. In particular, there is a simple relation between the spatial and the angular extent of a pixelated display with fixed number of pixels, see [Benton]. In short, if pixels are big, the spatial extent of the display is big as well, but the diffraction angle is small; and vice versa. Example on the slide shows that a "large" display with "reasonable" viewing angle (i.e. the angular extent) requires extraordinary number of pixels. Current displays offer pixel counts several orders of magnitudes smaller.

This means that a particular flat display, such as LCD with resolution 4000 × 2000 pixels, can be either large with a small viewing zone, or small with a big viewing zone.

θ

display

viewing zone

- space-bandwidth product constant:
  – either large display,
    narrow viewing zone
  – or small display,
    wide viewing zone

---

negative (concave) lens

display

small virtual image of the display

wider viewing zone
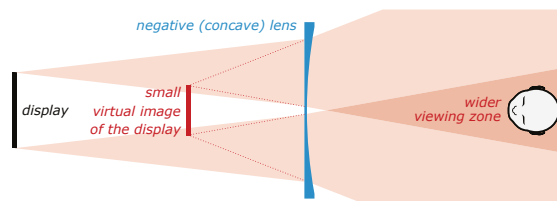
- use a lens to enlarge the angle?
  ⇒ virtual image of the display gets smaller
- space-bandwidth product is hard to cheat!

One may think for a while that the angular extent can be increased using some additional optics, for example a negative lens. It definitely makes the angular extent bigger – but the observer then sees a virtual image of the display, which is small. Thus, it is not easy to cheat the space-bandwidth product!

SLM based holographic display systems
- multiple spatial light modulators
  - planar layout – does not increase viewing angle
  - circular layout – increases horizontal view. angle
  - additional optical elements (e.g. one way mirror) for seamless stitching
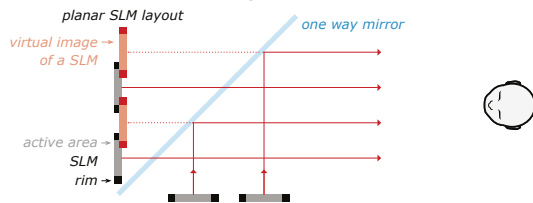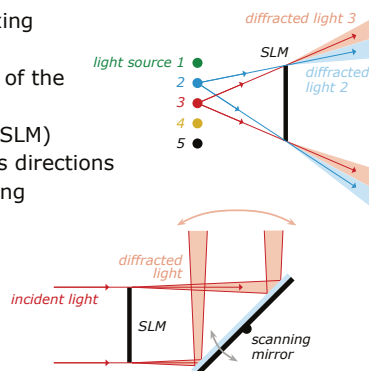
planar SLM layout

one way mirror

virtual image of a SLM

active area

SLM

rim

The most straightforward solution of the problem is to take many individual SLMs to increase pixel count. They can be arranged in various ways, for example in a plane or a cylindrical surface, see e.g. [Kurihara, Kozacki]. The biggest problem is proper alignment of individual SLMs and their seamless stitching.

- time multiplexing
  - sequential illumination of the spatial light modulator (SLM) from various directions
  - beam steering (horizontal, vertical or both)

diffracted light 3

SLM

light source 1
2
3
4
5

diffracted light 2

diffracted light

incident light

SLM

scanning mirror

An alternative approach exploits *persistence of vision* of a human eye – we see a flash of light much longer than its actual duration. See [Higashino, Kang] for examples of such displays. The biggest problem is that the SLM must provide very high frame rate.

- speckle field
  - diffuser changes phase randomly at each point
  - ⇒ display size = diffuser size (good),
    viewing angle = diffusing angle (very good),
    obscures image due to phase breakup (bad)
  - phase breakup can be compensated at SLM
  - ⇒ big display, big viewing angle
  - trades space-bandwidth product for contrast
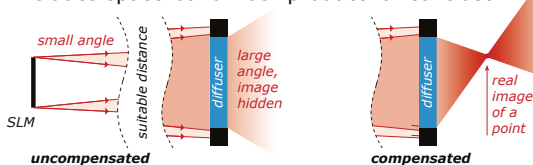
small angle

suitable distance

large angle, image hidden

diffuser

SLM

diffuser

real image of a point

uncompensated　　　　**compensated**

A completely different approach was recently described by [Yu]. A SLM illuminates a diffuser, for example a piece of ground glass, in such a way that each pixel of the SLM illuminates the whole area of the diffuser. As the diffuser changes phase of transmitted light randomly, the optical field created by the SLM is damaged and the observer behind the diffuser does not see any reasonable image. But the diffuser is constant in time, i.e. its phase characteristics can be measured and compensated by the SLM. Thus, it is possible to create an image that is as large as the diffuser, and its viewing angle is as large as the diffusing angle of the diffuser. The trick is that phase compensation cannot be done exactly with limited number of SLM pixels, i.e. there is a tradeoff between the space-bandwidth product and the image contrast and resolution.

Modulation types
- original holography: amplitude modulation
  (transmittance of a point of a hologram $\in [0, 1]$)
  – calculate phasors U_object, $U_\text{reference}$
  – transmittance $\propto |U_\text{object} + U_\text{reference}|^2$
  – attenuates light $\Rightarrow$ low brightness
- bipolar intensity: alternative amplitude modulation
  – transmittance $\propto \text{Re}\{U_\text{object}\, U_\text{reference}^*\}$
  – does not contain fringes that form unwanted diffraction

Until now, we have assumed that each pixel of the SLM (or each point of a hologram) is transparent, opaque or something in between. Such a pixel modulates light amplitude (and thus light intensity). It works well, but as the display absorbs some light, the image is quite dim [Goodman].

Please note that it is possible to calculate the pixel value in a way as we did, i.e., by calculation of the intensity of the sum of an object wave and a reference wave. Such a calculation contains the interference pattern of object points and the reference wave, as well as the mutual interference pattern of object points. The second pattern is not important for display purposes; in fact, it increases noise in the image. Thus, there are alternative approaches such as bipolar intensity [Lucente] that calculate just the pattern we want.

- bleached holograms: phase modulation
  (transmittance of a point of a hologram = $\exp(j\varphi)$)
  – $I = |U_\text{object} + U_\text{reference}|^2$
    transmittance $\propto \exp(jI)$
  – much brighter than amplitude hologram,
    more noise
- kinoform: alternative phase modulation
  – calculate the object wave only (complex values)
  – set amplitude = 1, keep phase
    transmittance = $U_\text{object}/ |U_\text{object}|$
  $\Rightarrow$ phase only modulation
  – similar properties as bleached holograms
  – physically (almost) impossible, calculation easy

It was recognized shortly after invention of holography that black-and-white pattern on a photographic film can be bleached, i.e., the film becomes transparent. But, the points that were black change phase of light in some way, and the points that were grey change it differently. Thus, the amplitude modulation was changed to phase modulation. Surprisingly, diffraction on such a phase pattern is very similar to diffraction on the original black-and-white pattern. The image created by diffraction is a bit noisier, but it is much brighter as the pattern does not absorb any light [Goodman]. Thus, SLMs that modify phase rather than amplitude are often preferred.

Again, there are other approaches to phase modulation. For example, if we calculate the object wave, set its amplitude to 1, keep the phase and "print" such a pattern, we get phase only modulation again. Such a pattern diffracts an illumination wave at normal incidence similarly as a proper hologram, see "kinoform" in [Goodman]. There are two key ideas. First, the object-reference waves interference used in holography is just a trick that allows us to capture phase of the object wave; in calculations, getting phase of the object wave is trivial.

Second: it is actually necessary to take the illumination wave and to change its phase and amplitude to match the original object wave. Kinoform changes just the phase.

Change of both amplitude and phase is called "full complex modulation". Unfortunately, we do not still know how to make it well. For some approaches, see for example "ROACH" in [Goodman] or [Ulusoy].
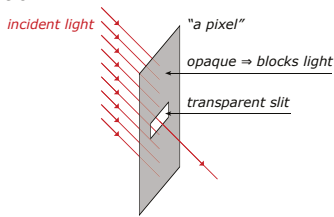
- full complex modulation
  – both amplitude and phase modulation at once
  – "recording" does not require reference wave
  – transmittance = $U_\text{object}/ U_\text{illumination}$
  – perfect image – exact copy of the object wave
  – we don't know how to make it well
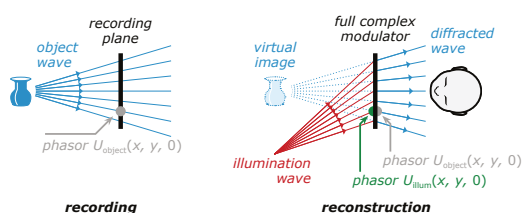
- phase detour
  - early approximation of the complex modulation
  - width of the slit modulates light intensity
    ⇒ amplitude modulation
  - position of the slit picks light with proper phase
    ⇒ phase modulation
  - dim image, requires very high resolution
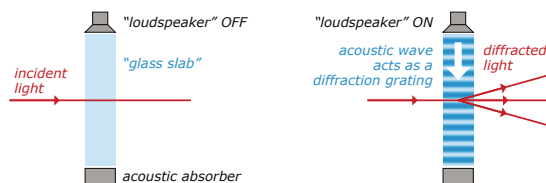 - other, more advanced versions exist

Acousto-optical modulator
- does not display the hologram "pixel by pixel"
- displays an elementary diffractive grating at once
- incident sound wave ⇒ frequencies in the grating
- limited to 1-D gratings
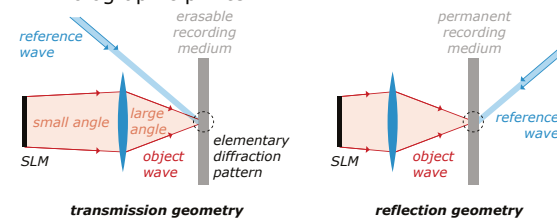 ⇒ suitable for horizontal parallax only displays

True holographic modulator
- elementary diffraction pattern – physically recorded hologram
- recording medium erasable ⇒ holographic display
- recording medium permanent
  ⇒ holographic printer

transmission geometry   **reflection geometry**

- transmission geometry: requires back illumination
  - more suitable for an enclosed device
- reflection geometry: requires front illumination
  - more suitable for prints
  - allows white light illumination and full colour reconstruction

transmission geometry   **reflection geometry**

Until now, we have assumed just "homogenous pixels" that simply change amplitude or phase of light. An early attempt to complex modulation – phase detour, see [Goodman] – shows that a single element of a digital hologram can have more complicated structure. Here, the element is opaque with an exception of small window. Its size modulates transmitted light amplitude (intensity), its position picks light of appropriate phase from the illumination wave. Please note that there are many other schemes, see for example [Lee78].

It follows that an individual hologram element (hogel, in the same way as a pixel is a picture element) could in principle modify light in a complicated way. A simple example of a physical implementation of a hogel is an acousto-optical modulator (AOM) that deflects (diffracts) incoming light in a single plane, say xz. See [Lucente] for a classical holographic display employing AOM's or [Smalley] for a modern approach.

It would be much better if a hogel was able to diffract light in both xz and yz planes, exactly as a hologram does. The simplest thing we can do is to take a holographic recording material, to partition it to small rectangular areas, and to make a hologram inside each of them. This is, in fact, the principle of a holographic printer, see [Bjelkhagen13, Saxby]. It is also possible to take a recording medium that can be erased, such as a photorefractive material; then, we can update the hologram once a while, and get a live holographic display, see [Blanche].
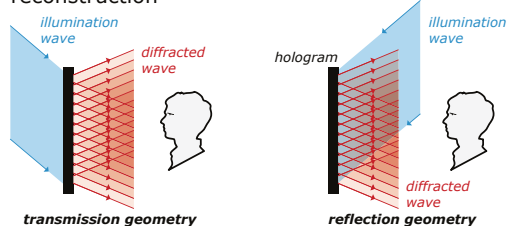
Please note that it is possible to design both setups for a transmission or a reflection hologram.

- examples of alternative hologram geometries

cylindrical hologram

illumin. wave

diffracted wave

virtual image

tabletop hologram

illumination wave

diffracted wave

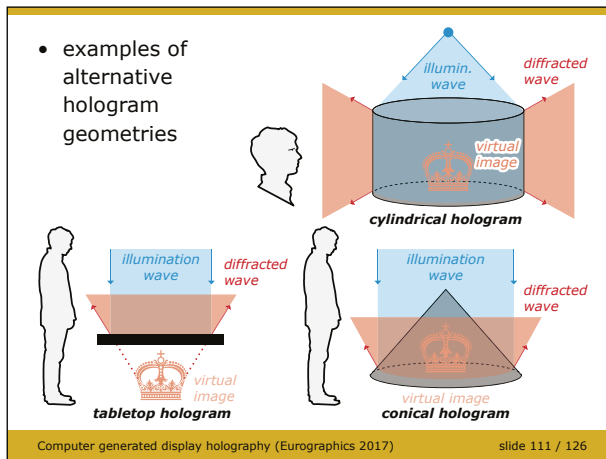virtual image

conical hologram

illumination wave

diffracted wave

virtual image

Finally, it is worth mentioning that idea of holography does not require just geometries I have described until now. Other geometries allow 360° viewing angle, are better suited for a real image display, etc. See for example [Benton, Saxby] for details.

Holographic displays in general

- pros:
  – offer great 3-D experience,
    hardly distinguishable from reality
- cons:
  – require high space-bandwidth product
  – hologram calculation is slow
  – require coherent light,
    which causes speckle noise
  – technology for both display and calculation
    is far from optimal

In the end, I should admit that while holography offers great viewing experience, it has many drawbacks. It seems that 3-D displays of the near future will be based on other than holographic principles. Anyway, as they have their drawbacks too, it is very likely that at least some ideas from holography will be employed there as well.
It is also worth repeating that holography has much more applications than just display technology.
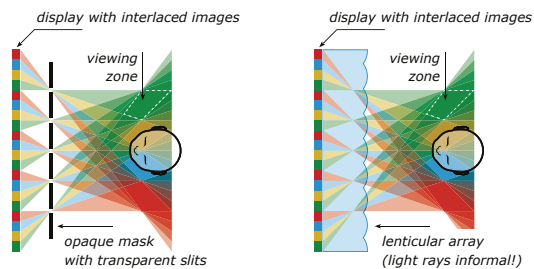
COMPETING TECHNOLOGIES

In the rest of the tutorial, I will briefly mention the most important competing technologies for 3-D display.

Light field displays
- parallax barrier displays, lenticular displays
  – established technology
  – problematic viewing outside viewing zones

The oldest ideas of autostereoscopic display construction are more than a century old – these are, in particular, the parallax barrier and the integral (lenticular) technology [Lueder, Okoshi]. While they can be thought as technologies that deliver correct images to the left and right eyes of the observer (i.e. autostereoscopic multiview displays), they can be also thought as simple light field displays.

- idea of "light field" L($\mathbf{x}$, $\mathbf{n}$):
  light ray from a point $\mathbf{x}$ (usually in $z = 0$)
  in a direction $\mathbf{n}$ has luminance L($\mathbf{x}$, $\mathbf{n}$)
  ⇒ light field is 4-D (at least)
- based on ray optics
  – very intuitive
  – easy to define light field that cannot exist
    (e.g. a single ray)
  – does not work well with small distances and
    angles that are required for high fidelity display
- it is not clear how to sample L($\mathbf{x}$, $\mathbf{n}$)
  ⇒ limited resolution and depth of field

While idea of light field is very old, it became "mainstream" in computer graphics due to [Levoy]. Naturally, both integral and parallax barrier displays sample a light field at pre-defined points, where finite number of rays is defined. As both types of displays are far from ultrarealistic (contrary to holography), it is clear that at least sampling of the light field is not sufficient. Here comes the first problem: it is not clear how to sample light field properly. In fact, there is nothing like "infinitesimally thin ray". Moreover, the wave nature of light imposes strict limits on what details are

actually possible in the light field, while the idea of the light field itself does not impose any limits.

The consequence is clear: if we try to make either integral of parallax barrier display "too fine", diffraction becomes important and the ray model of light is no longer valid.

Anyway, both technologies offer technically feasible autostereoscopic displays that can work very well if we have realistic expectations; see for example [Efrat, Takaki].

- multilayer (tensor) displays
  - attempts to approximate light field directly
  - ray 1 (luminance L_1) crosses A', B, C
    ray 2 (luminance L_2) crosses A, B, C'
    ... } system of equations
  - how to set transparency at points A, B, C, ...?
    $\Rightarrow$ solve the system of equations
  - combines well with other technologies, affordable
 - solution is only approximate
    $\Rightarrow$ ghosting
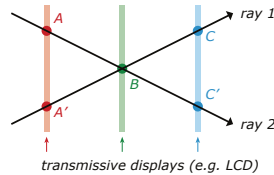
transmissive displays (e.g. LCD)

More direct approach to light field displays was proposed by [Lanman10, Wetzstein]. While it has its specific limitations too (for example depth of field is quite limited), it shows that original approaches are still possible in light field displays.

Multiplane displays

- decomposition of a scene to planar slices
- display slices on real or virtual places in space
- no vergence-accommodation conflict
- problems with hidden surface elimination
 $\Rightarrow$ usually requires user tracking
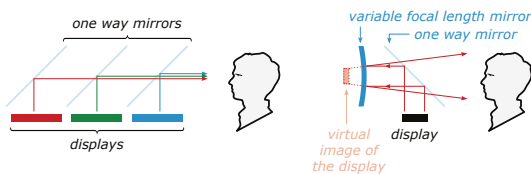
A common problem of multiview displays, such as common parallax barrier and lenticular ones, is the vergence-accommodation conflict. It means that the eye does not focus to a correct distance of a 3-D illusion, which causes visual discomfort. The visual perception can be improved by slicing the 3-D illusion, and displaying individual slices in various planes, either real or virtual [Ravikumar, Traub, Lueder]. Please note that it is assumed there is just one observer at a fixed location; otherwise, errors in occlusion between virtual scene objects appear.

Near eye display

- for augmented / virtual reality
- requires precise head tracking and fast response
- only small display area required
- should provide accommodation eye response
  - natural with holography
  - possible with light field displays
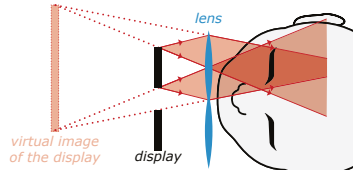
This, however, is not important in near eye displays. While I have been talking about autostereoscopic displays until now, systems for augmented and virtual reality should be mentioned as well. Current commodity near eye displays usually use a common flat microdisplay to show an image and additional optics to enlarge it and move it virtually far away from the viewer. However, if the illusion is supposed to be in a close distance from the observer, accommodation mismatch appears, as the virtual image of the microdisplay is still located far away from the viewer. Thus, a good microdisplay should provide proper accommodation response. It is naturally provided by a true holographic microdisplay; but other technologies such as light field displays can provide it as well [Lanman13].

RECOMMENDED READING

I have selected a few books that, I think, are suitable for beginners. The best approach to computer generated display holography is, I think, to have some hands-on experience with classical holography, to understand theory of classical holography and Fourier optics, and indeed to study digital techniques. In case of any questions, do not hesitate to contact me. Thank you for your attention!

General holography

- **Holographic Imaging**
  S. A. Benton, V. M. Bove Jr.; Wiley 2008
  excellent introduction to general holography and display holography, touches digital holography a bit

- **Optical Holography**
  R. J. Collier, C. B. Burckhardt, L. H. Lin; Academic Press 1971
  classic textbook, holography in depth; maybe not suitable as a first book on holography you read, but definitely worth reading after gaining some experience

---

Display holography

- **Three-Dimensional Imaging Techniques**
  T. Okoshi; Academic Press 1976
  considerations on 3-D imaging, both holography and integral imaging; still very relevant book

- **Practical Holography**
  G. Saxby, S. Zacharovas; CRC Press 2015
  a must for anyone making classical display holograms, also covers holographic printers

- **Ultra-Realistic Imaging**
  H. Bjelkhagen, D. Brotherton-Ratcliffe; CRC Press 2013
  full colour holography and holographic printing in depth

Fourier optics

- **Introduction to Fourier Optics**
  J. W. Goodman; Roberts and Company Publishers 2004
  classic textbook, diffraction and related phenomena including holography in (reasonable) depth; every digital holographer should have it at hand

- **Computational Fourier Optics: A MATLAB Tutorial**
  David G. Voelz; SPIE Press 2011
  nice and short introduction to the topic, works well as a supplement to the Goodman's book

---

Digital holography

- **Digital Holography**
  P. Picart, J.-C. Li; Wiley-ISTE 2012
- **Digital Holography and Digital Image Processing**
  L. Yaroslavsky; Springer 2004
- **Digital Holography and Wavefront Sensing**
  U. Schnars, C. Falldorf, J. Watson, W. Jüptner; Springer 2015
- **Introduction to Modern Digital Holography**
  T.-C. Poon; Cambridge University Press 2014

each book contains general introduction, then focuses on different aspects and applications

---

Journals

- Optics Express, Applied Optics, Optics Letters
  – *most CGDH articles is published here nowadays*
- Journal of the Optical Society of America A, Optical Engineering, Optics Communications, Journal of Display Technology
  – *worth checking regularly*

Conferences

- International Symposium on Display Holography
  – *mostly display holography, art and technology, both classical and digital*
- Practical Holography (SPIE)
  – *art and technology, both classical and digital*
- Digital Holography & 3-D Imaging (OSA)
  – *general digital holography, 3-D imaging*
- The Holography Conference (Reconnaissance)
  – *mostly security holography and packaging*

Bibliography

[Benton]   Benton, S. A., Bove, Jr., V. M. *Holographic Imaging.* Wiley, 2008. ISBN: 9780470224120.

[Blanche]   Blanche, P. A., et al. "Holographic three-dimensional telepresence using large-area photorefractive polymer". In: *Nature,* 468.7320 (03 November 2010), pp. 80–83. DOI: 10.1038/nature09521.

[Bjelkhagen13]   Bjelkhagen, H., Brotherton-Ratcliffe, D. *Ultra-Realistic Imaging: Advanced Techniques in Analogue and Digital Colour Holography.* Taylor & Francis, 2013. ISBN: 9781439827994.

[Bjelkhagen16]   Bjelkhagen, H. I., Lembessis, A., Sarakinos, A. "Ultra-realistic imaging and OptoClones". In: *Proc. SPIE 9771, Practical Holography XXX: Materials and Applications* (2016), pp. 977105–977105–8. DOI: 10.1117/12.2207524.

[Collier]   Collier, R., Burckhardt, C., Lin, L. *Optical holography.* Student editions. Academic Press, 1971. ISBN: 9780121810528.

[Efrat]   Efrat, N., Didyk, P., Foshey, M., Matusik, W., Levin, A. "Cinema 3D: Large Scale Automultiscopic Display". In: *ACM Trans. Graph.,* 35.4, Article 59 (July 2016). DOI: 10.1145/2897824.2925921.

[Favalora]   Favalora, G. E. et al. "100-million-voxel volumetric display". In: *Proc. SPIE 4712, Cockpit Displays IX: Displays for Defense Applications.* (August 28, 2002). DOI: 10.1117/12.480930.

[Frère]   Frère, C., Leseberg, D., Bryngdahl, O. "Computer-generated holograms of three-dimensional objects composed of line segments". In: *J. Opt. Soc. Am. A,* 3.5 (May 1986), pp. 726–730. DOI: 10.1364/JOSAA.3.000726.

[Gabor48]   Gabor, D. "A New Microscopic Principle". In: *Nature,* 161 (May 1948), pp. 777–778. DOI: 10.1038/161777a0.

[Gabor49]   Gabor, D. "Microscopy by Reconstructed Wave-Fronts". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* 197.1051 (1949), pp. 454–487. DOI: 10.1098/rspa.1949.0075.

[Hanák]   Hanák, I., Janda, M., Skala, V. "Detail-driven digital hologram generation". In: *Visual Computer,* 26 (January 2010), pp. 83–96. DOI: 10.1007/s00371-009-0378-5

[Higashino]   Higashino, Y., Tsuchioka, T., Matsushima, K. "Light-Source Switching Time-Division Multiplexing Holographic Display and Reduction of Degradation by Higher Order Diffraction Images". In: *Imaging and Applied Optics 2016,* OSA Technical Digest (online) (Optical Society of America, 2016), paper DM4E.1. DOI: 10.1364/DH.2016.DM4E.1

[Huang]   Huang, Y., Zhao, K., Yan, X., Pei, C., Jiang, X. "Phase-searching look-up table method for computer-generated holograms". In: *Journal of Electronic Imaging,* 23.4, p. 043013, 2014. DOI: 10.1117/1.JEI.23.4.043013.

[Jones]   Jones, A. et al. "Rendering for an Interactive 360 Degree Light Field Display". In: *ACM SIGGRAPH conference proceedings.* San Diego, CA, Aug. 2007. DOI: 10.1145/1275808.1276427.

[Kang]   Kang, H., Lim, S., Stoykova, E., Kim, Y., Hong, J., Hong, S. "Time Angular Multiplexing Holographic Display System". In: *Imaging and Applied Optics 2016,* OSA Technical Digest (online) (Optical Society of America, 2016), paper DM3E.5. DOI: 10.1364/DH.2016.DM3E.5.

[Kim]	Kim, S.-C., Kim, E.-S. "Effective generation of digital holograms of three-dimensional objects using a novel look-up table method". In: *Appl. Opt.,* 47.19 (Jul 2008), pp. D55–D62. DOI: 10.1364/AO.47.000D55.
[Koenig]	Koenig, M., Deussen, O., Strothotte, T. "Texture-based hologram generation using triangles". In: *Proc. SPIE 4296, Practical Holography XV and Holographic Materials VII*, 2001, pp. 1–8. DOI:10.1117/12.429440.
[Kozacki]	Kozacki, T., Kujawińska, M., Finke, G., Hennelly, B., Pandey, N. "Extended viewing angle holographic display system with tilted SLMs in a circular configuration." In: *Appl. Opt.,* 51.11 (2012), pp. 1771–1780. DOI: 10.1364/AO.51.001771
[Kurihara]	Kurihara, T., Takaki, Y. "Improving viewing region of 4f optical system for holographic displays". In: *Opt. Express,* 19.18 (2011), pp. 17621–17631. DOI: 10.1364/OE.19.017621
[Lanman10]	Lanman, D., Hirsch, M., Kim, Y., Raskar. R. "Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization". In: *ACM SIGGRAPH Asia 2010 papers* (SIGGRAPH ASIA '10). DOI: 10.1145/1866158.1866164.
[Lanman13]	Lanman, D., Luebke, D. "Near-eye light field displays". In: *ACM SIGGRAPH 2013 Emerging Technologies* (SIGGRAPH '13). DOI: 10.1145/2503368.2503379.
[Lee78]	Lee, W. H. "Computer-Generated Holograms: Techniques and Applications". In: *Progress in Optics,* 16 (1978), pp. 119–232. DOI: 10.1016/S0079-6638(08)70072-6.
[Lee12]	Lee, S., Wey, H. C., Nam, D. K., Park, D. S., Kim, C. Y. "Fast hologram pattern generation by radial symmetric interpolation". In: *Proc. SPIE 8498, Optics and Photonics for Information Processing VI* (2012). DOI: 10.1117/12.928904.
[Leseberg]	Leseberg, D., Frère, C. "Computer-generated holograms of 3-D objects composed of tilted planar segments". In: *Appl. Opt.,* 27.14 (Jul 1988), pp. 3020–3024. DOI: 10.1364/AO.27.003020.
[Levoy]	Levoy, M., Hanrahan, P. "Light Field Rendering". In: *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (SIGGRAPH '96)*, pp. 31–42. DOI: 10.1145/237170.237199
[Lobaz11]	Lobaz, P. "Reference calculation of light propagation between parallel planes of different sizes and sampling rates". In: *Opt. Express,* 19.1 (Jan 2011), pp. 32–39. DOI: 10.1364/OE.19.000032.
[Lobaz15]	Lobaz, P., Vaněček, P. "Safe range of free space light propagation calculation in single precision". In: *Opt. Express,* 23.3 (Feb 2015), pp. 3260–3269. DOI: 10.1364/OE.23.003260.
[Lohmann]	Lohmann, A. W. "Three-dimensional properties of wave-fields". In: *Optik,* 51.2 (1978), pp. 105–117.
[Lucente]	Lucente, M. *Diffraction-specific fringe computation for electro-holography.* Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, Sept. 1994.
[Lueder]	Lueder, E. *3D Displays*. Wiley Series in Display Technology. Wiley, 2011. ISBN: 9781119963042.
[Maas]	Maass, U. *Device for displaying moving images in the background of a stage.* US Patent US 5865519 A. Publication date Feb 2, 1999.

[Matsushima02] Matsushima, K., Schimmel, H., Wyrowski, F. "New creation algorithm for digitally synthesized holograms in surface model by diffraction from tilted planes". In: *Proc. SPIE 4659, Practical Holography XVI and Holographic Materials VIII* (2002), pp. 53–60. DOI:10.1117/12.469291.

[Matsushima09] Matsushima, K., Shimobaba, T. "Band-Limited Angular Spectrum Method for Numerical Simulation of Free-Space Propagation in Far and Near Fields". In: *Opt. Express,* 17.22 (Oct. 2009), pp. 19662–19673.
DOI: 10.1364/OE.17.019662.

[Matsushima14] Matsushima, K., Nakamura, M., Nakahara, S. "Silhouette method for hidden surface removal in computer holography and its acceleration using the switch-back technique". In: *Opt. Express,* 22.20, pp. 24450–24465 (2014).
DOI: 10.1364/OE.22.024450.

[Nakahara] Nakahara, S., Matsushima, K. "Fabrication of Computer Generated Holograms Constituted from Sub-Micrometer Pixel for Wide Viewing Angle Using Laser Lithography". In: *Imaging and Applied Optics 2014*, OSA Technical Digest (online). DOI: 10.1364/AIO.2014.JTu4A.10.

[Nishitsuji] Nishitsuji, T., Shimobaba, T., Kakue, T., Ito, T. "Fast calculation of computer-generated hologram using run-length encoding based recurrence relation". In: *Opt. Express,* 23.8, pp. 9852–9857, Apr 2015.
DOI: 10.1364/OE.23.009852.

[Ochiai] Ochiai, Y., et al. "Fairy Lights in Femtoseconds: Aerial and Volumetric Graphics Rendered by Focused Femtosecond Laser Combined with Computational Holographic Fields". In: *ACM Trans. Graph.,* 35.2 (Feb 2016), 17:1–17:14. ISSN: 0730-0301. DOI: 10.1145/2850414.

[Okada] Okada, N., et al. "Fast calculation of a computer-generated hologram for RGB and depth images using a wavefront recording plane method". In: *Photonics Letters of Poland,* 6.3 (Sep 2014), pp. 90–92.

[Okoshi] Okoshi, T. *Three-dimensional imaging techniques.* Academic Press, 1976.
ISBN: 9780125252508.

[Ravikumar] Ravikumar, S., Akeley, K., Banks, M. "Creating effective focus cues in multi-plane 3D displays". In: *Opt. Express,* 19.21 (2011), pp. 20940–20952.
DOI: 10.1364/OE.19.020940

[Renesse] Renesse, R. L. van. *Optical Document Security*. 3rd edition. Artech House, 2005. ISBN: 1580532586.

[Ritter] Ritter, A., Böttger, J., Deussen, O., König, M., Strothotte, T. "Hardware-based rendering of full-parallax synthetic holograms". In: *Appl. Opt.,* 38.8 (Mar 1999), pp. 1364–1369. DOI: 10.1364/AO.38.001364.

[Saleh] Saleh, B. E. A., Teich, M. C. *Fundamentals of photonics.* Wiley, 1991.
ISBN 978-0-471-35832-9.

[Saxby] Saxby, G., Zacharovas, S. *Practical Holography*. Fourth Edition. CRC Press, 2015. ISBN: 9781482251586.

[Schnars] Schnars, U., Falldorf, C., Watson, J., Jüptner, W. *Digital Holography and Wavefront Sensing: Principles, Techniques and Applications*. 2nd. Springer Publishing Company, Incorporated, 2014. ISBN: 9783662446928.

[Shimobaba] Shimobaba, T., Masuda, N., Ito, T. "Simple and fast calculation algorithm for computer-generated hologram with wavefront recording plane". In: *Opt. Lett.,* 34.20 (Oct 2009), pp. 3133–3135. DOI: 10.1364/OL.34.003133.

[Simonsen]	Simonsen, P., Christensen, C. D. *3D imaging device*. US Patent US D617361 S1. Publication date Jun 8, 2010
[Smalley]	Smalley, D. E., Smithwick, Q. Y. J. , Bove, Jr., V. M., Barabas J., Jolly, S. "Anisotropic leaky-mode modulator for holographic video displays". In: *Nature,* 498, 313–317 (20 June 2013). DOI: 10.1038/nature12217
[Smith]	Smith, S. *Digital Signal Processing: A Practical Guide for Engineers and Scientists.* Newnes, 2013. ISBN: 9780080477329.
[Takaki]	Takaki, Y., Tanaka, K., Nakamura, J. "Super multi-view display with a lower resolution flat-panel display". In: *Opt. Express*, 19.5 (2011), pp. 4129–4139. DOI: 10.1364/OE.19.004129.
[Traub]	Traub, A. "Stereoscopic Display Using Rapid Varifocal Mirror Oscillations". In: *Appl. Opt.,* 6.6 (1967), pp. 1085–1087. DOI: 10.1364/AO.6.001085.
[Ulusoy]	Ulusoy, E., Onural, L., Ozaktas, H. "Full-complex amplitude modulation with binary spatial light modulators". In: *J. Opt. Soc. Am. A,* 28.11 (2011), pp. 2310–2321. DOI: 10.1364/JOSAA.28.002310.
[Verheijen]	Verheijen, M. J. "E-beam Lithography for Digital Holograms". In: *Journal of Modern Optics*, 40.4 (1993). DOI: 10.1080/09500349314550761.
[Wetzstein]	Wetzstein, G., Lanman, D., Heidrich, W., Raskar, R. "Layered 3D: tomographic image synthesis for attenuation-based light field and high dynamic range displays". In: *ACM Trans. Graph.,* 30.4, Article 95 (July 2011). DOI: 10.1145/2010324.1964990
[Yatagai]	Yatagai, T. "Stereoscopic approach to 3-D display using computer-generated holograms". In: *Appl. Opt.,* 15.11 (Nov 1976), pp. 2722–2729. DOI: 10.1364/AO.15.002722
[Yoshikawa]	Yoshikawa, H., Taniguchi, H. "Computer Generated Rainbow Hologram". In: *Optical Review,* 6.2 (March 1999), pp. 118–123. DOI: 10.1007/s10043-999-0118-0
[Yu]	Yu, H., Lee, K. R., Park, J., Park, Y. K. "Ultrahigh-definition dynamic 3D holographic display by active control of volume speckle fields". In: *Nature Photonics,* 11 (2017), pp. 186–192. DOI: 10.1038/nphoton.2016.272.