Recent progress in computer-generated holography for three-dimensional scenes

Jae-Hyeung Park
School of Information and Communication Engineering, Inha University, Incheon, South Korea

ABSTRACT
Computer-generated holography (CGH) is a crucial technique in preparing contents for holographic three-dimensional displays. In this paper, the recent progress in the CGH techniques is reviewed, covering the point-cloud-, light-ray-field-, layer-, and polygon-based techniques.

ARTICLE HISTORY
Received 29 August 2016
Accepted 4 October 2016

KEYWORDS
Computer-generated holography; holographic display

1. Introduction
Holography enables the optical acquisition and the reconstruction of the complex optical field, including the amplitude and phase distribution of the light, by means of the interference. By capturing the complex field of the light reflected from a three-dimensional (3D) scene based on the holography principle, it is possible to obtain the 3D geometry of the 3D scene digitally, or to reconstruct it optically as a 3D image. Ever since the holography principle was proposed, holography has been evolving, and it is now a representative technique in capturing and reconstructing the complex optical field.

Computer-generated holography (CGH) is a technique used for obtaining a complex optical field not by optical capture but by numerical calculation. Given the digital data of a 3D scene, the complex field of the light from the 3D scene is calculated in the hologram plane using the diffraction theory. In spite of the dimension reduction from the 3D object volume to the two-dimensional (2D) hologram plane, the calculated complex field still contains the 3D information of the recorded scene. The calculated complex field can then be further processed for numerical interference with a virtual reference wave to yield the interference pattern, or it can be encoded in various ways for efficient optical reconstruction using holographic 3D displays.

In this paper, progress made in the last six years in the field of CGH for representing a 3D scene is reviewed. The basic principle of CGH is explained in Sections 2 and 3, then several approaches to enhancing the calculation speed, reducing the speckle noise, and enabling occlusion culling are introduced in the following sections.

2. Numerical propagation using the angular spectrum
One basic tool for hologram synthesis is the numerical propagation of the optical field—i.e. the calculation of the corresponding optical field when the optical field at the other plane is given. Before reviewing the specific CGH methods, the numerical propagation will be explained in this session. Let the optical field at the \( z = 0 \) plane be represented by \( U(\xi, \eta; z = 0) \), as shown in Figure 1. From the scalar diffraction theory, it can be shown that the corresponding optical field at the \( z = z \) plane is given by

\[
U(x, y; z) = \frac{z}{j\lambda} \int \int U(\xi, \eta; z = 0) \exp \left[ j \frac{2\pi}{\lambda} \sqrt{(x-\xi)^2 + (y-\eta)^2 + z^2} \right] \frac{d\xi}{(x-\xi)^2 + (y-\eta)^2 + z^2} d\eta,
\]

which is called the ‘first Rayleigh–Sommerfeld formula’ [1], or equivalently, it can be represented in the Fourier
is obtained by taking the inverse Fourier transform to obtain the angular spectrum in the output plane \( \Delta \xi \) and \( \Delta \eta \) where the sampling points are the numbers of sampling points. Then transfer function \( T_\xi(f_x, f_y) \) is prepared with sampling intervals \( \Delta f_x \) and \( \Delta f_y \) and is multiplied with \( G(f_x, f_y; z = 0) \) to obtain the angular spectrum in the output plane \( G(f_x, f_y; z = z) \). Finally, output optical field \( U(x, y; z = z) \) is obtained by taking the inverse Fourier transform of \( G(f_x, f_y; z = z) \). The sampling interval in the output optical field is given by \( \Delta x = 1/N_x \Delta f_x = \Delta \xi \) and \( \Delta y = 1/N_y \Delta f_y = \Delta \eta \). Therefore, the output optical field can be obtained with the same sampling interval as that of the input optical field.

### 3. Principle of CGH

In CGH, a 3D scene is decomposed into many primitives. The optical field for each primitive is calculated and accumulated to yield the optical field for the entire 3D scene. Depending on the primitives, the CGH techniques can be classified into the point-cloud-, ray-, polygon-, and layer-based methods, as shown in Figure 2. The principle of each technique is explained in the following.

#### 3.1. Point-cloud-based synthesis

In the point-cloud-based methods, the 3D scene is considered a collection of the self-luminous points, which emit spherical waves. Letting \( z = 0 \) be the hologram plane, the optical field in the hologram plane of a single object point at \((x_0, y_0, z_0)\) is given by

\[
H_0(x, y) = \frac{a_0}{\sqrt{(x-x_0)^2 + (y-y_0)^2 + z_0^2}} \exp \left[ -j \frac{2\pi}{\lambda} \sqrt{(x-x_0)^2 + (y-y_0)^2 + z_0^2} \right]
\]

where \( \lambda \) is the wavelength and \( a_0 \) is the complex constant representing the initial phase and the amplitude of the object point. In the first line of Equation (4), the denominator in the amplitude part is sometimes approximated to a constant as it hardly affects the quality of the reconstructed scene.

In the calculation of Equation (4), the calculation area in the hologram plane should be restricted to avoid aliasing due to the finite sampling grid in the hologram plane. The local spatial frequency of the phase part of \( H_0(x, y) \) is given by

\[
f_\xi(x, y) = \frac{1}{2\pi} \frac{\partial \angle H_0(x, y)}{\partial x} = \frac{\sin \theta_x}{\lambda \sqrt{(x-x_0)^2 + (y-y_0)^2 + z_0^2}}
\]

\[
f_\eta(x, y) = \frac{1}{2\pi} \frac{\partial \angle H_0(x, y)}{\partial y} = \frac{\sin \theta_y}{\lambda \sqrt{(x-x_0)^2 + (y-y_0)^2 + z_0^2}}
\]

where \( \theta_{x,y} \) indicates the angle between the \( z \)-axis and the line joining the object point \((x_0, y_0, z_0)\) and point \((x, y, 0)\) in the hologram plane. Considering the sampling conditions \(1/\Delta x > 2f_\xi \) and \(1/\Delta y > 2f_\eta \) with sampling intervals \( \Delta x \) and \( \Delta y \) in the hologram plane, the calculation area in the hologram plane for a single object point is roughly limited by an area that satisfies...
3.2. Ray-based synthesis

In the ray-based methods, the target 3D scene is prepared in the form of a light ray field, which is a spatio-angular distribution of the radiance of the light rays corresponding to the target 3D scene. The light ray field can be prepared considering multiple reflections, refraction through a volumetric material, and the reflectance model of the 3D image surface for the given illumination for a more realistic representation of the 3D scene \([4,5]\). Suppose \(L(x, y; \theta_x, \theta_y)\) represents the light ray field – i.e. the amplitude of a light ray that passes through a plane at position \((x, y)\) with angle \((\theta_x, \theta_y)\). The plane where the light ray field is represented is called ‘ray-sampling plane.’

The ray-based method generally uses a two-step process. In the first step, it generates an intermediate optical field in the ray-sampling plane from light ray field \(L(x, y; \theta_x, \theta_y)\) and then propagates the generated optical field to the hologram plane to obtain the final hologram \([6]\).

In the first step, the intermediate optical field plane is divided into many sections of equal size \(s\), and the processing is performed section-wise. For section \(A_{m,n}(x_l, y_l)\), where \((x_l, y_l)\) is the local position with respect to the section center at \((x, y) = (ms, ns)\), the corresponding light ray field at the center of section \(L(x = ms, y = ns; \theta_x, \theta_y)\) is Fourier-transformed after it is multiplied with a random phase distribution to yield \(A_{m,n}(x_l, y_l)\). More specifically,

\[
A_{m,n}(x_l, y_l) = \int \int L(ms, ns; \theta_x, \theta_y) \exp[j\psi(\theta_x, \theta_y)] \\
\times \exp \left[ -\frac{2\pi}{\lambda} (\theta_x x_l + \theta_y y_l) \right] \, d\theta_x \, d\theta_y,
\]

where \(\theta_x, \theta_y\) are represented in radian unit, and \(\psi(\theta_x, \theta_y)\) is a random phase distribution in \([0, 2\pi]\). The angular range of light ray field \(L\) in Equation (7) is related to the sampling pitch of intermediate optical-field planes \(\Delta x_l\)

\[\sin \theta_{x,y} < \lambda / 2\Delta x, y\text{ in the rectangular sampling grid. The optical field for the entire 3D scene is simply given by}

\[H(x, y) = \sum_m H_m(x, y). \tag{6}\]
and $\Delta y_l$ by

$$-\frac{\lambda}{2\Delta x_l} \leq \theta_x \leq \frac{\lambda}{2\Delta x_l} \quad \text{and} \quad -\frac{\lambda}{2\Delta y_l} \leq \theta_y \leq \frac{\lambda}{2\Delta y_l}. \quad (8)$$

Intermediate optical-field $H_{\text{intermediate}}(u, v)$ in the ray-sampling plane is completed by tiling sections $A_m,n$. Note that light ray field $L(x = ms, y = ns; \theta_x, \theta_y)$ for each section is actually a perspective view of the 3D scene at the section center $(x = ms, y = ns)$, with the field of view given by Equation (8). Therefore, a few techniques that use the multiple-perspective views to synthesize the hologram [7,8] can also be classified as ray-based methods.

In the second step, the hologram is finally synthesized by propagating the intermediate optical field to the hologram plane numerically using the method explained in Section 2.

The resolution of the 3D image represented in the light ray field is generally degraded as the distance between the 3D image and the ray-sampling plane increases. Therefore, in the ray-based method, the ray-sampling plane is usually located around the 3D image [6]. The two-step process enables the conversion of the light ray field to the optical field around the 3D image, not directly in the hologram plane, which enhances the reconstruction quality of large-depth 3D objects.

As acquisition methods of the light ray field of real-life-sized objects have been developed of late with a light field camera or an integral imaging camera [9], the ray-based method has gained considerable attention as a practical method of obtaining the hologram of real objects.

**3.3. Triangular-mesh-based synthesis**

The triangular-mesh-based method uses the 3D scene represented in the polygon mesh model. The optical field of each triangular facet is calculated and accumulated in the hologram plane to obtain the final hologram of the 3D scene. The calculation of the optical field of the triangular facet is performed by finding the relationship between the angular spectra represented in the hologram plane and in the local plane containing the triangular facet [10–13].

Suppose that global coordinates $(x, y, z)$ and local coordinates $(x_l, y_l, z_l)$ are defined such that the hologram plane is at $z = 0$ and the triangle is at the $z_l = 0$ plane. A $3 \times 3$ rotation matrix $R$ and a $3 \times 1$ translation vector $c$ can be found, which satisfy

$$r_{x_l,y_l,z_l} = R r_{x,y,z} + c, \quad R = \begin{bmatrix} u_{x_l}^T \\ u_{y_l}^T \\ u_{z_l}^T \end{bmatrix}, \quad (9)$$

where $r_{x,y,z} = [x, y, z]^T$ and $r_{x_l,y_l,z_l} = [x_l, y_l, z_l]^T$ are the position vectors in the global and local coordinates. In Equation (9), $u_{x_l}$, $u_{y_l}$, and $u_{z_l}$ are the $3 \times 1$ unit vectors of the local $x_l$, $y_l$, and $z_l$ axes represented in the global coordinates. The optical field represented in global coordinates $U(x, y, z) = U(r_{x,y,z})$ and local coordinates $U_l(x_l, y_l, z_l) = U_l(r_{x_l,y_l,z_l})$ is related by

$$U(r_{x,y,z}) = U_l(R r_{x,y,z} + c) = \int \int G_l(f_{x_l,y_l}) \exp[j 2\pi f_l r_{x_l,y_l,z_l} (R r_{x,y,z} + c)] \, df_{x_l,y_l}, \quad (10)$$

where $f_{x_l,y_l} = [f_{x_l}, f_{y_l}]^T$ and $f_{x_l,y_l,z_l} = [f_{x_l}, f_{y_l}, f_{z_l}]^T = [(f_{x_l}, f_{y_l}, ((1/\lambda)^2 - f_{x_l}^2 - f_{y_l}^2)^{1/2})]^T$ are the local spatial-frequency vectors, and $G_l$ is the local angular spectrum of the triangular facet—i.e. the Fourier transform of $U_l(x_l, y_l, z_l)$. Using the fact that global spatial frequency $f_{x,y,z} = [f_x, f_y, f_z]^T = [(f_x^2 + f_y^2 + (1/\lambda)^2 - f_{x_l}^2 - f_{y_l}^2)^{1/2}]$ is related to the local spatial frequency by $f_{x,y,z} = R^{-1} f_{x_l,y_l,z_l}$ and $R^T = R^{-1}$, Equation (10) can be rewritten as

$$U(r_{x,y,z}) = \int \int G_l(f_{x_l,y_l}) \exp[f_l R r_{x,y,z}] \, df_{x_l,y_l} \times \exp[j 2\pi f_l R (R r_{x,y,z} + c)] \, df_{x_l,y_l} = \int \int G_l(f_{x_l,y_l}) \exp[f_l R (R r_{x,y,z} + c)] \, df_{x_l,y_l} \times \exp[j 2\pi f_l R (R r_{x,y,z} + c)] \, df_{x_l,y_l}, \quad (11)$$

where $f_{x,y} = [f_x, f_y]^T$ is the global spatial frequency and $f_l/f_z = |\, df_{x_l,y_l} / df_{x,y} \,|$ is the Jacobian determinant. As the optical field in global coordinates $U(r_{x,y,z})$ is given by its global angular spectrum $G(f_{x,y})$ by

$$U(r_{x,y,z}) = \int \int G(f_{x,y}) \exp[j 2\pi f_l R (R r_{x,y,z})] \, df_{x,y}, \quad (12)$$

the relationship between the global and local angular spectra can be found by comparing Equations (11) and (12)

$$G(f_{x,y}) = G_l(f_{x_l,y_l}) \exp[j 2\pi f_{x_l,y_l,z_l} (R r_{x,y,z} + c)] f_{z_l}/f_{z}. \quad (13)$$

Therefore, once local angular spectrum $G_l(f_{x_l,y_l})$ is calculated, the global angular spectrum of polygon facet $G(f_{x,y})$ can be obtained from local angular spectrum.
$G_l(f_{xl,yl})$ using Equation (13), and then the optical field in hologram plane $z = 0$ is finally given by

$$H(r_{x,y}) = U(x, y, z = 0) = \int \int G(f_{x,y}) \exp\left[j2\pi f_{x,y}^T r_{x,y}\right] df_{x,y}. \quad (14)$$

There are two approaches according to the way local angular spectrum $G_l(f_{xl,yl})$ is obtained. In one approach called 'Fast Fourier Transform (FFT)-based method' in this paper, the local angular spectrum is obtained in discrete form by taking the discrete Fourier transform of the triangular facet in the local plane [12–14]. As the discrete Fourier transform is usually performed in the rectilinear uniform local spatial-frequency $f_{xl,yl}$ grid, the acquisition of the global angular spectrum in the rectilinear uniform global spatial-frequency $f_{x,y}$ grid using Equation (13) requires resampling or the interpolation of the discrete local angular spectrum.

In another approach called 'fully analytic method,' the local angular spectrum is related to the analytic formula of the angular spectrum of a reference triangle [10,11]. Suppose that reference triangle $g_o(r_{xr,yr})$ is defined and the analytic formula of its angular spectrum $G_o(r_{xr,yr})$ is known. An arbitrary triangle represented in local coordinates $g_l(r_{xl,yl})$ is related to the reference triangle by

$$g_l(r_{xl,yl}) = g_o(A r_{xl,yl} + b) \exp\left[j\frac{2\pi}{\lambda}\left(\begin{bmatrix} u_{xl}^T \\ u_{yl}^T \end{bmatrix} - u_c^T \right) r_{xl,yl}\right], \quad (15)$$

where $A$ is a $2 \times 2$ affine transform matrix and $b$ is a $2 \times 1$ translation vector. The last term accounts for the phase distribution on the triangular-mesh surface due to the carrier wave of direction $u_c$. From Equation (15), the local angular spectrum is given by

$$G_l(f_{xl,yl}) = \int \int g_o(r_{xl,yl}) \exp\left[-j2\pi f_{xl,yl}^T r_{xl,yl}\right] dr_{xl,yl} \exp\left[j2\pi \left(f_{xl,yl} - \left[u_{xl}^T \\ u_{yl}^T \right] u_c^T \right] b\right] \det(A)

\times G_o\left(A^{-T} \left(f_{xl,yl} - \left[u_{xl}^T \\ u_{yl}^T \right] u_c^T \right)\right). \quad (16)$$

Therefore, the local angular spectrum of an arbitrary triangle can be obtained from the analytic formula of the angular spectrum of the reference triangle using Equation (16). In this approach, the global angular spectrum is directly obtained from the analytic formula of the reference angular spectrum by the successive application of Equations (13) and (16), which eliminates the need for any resampling or interpolation, ensuring exact hologram synthesis [15,16].

Reference triangle $g_o(r_{xr,yr})$ is usually assumed to have a uniform amplitude and phase to enable the analytic representation of its angular spectrum. A linearly interpolated amplitude [17] or a textured pattern [18,19] are also used, however, for a more realistic representation of 3D objects with the given number of triangular meshes.

Generally, the triangular-mesh-based method has the advantages of being compatible with computer graphics technologies that represent 3D objects using a polygon mesh. The triangular-mesh-based method is also more computationally efficient than the point-cloud-based method, especially for large objects, as it does not need to fill the whole area with numerous points.

### 3.4. Layer-based synthesis

The layer-based synthesis method prepares the 3D scene as a form of a few layers at different depths. The hologram is obtained by numerically propagating individual layers to the hologram plane and accumulating them. Let $I(\xi, \eta; z = z_i)$ be the intensity distribution of the object slice at the $z = z_i$ plane. The optical field at the $z = z_i$ plane is calculated using the following equation:

$$U(\xi, \eta; z = z_i) = \sqrt{I(\xi, \eta; z = z_i)} e^{j\phi(\xi, \eta)}, \quad (17)$$

where phase term $\phi(\xi, \eta)$ represents the random phase distribution on the object layer, which gives the diffusiveness of the reconstructed image. The corresponding optical field in the hologram plane is obtained by propagating $U(\xi, \eta; z = z_i)$ to the $z = 0$ plane numerically using the method explained in Section 2. This process is repeated for all the layers, and the optical field in the hologram is accumulated, resulting in the final hologram [20].

The layer-based method is simple and is usually computationally efficient as the number of layers is smaller than that of the other primitives, such as the point, triangle, and light ray. The coarse layer density, however, may lead to a layered appearance rather than a continuous 3D appearance. The reflectance model of the object surface (i.e. the angular distribution of the reflectance) is also not easy to encode in the layer-based method.

### 4. Speed enhancement

The naive implementation of CGH usually requires a rather long computation time. In this section, a few recent techniques for reducing the computation time especially
for the point-cloud- and triangular-mesh-based methods are introduced.

4.1. Point-cloud-based synthesis

In the point-cloud-based method, calculation is performed for every 3D point using Equation (4). For the $N_x \times N_y$ hologram resolution, Equation (4) requires $N_x \times N_y$ calculations for each 3D object point. Therefore, for $N_o$ object points, the total amount of calculations is given by $N_x \times N_y \times N_o$.

One method of reducing the calculation time is by using look-up tables (LUTs). It can be observed from Equation (4) that the optical field in the hologram plane corresponding to a single 3D object point is shift-invariant if the 3D object point remains in the same distance $z_o$. Therefore, by pre-calculating the optical fields for every distance and storing them, the LUTs can be constructed. The hologram of a 3D object can be obtained simply by adding the LUTs that are shifted according to the lateral position and multiplied with the complex amplitude of the 3D object points, as shown in Figure 3(a) [21,22]. In the case of a moving 3D scene, the temporal redundancy can also be considered to reduce the computational load. For every frame, the hologram pattern only for the moving part in the scene is updated, and the stationary part is left unchanged. Various techniques that use the simple frame difference, motion estimation, and motion compensation, and the MPEG-compatible techniques, have been reported [23,24].

One problem of the LUT-based method is that the memory requirement for storing $N_x \times N_y$-resolution LUTs for every possible distance is too large. One way of reducing the memory requirement is to use the rotationally symmetric property of the spherical-wave pattern. In Ref. [25], it is shown that the original 2D distribution of the spherical-wave pattern can be obtained by multiplying a pair of 1D distribution functions. Therefore, only two 1D distributions need to be stored instead of the full 2D distribution for each distance. In Refs. [26,27], radial interpolation of a single 1D distribution function is performed to obtain the full 2D distribution. These methods, however, still require the LUTs for every possible distance. Recent papers [28,29] report techniques that calculate the LUTs for different distances from a few pre-calculated patterns to address this issue.

Another approach to accelerating the calculation is to introduce an intermediate plane, as shown in Figure 3(b). From Equation (5), it can be seen that the non-aliased area on the hologram plane for a single object point increases as the distance of object point $z_o$ increases. As the hologram calculation based on Equation (4) needs to be performed only within the non-aliased area for each object point, the amount of calculation is smaller for small-distance object points than for larger-distance object points. Using this property, an intermediate plane is introduced close to the 3D objects, and the optical field is calculated on that plane first. The final hologram is then obtained by numerically propagating the optical field on the intermediate plane using the techniques explained in Section 2.

The intermediate plane is usually called ‘wavefront-recording plane’ in these techniques. Advanced techniques, including combination with the LUT-based method [30], implementation using a graphic processing unit (GPU) [31], and the tilted wavefront-recording plane [32], have been reported.

In the construction of the LUTs or the optical-field calculation on the wavefront-recording plane, the spherical wave should be calculated. Apart from decomposition into a pair of 1D distributions [25] or the radial interpolation [26,27] mentioned earlier, the FFT-based method has also been proposed [33]. In this method, the spherical-wave pattern is divided into small rectangular sections, each of which is approximated by a linear-phase
distribution, as shown in Figure 4. Linear-phase distribution approximation enables efficient pattern generation using FFT operation, significantly reducing the overall processing time. The relative phase between the divided sections can be matched to minimize the approximation error. The photo-realistic color reproduction of the 3D scene using this technique has been demonstrated [33].

Speed enhancement using GPUs or a specialized processor is also an active research field. In the point-cloud-based hologram calculation, the hologram patterns for the object points have a similar shape and are obtained from the same mathematical formula. The calculation of the hologram pattern for each object point is independent from that for the other object points. This property is highly suitable for parallel processing. Using this property, several researches have been reported to develop a computation system for the high-speed generation of the hologram [34–37].

### 4.2. Triangular-mesh-based synthesis

In the triangular-mesh-based methods, the hologram pattern is calculated for each triangular facet, and is accumulated to yield the final hologram. Unlike the point-cloud-based methods, the elementary hologram patterns for triangular facets are all different according to the size, shape, and orientation of the triangles. Therefore, it is not possible to use LUT, as the point-cloud-based methods do. Speed enhancement is usually achieved in the triangular-mesh-based methods by limiting the calculation window or approximating the calculation procedures.

In the FFT-based method, each triangular facet is Fourier-transformed using the discrete FFT in the local plane to obtain $G_i(f_{d_{x,y}})$ before it is related to global angular spectrum $G(f_{x,y})$ using Equation (13). The research reported in Refs. [38,39] performed only a single FFT operation on a reference triangle rather than repeated FFT operations for every triangle to reduce the computation time. The local angular spectrum for an arbitrary triangle is obtained by resampling the obtained discrete local angular spectrum of the reference triangle considering the geometric relationship between the given and reference triangles. The computation time for resampling is smaller than that for FFT operation, and hence, the total computation time is reduced.

In the fully analytic method, resampling should be avoided as it degrades the benefit of the method: the precision of the calculated hologram free from errors caused by any resampling and interpolation. Instead, speed enhancement can be achieved by limiting the spatial-frequency range where the angular spectrum is actually calculated [40]. In the fully analytic method, the analytic formula of the angular spectrum of the reference triangle is prepared. The global angular spectrum is calculated by addressing the analytic formula in a rectangular and regular global spatial-frequency grid after transforming the global grid to a reference triangle plane. By observing the analytic formula of the angular spectrum of the reference triangle, it can be found that much of the energy is concentrated around the low-spatial-frequency area and the three radial lines that correspond to the three sides of the reference triangle, as shown in Figure 5. In the other area, the angular spectrum of the reference triangle has a very low value. This is because the reference triangle in the fully analytic method is usually assumed to have a uniform amplitude and phase inside the triangular area. Using this fact, the addressing and calculation of the global angular spectrum can be performed not in the whole global spatial-frequency grid but only where the reference angular spectrum has a significant value without loss of precision. Reference [40] reports significant computation time reduction using this technique.

A hybrid method between the triangular-mesh- and point-cloud-based methods has also been reported [41]. In this method, a 3D object is initially prepared in the triangular-mesh model. The hologram calculation, however, is performed on a point-by-point basis. Each
triangular facet is represented by several equally spaced points, and the hologram is calculated for each point. As all the points for each triangular facet are on the same plane and are close to one another, some approximation can be justified, which leads to speed enhancement in this technique.

Finally, speed enhancement for a large-viewing-angle hologram is worth mentioning. The hologram is calculated for a given viewing direction and angular range. Changing the viewing direction in the fully analytic method can be done simply in principle, by changing the carrier wave direction, but it is actually time-consuming as it requires repetition of the whole calculation. Reference [42] reports a method of avoiding the whole calculation. Reference [42], the elementary hologram pattern for each triangular facet is calculated for a viewing direction, and is stored separately. Viewing direction update is performed by shifting the stored elementary hologram pattern laterally and multiplying it with a linear-phase function. Although this method requires a large memory for storing elementary hologram patterns and causes errors when the change of the viewing direction is large, it is efficient when one needs to update the hologram fast within a small range of viewing direction changes. Hologram generation for a very large viewing angle can be performed by using a pre-calculated optical field not on a flat plane but on a cylindrical or spherical plane. In this case, the hologram for the given viewing direction is obtained by transforming the optical field pre-calculated on the cylindrical or spherical plane into a hologram plane [43,44].

5. Reflectance model of a 3D object

The reflectance model of a 3D object defines the angular distribution of the reflectance of an individual 3D object surface for the given illumination, which is crucial in the realistic representation of the object. Phong’s reflection model is a simple and common example of this [45]. In the usual computer graphics rendering, the viewpoint of the 3D scene is fixed, and thus, the reflectance of each surface is given by a scalar value. In a hologram, however, the 3D object is reconstructed within a viewing angle; thus, the whole angular reflectance distribution should be encoded into a single hologram.

For the point-cloud-based method, a simple but effective method has been proposed [46]. In this method, the amplitude of the spherical-wave pattern for each 3D point is spatially modulated to represent the reflectance model of the 3D point, as shown in Figure 6. As the brightness of the reconstructed 3D object point for a given observation direction is determined by the amplitude of the corresponding area in the spherical-wave pattern, the angular distribution of the brightness can be controlled by spatially modulating the amplitude of the spherical-wave pattern.

In the triangular-mesh-based method shown in Figure 7, such simple correspondence does not exist as the triangle may have an arbitrary shape and orientation. In the FFT-based method, where the local angular spectrum of the triangle is obtained in discrete form by performing the FFT operation on the local plane, the phase distribution on the triangle has been found to achieve the desired reflectance model. As the complex field of the triangle on its local plane and the angular distribution of the reconstructed light have a Fourier transform relationship, the corresponding phase distribution in the triangle domain can be obtained by simultaneously forcing the angular radiance in the Fourier domain to the desired reflectance model, and the amplitude distribution in the local triangle domain to the shape of the triangle [47]. The iterative Fourier transform algorithm can also be used in this regard. It has also been reported to model each triangular facet with microfacets [48], or to analyze a real-object facet using the finite-difference time domain method [49].

In the fully analytic method, arbitrary phase distribution in the triangle cannot be considered because its angular spectrum is not analytic. In the initial proposal of the fully analytic method [10], the division of each triangle and random phase assignment was proposed to control the diffusiveness of the surface. The arbitrary-reflectance model, however, cannot be implemented in this case. Recently, a convolution-based method was proposed for the implementation of the arbitrary-reflectance model [50]. In this method, the global angular spectrum for a triangular facet with a uniform amplitude and linear phase corresponding to a carrier wave is calculated first. The calculated global angular spectrum is then
convoluted with a kernel that corresponds to the desired reflectance model of the triangular facet. The convolution is performed not on the local plane but on the global plane, and thus, it maintains a fully analytic framework. The error caused by the convolution is also negligible, which makes this technique attractive.

6. Speckle reduction

Speckle is a spatial random fluctuation of the intensity in the observed 3D images when they are reconstructed using a coherent light source such as laser. In CGH, the 3D image is reconstructed such that an individual point in the reconstruction has a random or significantly different phase with respect to the neighboring ones to ensure a large viewing angle. These reconstructed 3D image points are interfered with on the observation plane, such as the retina of the observer, resulting in speckle noise.

One obvious solution to this problem would be to use a less-coherent light source, such as the light-emitting diode (LED). The increased spectral bandwidth of LED, however, will result in the blurring of the reconstructed images, degrading the image quality. Another obvious solution would be to calculate many CGHs with different random phases, and to present them fast. The speckle patterns in the observation plane will then be washed out, resulting in clean images. The required refresh rate of the spatial light modulator (SLM), however, is too high to achieve the desired speckle reduction as the speckle contrast is reduced only to $M^{1/2}$ when $M$ images with different random phases are presented [51].

In the point-cloud-based method, a simple technique that removes the speckle in the reconstruction using time-multiplexing has been proposed [52,53]. The proposed technique reconstructs 3D image points in a sparse grid at a time, as shown in Figure 8. By time-multiplexing the CGHs of different point groups, the whole 3D image with a dense point cloud can be reconstructed. As the close points are reconstructed at different times, interference does not occur between them, resulting in speckle-free images. Although this technique uses time-multiplexing, it requires much less time-multiplexing than the traditional method, which gives different random phases for averaging the speckle pattern. Sparse point reconstruction using point grouping has also been reported for image-quality enhancement [54].

In the ray-based method, a similar approach has been proposed [55]. Instead of the 3D point, the light rays are sparsely selected in the CGH generation. In the reconstruction, multiple CGHs for different groups of rays are reconstructed in a time-multiplexing manner, reproducing 3D images without speckle.
7. Occlusion

A hologram should properly present the occlusion of the objects behind the closer one for the realistic representation of 3D images. As the hologram has a viewing angle rather than a fixed viewpoint, this occlusion processing is more difficult than the usual computer graphics rendering. A simple method that selects the visible parts of the object primitives (i.e., 3D object points or triangular meshes) to a specific viewpoint fails to deliver a realistic reproduction as the visible parts vary according to the viewing direction within the viewing angle.

In the point-cloud-based method, an occlusion mask in the wavefront-recording plane has been proposed to address the occlusion issue [56]. The spherical-wave patterns are accumulated in the wavefront-recording plane from the rear-object points to the front-object points. When the front-object point is processed, the occlusion mask is multiplied in the wavefront-recording plane before the spherical-wave pattern for the object point is accumulated to occlude the rear-object points. The complement Gaussian amplitude distribution is used as the occlusion mask in this technique. The width of the Gaussian distribution is selected considering the point density of the point cloud. This method, however, does not give the exact occlusion of the 3D object.

The occlusion between the light ray fields has been well developed in computer graphics. In the ray-based CGH, this ray occlusion technique can be used to realize the occlusion effect in the hologram. In Ref. [57], the ray-sampling planes are placed around each 3D object. The intermediate optical field in the rear ray-sampling plane for the rear object is propagated to the front ray-sampling plane and converted to a light ray field. This light ray field for the rear object is occluded by the light ray field of the front object. The composite light ray field containing the front object and the occluded rear object is then transformed into an optical field in the front ray-sampling plane. Finally, this intermediate optical field is propagated to the hologram plane. This technique achieves occlusion by successive conversion between the optical and light ray fields.

A silhouette method that can be applied to general CGHs has also been proposed [58]. In this method, an object plane is defined around each object, and the binary occlusion mask is set as the cross-section of the object with the plane, as shown in Figure 9. The optical field propagated from the rear object to the front-object plane is multiplied with the occlusion mask to block the light meeting the cross-section of the front object. The addition of the unblocked optical field for the rear object and the optical field for the front object is propagated to the hologram plane. This basic method can be further elaborated to ensure more exact occlusion in the case of the triangular-mesh-based CGH [58]. A naive approach would be to set a binary occlusion mask for each triangle of the 3D object, and to repeat the occlusion process for every triangle, which will result in excessive computation time. In Ref. [58], it was instead proposed that each triangle not be treated as an occluding mask but as a transparent aperture. The final hologram with correct occlusion is obtained by subtracting the non-occluded optical field with the optical field through the triangle aperture using Babinet’s principle. In this way, only a small part of the optical field that passes through the triangle aperture needs to be calculated, resulting in a significant reduction of the computation time.

8. Conclusion

In this paper, the recent progress in the CGH research field is introduced. The basic principles of the point-cloud-
triangular-mesh-, light-ray-field-, and layer-based methods are explained, and their recent progress, including speed enhancement, reflectance model implementation, speckle suppression, and occlusion culling, is presented. Although various relevant research studies have been actively conducted of late, the CGH techniques still have much room for enhancement. The real-time generation of a high-resolution hologram, the photo-realistic reproduction of a 3D scene without speckle, and the generation of a wide-viewing-angle CGH with the correct reflectance model and the occlusion effects are a few examples of relevant matters that require further research. Although they are not covered in this paper, the quantitative assessment of the CGH quality, the optimal encoding of the CGH to suppress the noise from the limited modulation and defects of the SLM, the development of an authoring framework of CGH, and an efficient compression scheme for CGH, and the conversion of CGH for holographic displays with different specifications and viewing conditions and maximizing its compatibility with other 3D content formats are also important topics related to CGH. CGH is a highly promising and demanding technique. It is believed that the recent progress in CGH will provide a new standard for 3D content representation and will play a vital role in many applications, including holographic 3D displays.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by MSIP (Ministry of Science, ICT, and Future Planning) of South Korea, under the ITRC (Information Technology Research Center) support program (IITP-2016-R0992-16-1008) supervised by IITP (Institute for Information & Communications Technology Promotion).

Notes on contributor

Jae-Hyeung Park received his B.S., M.S., and Ph.D. degrees from Seoul National University in 2000, 2002, and 2005, respectively. After graduation, he joined Samsung Electronics, where he worked on the development of motion blur reduction techniques for LCDs. From 2007 to 2012, he was on the faculty of Chungbuk National University in South Korea. In 2013, he joined the faculty of Inha University, where he is now an associate professor. He has been working on the acquisition, processing, and display of three-dimensional information using the holography and light field techniques.

References

[1] J.W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill Book Co., New York, 1996).
[2] R.P. Muffoletto, J.M. Tyler and J.E. Tohline, Opt. Exp. 15, 5631 (2007).
[3] N. Okada, T. Shimobaba, Y. Ichihashi, R. Oi, K. Yamamoto, M. Oikawa, T. Kakue, N. Masuda and T. Ito, Opt. Exp. 21, 9192 (2013).
[4] T. Ichikawa, K. Yamaguchi and Y. Sakamoto, Appl. Opt. 52, A201 (2013).
[5] T. Ichikawa, T. Yoneyama and Y. Sakamoto, Opt. Exp. 21, 32019 (2013).
[6] K. Wakunami and M. Yamaguchi, Opt. Exp. 19, 9086 (2011).
[7] H. Zhang, Y. Zhao, L. Cao and G. Jin, Opt. Exp. 23, 3901 (2015).
[8] Y. Ohsawa, K. Yamaguchi, T. Ichikawa and Y. Sakamoto, Appl. Opt. 52, A167 (2013).
[9] S.-K. Lee, S.-I. Hong, Y.-S. Kim, H.-G. Lim, N.-Y. Jo and J.-H. Park, Opt. Exp. 21, 23662 (2013).
[10] H. Kim, J. Hahn and B. Lee, Appl. Opt. 47, D117 (2008).
[11] L. Ahrenberg, P. Benzie, M. Magnor and J. Watson, Appl. Opt. 47, 1567 (2008).
[12] K. Matsushima and S. Nakahara, Appl. Opt. 48, H54 (2009).
[13] K. Matsushima, H. Nishi and S. Nakahara, J. Electron. Imaging. 21, 23002 (2012).
[14] Y. Pan, Y. Wang, J. Liu, X. Li, J. Jia and Z. Zhang, Appl. Opt. 52, 4391 (2013).
[15] J.-H. Park, H.-J. Yeom, H.-J. Kim, H. Zhang, B. Li, Y.-M. Ji and S.-H. Kim, Opt. Exp. 23, 8006 (2015).
[16] D. Im, E. Moon, Y. Park, D. Lee, J. Hahn and H. Kim, Opt. Lett. 39, 3642 (2014).
[17] J.-H. Park, S.-B. Kim, H.-J. Yeom, H.-J. Kim, H. Zhang, B. Li, Y.-M. Ji, S.-H. Kim and S.-B. Ko, Opt. Exp. 23, 33893 (2015).
[18] W. Lee, D. Im, J. Paek, J. Hahn and H. Kim, Opt. Exp. 22, 31180 (2014).
[19] H.-G. Lim, N.-Y. Jo and J.-H. Park, *Digital Holography and Three-dimensional Imaging* (Optical Society of America, Hawaii, 2013), p. DW2A.8.
[20] Y. Zhao, L. Cao, H. Zhang, D. Kong and G. Jin, Opt. Exp. 23, 25440 (2015).
[21] M.E. Lucente, J. Electron. Imaging. 2, 28 (1993).
[22] S.-C. Kim and E.-S. Kim, Appl. Opt. 47, D55 (2008).
[23] S.-C. Kim, X.-B. Dong, M.-W. Kwon and E.-S. Kim, Opt. Exp. 21, 11568 (2013).
[24] X.-B. Dong, S.-C. Kim and E.-S. Kim, Opt. Exp. 22, 8047 (2014).
[25] S.-C. Kim, J.-M. Kim and E.-S. Kim, Opt. Exp. 20, 12021 (2012).
[26] S. Lee, H. Chang, H. Wey and D. Nam, Appl. Opt. 55, A104 (2016).
[27] T. Nishitsuji, T. Shimobaba, T. Kakue and T. Ito, Opt. Exp. 23, 9852 (2015).
[28] S.-C. Kim and E.-S. Kim, Opt. Exp. 22, 22513 (2014).
[29] C. Gao, J. Liu, X. Li, G. Xue, J. Jia and Y. Wang, Opt. Exp. 23, 33194 (2015).
[30] T. Shimobaba, H. Nakayama, N. Masuda and T. Ito, Opt. Exp. 18, 19504 (2010).
[31] J. Weng, T. Shimobaba, N. Okada, H. Nakayama, M. Oikawa, N. Masuda and T. Ito, Opt. Exp. 20, 4018 (2012).
[32] D. Arai, T. Shimobaba, K. Murano, Y. Endo, R. Hirayama, D. Hiyama, T. Kakue and T. Ito, Opt. Exp. 23, 1740 (2015).
[33] H. Kang, E. Stoykova and H. Yoshikawa, Appl. Opt. 55, A135 (2016).
[34] T. Shimobaba, T. Ito, N. Masuda, Y. Ichihashi and N. Takada, Opt. Exp. 18, 9955 (2010).
[35] Y.-H. Seo, H.-J. Choi, J.-S. Yoo and D.-W. Kim, Opt. Exp. 19, 8750 (2011).
[36] N. Takada, T. Shimobaba, H. Nakayama, A. Shiraki, N. Okada, M. Oikawa, N. Masuda and T. Ito, Appl. Opt. 51, 7303 (2012).
[37] Y.H. Lee, Y.H. Seo, J.S. Yoo and D.W. Kim, J. Opt. Soc. Korea 18, 698 (2014).
[38] F. Yang, A. Kaczorowski and T.D. Wilkinson, Appl. Opt. 53, 8261 (2014).
[39] Y. Pan, Y. Wang, J. Liu, X. Li and J. Jia, Appl. Opt. 52, A290 (2013).
[40] D. Im, J. Cho, J. Hahn, B. Lee and H. Kim, Opt. Exp. 23, 2863 (2015).
[41] Y. Ogihara and Y. Sakamoto, Appl. Opt. 54, A76 (2014).
[42] J. Cho, J. Hahn and H. Kim, Opt. Exp. 20, 28282 (2012).
[43] K. Hosoyachi, K. Yamaguchi, T. Ichikawa and Y. Sakamoto, Appl. Opt. 52, A33 (2013).
[44] H. Sakata, K. Hosoyachi, C.-Y. Yang and Y. Sakamoto, Appl. Opt. 50, H306 (2011).
[45] B.T. Phong, Commun. ACM 18, 311 (1975).
[46] T. Kurihara and Y. Takaki, Opt. Exp. 20, 3529 (2012).
[47] H. Nishi, K. Matsushima and S. Nakahara, Appl. Opt. 50, H245 (2011).
[48] K. Yamaguchi, T. Ichikawa and Y. Sakamoto, Appl. Opt. 50, H195 (2011).
[49] T. Ichikawa, Y. Sakamoto, A. Subagyo and K. Sueoka, Appl. Opt. 50, H211 (2011).
[50] H.-J. Yeom and J.-H. Park, Opt. Exp. 24, 19801 (2016).
[51] J.W. Goodman, J. Opt. Soc. Am. 66, 1145 (1976).
[52] Y. Takaki and M. Yokouchi, Opt. Exp. 19, 7567 (2011).
[53] T. Kurihara and Y. Takaki, Opt. Exp. 21, 4044 (2013).
[54] F. Yang, A. Kaczorowski and T.D. Wilkinson, Appl. Opt. 54, 5726 (2015).
[55] T. Utsugi and M. Yamaguchi, Opt. Exp. 22, 17193 (2014).
[56] A. Symeonidou, D. Blinder, A. Munteanu and P. Schelkens, Opt. Exp. 23, 22149 (2015).
[57] K. Wakunami, H. Yamashita and M. Yamaguchi, Opt. Exp. 21, 21811 (2013).
[58] K. Matsushima, M. Nakamura and S. Nakahara, Opt. Exp. 22, 24450 (2014).