HOKKAIDO UNIVERSITY

| | |
|---|---|
| Title | Computer-generated holograms using multiview images captured by a small number of sparsely arranged cameras |
| Author(s) | Ohsawa, Yusuke; Yamaguchi, Kazuhiro; Ichikawa, Tsubasa et al. |
| Citation | Applied Optics, 52(1), A167-A176<br>https://doi.org/10.1364/AO.52.00A167 |
| Issue Date | 2013-01-01 |
| Doc URL | https://hdl.handle.net/2115/52063 |
| Rights | © 2013 Optical Society of America |
| Type | journal article |
| File Information | A052-1_A167-A176.pdf |

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

# Computer-generated holograms using multiview images captured by a small number of sparsely arranged cameras

Yusuke Ohsawa,* Kazuhiro Yamaguchi, Tsubasa Ichikawa, and Yuji Sakamoto

Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan

*Corresponding author: ohsawa@ist.hokudai.ac.jp

Received 9 August 2012; revised 3 October 2012; accepted 4 October 2012;
posted 5 October 2012 (Doc. ID 174083); published 13 November 2012

Computer-generated holograms (CGHs) using multiview images (MVIs) are holograms generated with multiple ordinary cameras. This process typically requires a huge number of cameras arranged at high density. In this paper, we propose a method to improve CGH using MVIs that obtains the MVIs by using voxel models rather than cameras. In the proposed method the voxel model is generated using the shape-from-silhouette (SFS) technique. We perform SFS using a small number of cameras arranged sparsely to create voxel models of objects and then generate the required number of images from these models by volume rendering. This enables us to generate CGHs using MVIs with just a small number of sparsely arranged cameras. Moreover, the proposed method arrange CGHs using MVIs at arbitrary positions. © 2012 Optical Society of America

*OCIS codes:* 090.0090, 090.1760.

## 1. Introduction

There has been much previous research on three-dimensional (3D) display technologies. Holography is a particularly ideal 3D display technology to record and reconstruct 3D images. In holography, 3D images are recorded as an interference pattern of an object light wave and a laser light wave; this interference pattern is called a hologram. To record 3D images in the hologram, it is necessary to prepare a precise optical system and to use laser light, so needless to say, the generation of holograms requires advanced expertise.

Computer-generated holograms (CGHs) are holograms generated by simulating the propagation, interference, and diffraction of light waves with computers. The generated interference pattern is two-dimensional (2D) image data, which means that CGHs can be printed onto film or displayed on a liquid crystal display. CGHs are advantageous in that they require no optical system or laser light to record 3D images. However, it is difficult to generate CGHs from real existing objects because 3D model data of the objects on a computer is required and creating precise models of complicated objects is difficult.

There are several methods to overcome this problem, such as using multiview images (MVIs) to generate CGHs. A direct calculation method of CGHs from MVIs [1] is one approach that uses MVIs. In this method, CGHs are generated by calculating the diffraction integral without transformation between real space and Fourier space, but a ridiculously huge number of images is required since each image corresponds to one individual pixel of the CGH. A method of generating CGHs from 3D Fourier spectra [2] has been reported. This method is based on computerized tomography, so it is not suitable to use MVIs captured by ordinary cameras. Generating CGHs using integral photography (IP) [3] is another approach. IP is a technique to record 3D information of a scene as IP images by lens array [4]. Using the 3D information obtained by IP, CGH is generated

1559-128X/13/01A167-10$15.00/0
© 2013 Optical Society of America

without creating 3D models of objects. However, this method requires specially designed photography devices and cannot produce high-resolution reconstructed images.

CGHs created using MVIs captured by ordinary digital cameras has been proposed [5,6]. These methods are based on holographic stereograms [7] generated by the spatial frequency distribution estimated from the MVIs. In this way, high-resolution reconstructed images are obtained by ordinary devices. There have been several methods proposed to improve CGHs using MVIs [8,9], but these methods still require a huge number of images from cameras arranged at high density to ensure CGHs of a sufficient continuous motion parallax. Reducing the number of cameras is a key practical issue in CGH using MVIs. There have been previous studies on reducing the number of cameras [10,11] in which a relatively small number of images are captured and new images are generated from them. These techniques are effective, but the required number of cameras are still a lot, and the density of cameras is not low enough.

In this paper, we propose a method to generate CGHs using MVIs with a small number of cameras arranged sparsely using voxel models. In this method, the voxel model is generated using shape-from-silhouette (SFS) [12,13]. SFS is advantageous because it can be performed with just a small number of sparsely arranged cameras. We create voxel models of objects using SFS and then generate the required number of images from these models by volume rendering. The proposed method enables us to generate CGHs using MVIs with a small number of cameras arranged sparsely; moreover, the generated CGHs have the same quality and provide the same motion parallax as CGHs generated from a large number of cameras. Furthermore, when the images are rendered correctly, the hologram planes are arranged at arbitrary positions regardless of the presence or absence of cameras at that position.

The rest of this paper is organized as follows. We describe a basic calculation method for CGHs using MVIs in Section 2. In Section 3, we discuss the algorithm we developed to reduce the required number of cameras. We performed experiments to determine the effectiveness of the proposed method and discuss the result in Section 4. We conclude the paper in Section 5.

## 2. CGHs Using MVIs

CGHs using MVIs are based on holographic stereograms [7] generated by calculating the spatial frequency distribution of object light geometrically from captured MVIs [5]. A CGH using MVIs is composed of many small holograms, called elementary holograms, tiled without gaps or overlaps. 3D images are reconstructed by superposing object light waves emitted from elementary holograms that are generated on the basis of MVIs. These MVIs are captured with cameras arranged at even intervals on a plane, with the camera direction being perpendicular to the plane. We call the MVIs for generating the elementary holograms "elementary images." In this section, we describe how to calculate the hologram from elementary images in two steps: (1) calculating the object light distribution and (2) generating hologram data.

### A. Calculating Object Light Distribution

Figure 1 shows a schematic diagram of the camera we used. An object is regarded as consisting of innumerable point light sources. When we capture an image with a camera, a light wave from the point light source passes through the camera lens and travels to the image plane of the camera. If the camera is a pinhole camera, the point light source will correspond to the point on the image plane. In this case, the light wave is considered a plane wave. The direction of the plane wave is the same as the direction of a ray that passes through the center of the camera lens. Let $D_L$ denote the distance between the lens plane $\xi-\eta$ and the image plane $X$-$Y$. The incidence angles of the plane wave on $\xi$ and $\eta$ axes $\theta_\xi$ and $\theta_\eta$ are given by

$$\theta_\xi = \tan^{-1}\frac{X}{D_L}, \tag{1}$$

$$\theta_\eta = \tan^{-1}\frac{Y}{D_L}, \tag{2}$$

where $X$ and $Y$ are the values of the coordinates on the image plane. The spacial frequencies on lens planes $f_\xi$ and $f_\eta$ are therefore written by

$$f_\xi = \frac{\sin\theta_\xi}{\lambda} = \frac{\sin\left(\tan^{-1}\frac{X}{D_L}\right)}{\lambda}, \tag{3}$$

$$f_\eta = \frac{\sin\theta_\eta}{\lambda} = \frac{\sin\left(\tan^{-1}\frac{Y}{D_L}\right)}{\lambda}, \tag{4}$$

where $\lambda$ is the wavelength of the plane wave. These equations indicate that only the direction of the

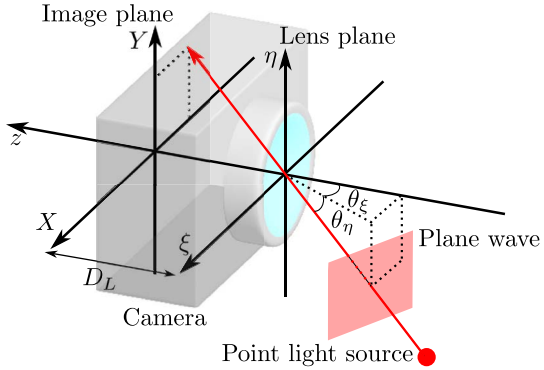



Fig. 1. (Color online) Schematic diagram of camera to capture elementary images.

plane wave determines the spatial frequencies. Let $g(X,Y)$ be the intensity distribution of the elementary image. The spatial frequency distribution on the lens plane $Q(f_\xi,f_\eta)$ is calculated from $g(X,Y)$ because the plane wave that enters the lens plane is in proportion to the intensity of the point on the image plane $(X,Y)$. Therefore, the spatial frequency distribution on the lens plane $Q(f_\xi,f_\eta)$ is calculated by converting the $X$ and $Y$ of $g(X,Y)$ into $f_\xi$ and $f_\eta$ in accordance with Eqs. (3) and (4). Here, the elementary image $g(X,Y)$ lacks phase information, so $Q(f_\xi,f_\eta)$ has only an intensity. Now, when we assume all surfaces of the object are perfectly diffuse, the phase of the object light wave is random. In this case, the spatial frequency distribution is calculated by adding a random phase, $Q'(f_\xi,f_\eta)$, given by

$$Q'(f_\xi,f_\eta) = Q(f_\xi,f_\eta)\exp[-j\sigma(f_\xi,f_\eta)], \tag{5}$$

where $\sigma(f_\xi,f_\eta)$ is a random phase function. The object light distribution on the lens plane $q'(\xi,\eta)$ is expressed using the inverse Fourier transform, as

$$q'(\xi,\eta) = \mathcal{F}^{-1}[Q'(f_\xi,f_\eta)]. \tag{6}$$

$q'(\xi,\eta)$ is the distribution of the object light wave, which travels in the neighborhood of the center of the lens of the camera. Henceforth, we consider the lens plane to be an elementary hologram plane $\xi-\eta$.

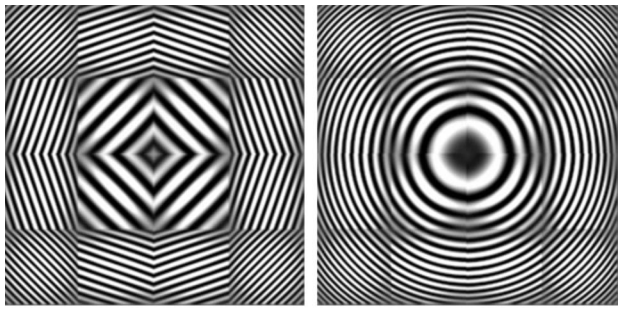
Equation (6) is calculated under the assumption that the object light wave is the plane wave; we call this "plane wave assumption." However, the object light emitted from the point is in actuality the spherical wave. The object light wave is therefore transformed into a spherical wave [6]; we call this transformation "spherical wave assumption." Now, we consider a convex lens placed in front of the elementary hologram plane $\xi-\eta$. The focal distance of the convex lens is the estimated depth of the object. In this arrangement, the convex lens transforms the object light wave into a spherical wave after passing through the elementary hologram plane $\xi-\eta$. Let $q'_s(\xi,\eta)$ be the object light distribution transformed into the spherical wave using a convex lens. $q'_s(\xi,\eta)$ is described as

$$q'_s(\xi,\eta) = q'(\xi,\eta)\exp\left[j\frac{\pi}{\lambda z_e}(\xi^2+\eta^2)\right], \tag{7}$$

where $z_e$ is the estimated depth of the object. Figure 2 shows the zone plates of CGH using MVI. To make the images easier to see, these zone plates have been calculated without random phase. Using the spherical wave assumption, we obtain a near-ideal zone plate.

### B. Generating Hologram Data

The hologram is generated by calculating the interference of the object light wave and a reference light wave. In this subsection, we describe how we generated a plane hologram. In this case, we capture



(a) (b)

Fig. 2. Zone plates of CGH using MVI: (a) generated by plane wave assumption, (b) generated by spherical wave assumption.

the elementary images with cameras arranged at even $\Delta S$ intervals on a hologram plane $u$-$v$. Figure 3 shows the camera arrangement and the hologram plane. When there are $N_H$ cameras in the horizontal direction and $N_V$ cameras in the vertical direction, let $q'_{s_{mn}}(\xi_{mn},\eta_{mn})$ denote the object light distribution $q'_s(\xi,\eta)$ obtained by the $(m,n)$th camera, where $m=1,2,\ldots,N_H$ and $n=1,2,\ldots,N_V$. The size of $q'_{s_{mn}}(\xi_{mn},\eta_{mn})$ is $\Delta S\times\Delta S$, and the center of $q'_{s_{mn}}(\xi_{mn},\eta_{mn})$ is located in the center of the lens of the $(m,n)$th camera. Then, an entire object light distribution $h(u,v)$ is generated using tiled $q'_{s_{mn}}(\xi_{mn},\eta_{mn})$ on the hologram plane $u$-$v$. The size of $h(u,v)$ is $N_H\Delta S\times N_V\Delta S$; this size will be the hologram size. Finally, the intensity distribution of the interference pattern $I(u,v)$ is given by

$$I(u,v) = \|h(u,v)+r(u,v)\|^2, \tag{8}$$

where $r(u,v)$ is the reference light distribution on the hologram plane $u$-$v$. The interference pattern $I(u,v)$ is the hologram data generated from the elementary image.

### C. Size of Elementary Holograms

The elementary holograms are segmentations of the hologram calculated by Eq. (8). Each elementary hologram corresponds to an object light distribution

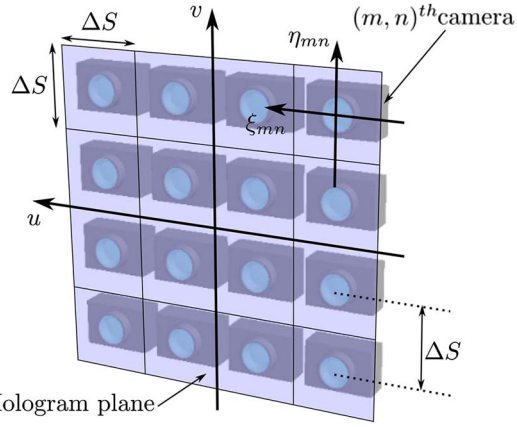



Fig. 3. (Color online) Hologram plane and camera arrangement for calculation of plane hologram.

$s_{mn}(\xi_{mn}, \eta_{mn})$, so its size is $\Delta S \times \Delta S$. Let us consider the size of an elementary hologram from the viewpoint of a motion parallax.

In CGH using MVIs, observers perceive 3D images by a superposition of the light waves emitted from the elementary holograms. They observe motion parallax when changes to their viewpoint cause changes to the reconstructed light waves entering their eyes. However, if $\Delta S$ is large, the observer cannot obtain sufficient motion parallax. We need to ensure that $\Delta S$ is small enough to prevent this lack of recognition.

Elementary holograms require $\Delta S$ designed to be smaller than the size of a pupil. A typical pupil ranges from about 2–7 mm, so $\Delta S$ should be smaller than 2 mm. Note that $\Delta S$ is the interval between cameras when capturing elementary images. To keep the size under 2 mm, we have to arrange the cameras densely at intervals of less than 2 mm. Such a camera arrangement is not realistic because it requires too many cameras. Even arranging the cameras at such dense intervals would be difficult.

## 3. Reducing the Number of Cameras for CGHs Using MVIs

As stated previously, to generate CGHs using MVIs, we have to arrange a huge number of cameras densely on a plane. There have been some studies on how to ease restrictions on this camera arrangement [8,9]. The methods proposed in these studies lift restrictions on both the depth and direction of the cameras, and the cameras do not have to be arranged at even intervals. However, the required number of cameras and their density are still unrealistically high.

In this paper, we propose a method to drastically reduce the required number of cameras for CGH using MVIs that consists of just two steps: (1) creating a voxel model of the object by SFS, and (2) generating the elementary images by volume rendering. SFS can be carried out by a small number of cameras arranged sparsely, so CGH using MVIs requires less restricted cameras. Conventionally, the holographic stereogram generated by a small number of elementary images is inferior in quality and provides an insufficient motion parallax. In the proposed method, the elementary images are rendered using a voxel model. We obtain a sufficient number of elementary images even with a small number of cameras. The proposed method therefore generates a hologram of the same quality and providing the same motion parallax as a CGH generated from a huge number of elementary images. In this section, we describe the calculation method used in the proposed method.

### A. Creating Voxel Model by SFS

There have been many studies on constructing the approximate shape of an object [14–17]. These methods can be used to construct a high-quality shape, but their algorithms are complex and susceptible to noise. In contrast, SFS [12,13] is quite robust due to its simplicity and less subject to the noise. Furthermore, various methods to improve SFS have been proposed [18–21]. These are the reasons we focused on SFS in this paper. SFS constructs the approximate shape as a voxel model by using silhouette images generated from captured MVIs. The silhouettes are binary images divided into a silhouette region and a background region. For successful SFS, the cameras used to capture the MVIs should be located over a wide range and be aimed in various directions. SFS create an adequate voxel model even if the number of cameras is small and they are arranged sparsely. Here, we describe how the voxel model is created by SFS.

Figure 4 shows the concept of SFS. An object is included in a voxel space and captured by $N_C$ cameras. Let us consider a cone $V_i$, called a visual cone, where $i$ is $1, 2, \ldots, N_C$. The vertex of $V_i$ is the viewpoint of the $i$th image $C_i$, and the silhouette $S_i$ is a cross section of $V_i$. The object exists at the intersection of all visual cones $V_1, V_2, \ldots, V_{N_C}$. Therefore, the approximate shape of the object $H$, called a visual hull, is written by

$$H = V_1 \cap V_2 \cap \cdots \cap V_i \cap \cdots \cap V_{N_C}. \tag{9}$$

This visual hull is the voxel model of the object.

We calculate Eq. (9) by removing any voxels that exist outside of the visual cones from the voxel space. This calculation is implemented by using perspective projection. Here, a 2D point on the $i$th image is denoted by $\boldsymbol{p}_i = [X_i, Y_i]^T$ and a voxel in the voxel space is denoted by $\boldsymbol{M} = [x, y, z]^T$. Using homogeneous coordinates, $\boldsymbol{p}_i$ is rewritten by $\tilde{\boldsymbol{p}}_i = [X_i, Y_i, 1]^T$ and $\boldsymbol{M}$ is rewritten by $\tilde{\boldsymbol{M}} = [x, y, z, 1]^T$. The relationship between the 3D point $\boldsymbol{M}$ and its image projection $\boldsymbol{p}_i$ is given by

$$\tilde{\boldsymbol{p}}_i = \boldsymbol{T}_i \tilde{\boldsymbol{M}}, \tag{10}$$

where $\boldsymbol{T}_i$ is the perspective projection matrix calculated by the camera parameters (focal distance,

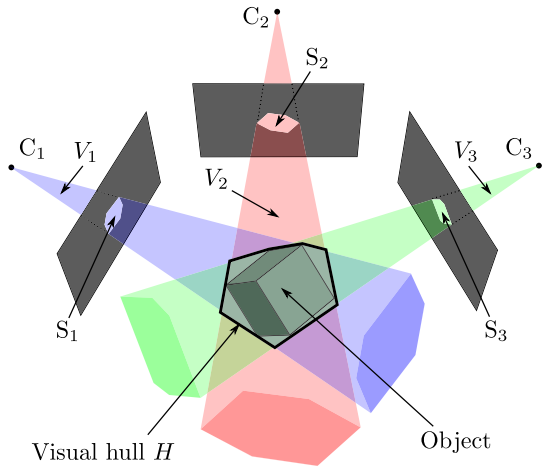



Fig. 4. (Color online) 3D model estimation using SFS.

distortion coefficients, rotation matrix, and translation vector) [22]. When the voxel expressed by $\boldsymbol{M}$ is projected into the silhouette region of the $i$th silhouette image, the voxel exists in a visual cone $V_i$. In contrast, when the voxel is projected into the background region, the voxel exists outside of the visual cone. We obtain the visual hull by taking the following steps for all voxels and silhouette images: (1) calculating Eq. (10), (2) judging whether the voxels are projected into the silhouette region or not, and (3) removing any voxels that exist outside of the visual cones.

There are two steps required before performing SFS. First, we determine the camera parameters by camera calibration [22]. These camera parameters are calculated from the captured images of a calibration pattern, which is like a checkerboard pattern. Second, we apply background subtraction to generate the silhouette images. We capture the calibration images and background images from the same viewpoints of each MVI.

### B. Rendering Elementary Images

After the voxel model is generated by SFS, we generate elementary images by volume rendering. Volume rendering is a technique for visualizing voxel models as 2D projection images. The elementary images are generated by selecting an image, determining the brightness, and volume rendering.

To render elementary images, we need to determine the brightness of the voxels with the captured MVIs. The proposed method assumes all surfaces of the object are perfectly diffuse and uses only a single image to determine the brightness, so first we need to select one of the MVIs. Let $\boldsymbol{n}_H$ be a normal unit vector of a hologram plane and $\boldsymbol{n}_{C_i}$ be a unit vector parallel to the optic axis of an $i$th camera. When we use a $k$th image to determine the brightness, $k$ satisfies the condition that the angle between $\boldsymbol{n}_H$ and $\boldsymbol{n}_{C_k}$ be minimum. The angle $\varphi_i$ between vectors $\boldsymbol{n}_H$ and $\boldsymbol{n}_{C_i}$ is represented as

$$\varphi_i = \cos^{-1}(\boldsymbol{n}_H \cdot \boldsymbol{n}_{C_i}). \tag{11}$$

$\boldsymbol{n}_{C_i}$ is calculated with the camera parameters. The number $k$ of the image to determine the brightness values is therefore written by

$$k = \underset{1 \le i \le N_C}{\operatorname{argmin}}\, \varphi_i. \tag{12}$$

The $k$th image is one of the images in which the angle between the said image and the hologram plane is minimum. We calculate the projection of the voxels into the $k$th image by Eq. (10). When a voxel is projected into the pixel of a $k$th image, let the brightness of the voxel be the brightness of the pixel. Here, $B_V(\tilde{\boldsymbol{M}})$ is the brightness of the voxel denoted by $\boldsymbol{M}$, and $B_I(\tilde{\boldsymbol{p}})$ is the brightness of the point on the image denoted by $\boldsymbol{p}$. Using Eq. (10), $B_V(\tilde{\boldsymbol{M}})$ is described as

$$B_V(\tilde{\boldsymbol{M}}) = B_I(\boldsymbol{T}_k \tilde{\boldsymbol{M}}). \tag{13}$$

The brightness of the voxels are determined by calculating Eq. (13) for all voxels.

Finally, we generate elementary images by volume rendering. Figure 5 shows a schematic diagram of the rendering. An elementary image is rendered by calculating the intersection points between the image plane and rays from each voxel to the center of the elementary hologram. Let $\boldsymbol{O}_E$ be position vectors of the center of the elementary hologram, which is the position of a virtual camera for rendering. Now, consider the ray from the voxel denoted by $\boldsymbol{M}$ to $\boldsymbol{O}_E$. The position vector of the intersection $\boldsymbol{p}_E$ is given by

$$\boldsymbol{p}_E = \boldsymbol{M} + t(\boldsymbol{O}_E - \boldsymbol{M}), \tag{14}$$

where parameter $t$ is written by

$$t = \frac{\boldsymbol{n}_H \cdot (\boldsymbol{M} - \boldsymbol{O}_E) - D_H}{\boldsymbol{n}_E \cdot (\boldsymbol{M} - \boldsymbol{O}_E)}, \tag{15}$$

where $D_H$ is the distance between the hologram plane and the image plane. We regard the brightness of the point denoted by $\boldsymbol{p}_E$ as the brightness of the voxel denoted by $\boldsymbol{M}$. Hence, the elementary image is rendered by calculating Eq. (14) for all voxels on the side of the hologram plane.

### C. Summary

The procedure of the proposed method is as follows.

1. Take pictures of calibration pattern and camera calibration.
2. Capture MVIs with calibrated cameras and generate their silhouette images.
3. Perform SFS using the silhouettes to obtain a voxel model.
4. Consider the position of the hologram plane; i.e., determine the normal unit vector of the hologram plane $\boldsymbol{n}_H$ and the position vector of the center of the hologram plane $\boldsymbol{O}_h$.
5. Determine the brightness of the voxel model.
6. Perform volume rendering to obtain elementary images.

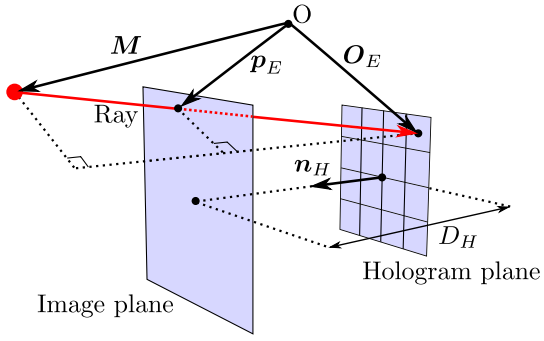



Fig. 5. (Color online) Schematic diagram of volume rendering for proposed method.

7. Generate hologram data from the elementary images using the calculation method described in Section 2.

We render the elementary images from arbitrary positions within the range in which the voxel model and the brightness have been obtained correctly. Even if the number of cameras to capture MVIs for SFS is small, we can obtain a sufficient number of elementary images, meaning that the proposed can method generate CGHs that provide a sufficient motion parallax with just a small number of cameras. Note that the elementary images are rendered from arbitrary viewpoints regardless of whether the cameras are located at the viewpoints or not. When the elementary images are rendered correctly, the hologram planes are also arranged at arbitrary positions regardless of the presence or absence of cameras at that position.

The proposed method has some restrictions stemming from SFS. First, a voxel model created with an extremely small number of cameras has lower quality, so it is necessary to use a certain number of cameras. Of course, the required number of cameras to generate a sufficient quality voxel model is different depending on the object and its position. We discuss the number of cameras when we discuss experimental result. A second restriction is that objects that have concave surfaces are not reconstructed correctly because concave surfaces cannot be reflected into silhouette images. However, various methods to solve this problem have been proposed [18,19]. If these methods are applied, the proposed method can generate holograms of more complex objects constructed by a combination of concave surfaces.

Our proposed method creates a voxel model, renders elementary images, and generates CGHs on the basis of the holographic stereogram. The voxel model is also able to generate CGHs directly. In this case, the phase of the light wave is calculated strictly, so it takes a long time to calculate CGH [23]. The holographic stereogram method has a great advantage in that it can be calculated much faster. Our method is significant in terms of improving this holographic stereogram.

## 4. Experimental Results

In Sections 2 and 3, we described how CGHs are generated using MVIs. In this section, we report the results of optical experiments we performed to determine the effectiveness of the proposed method. Table 1 lists the parameters of the holograms generated in the experiments.

**Table 1. Parameters of Experiments**

| | |
|---|---|
| **Voxel space** | |
| Number of voxels | 512 × 512 × 512 voxels |
| Voxel pitch | 97.7 μm |
| **Elementary images** | |
| Number of pixels | 800 × 600 pixels |
| Pixel pitch | 10.71 μm |
| **Holograms** | |
| Number of pixels | 4096 × 2048 pixels |
| Pixel pitch | 9.5 μm |
| Size | 38.9 mm × 19.4 mm |
| Wavelength | 632 nm |
| Propagation distance | 500 mm |
| **Elementary holograms** | |
| Number of pixels | 128 × 128 pixels |
| Pixel pitch | 9.5 μm |
| Size | 1.216 mm × 1.216 mm |

### A. Experiment I: Comparison with Conventional Method

We compared the reconstructed images of CGHs generated using the proposed method with those generated using the conventional method [5]. In this experiment, we did not use a real existing object but rather a 3D modeled object called "Bunny." A 38.9 mm × 19.4mm hologram plane was located 500 mm away from the object [Fig. 6(a)]. Elementary images were rendered at virtual camera positions by using a volume rendering method, and then CGHs were generated from obtained elementary images.

In the holographic stereogram method, camera intervals were determined by the pixel pitch of the LCD and the size of the pupil. In this case, it requires 512 elementary images captured by cameras arranged on the hologram plane at even intervals of about 1.2 mm. Using the proposed method, a CGH that originally required 512 elementary images is generated from just 14 cameras. The positions of those 14 cameras are shown in Fig. 6(b).

For comparison, we also generated CGHs using the conventional method is also generated. Using the

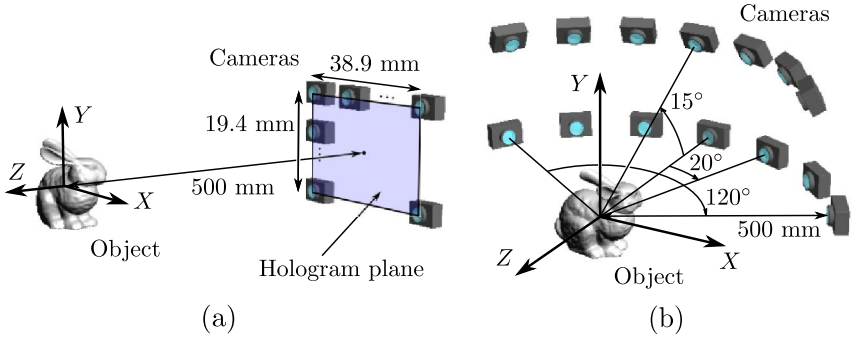


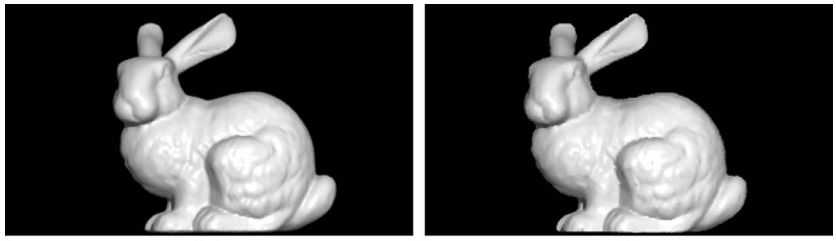

Fig. 6. (Color online) Camera arrangement: (a) for capturing elementary images for experiment I, (b) for SFS.

conventional method, the elementary images were captured by 512 cameras and the CGH was then generated by 512 captured elementary images. A selection of the elementary images captured by the conventional method is shown in Fig. 7(a) and a selection of those rendered by the proposed method is shown in Fig. 7(b). Figure 8 shows magnifications of Fig. 7. A comparison of Figs. 8(a) and 8(b) shows that a part of the elementary images rendered by the proposed method is not correct. We attribute this error to accidental errors during the SFS calculation. To measure the accuracy of the elementary images, we utilized the peak signal-to-noise ratio (PSNR), defined as $20\log_{10}(\sqrt{2^b-1}/\mathrm{MSE})$, where $b$ is the bit depth (in this experiment $b=8$) and MSE is the mean square error between the two images. The PSNR is commonly used as a measure of the quality of compressed images. We measured the PSNR of (1) 512 elementary images captured by cameras and (2) elementary images rendered by the proposed method. The average PSNR of each elementary image was about 34.3 dB. A processed image is generally acceptable if its PSNR is greater than 30 dB [24], so we conclude that the elementary images rendered by the proposed method have sufficient quality.

The reconstructed images of holograms by a computer simulation are shown in Fig. 9. Figures 9(a) and 9(b) show reconstructed images by the conventional method and by the proposed method, respectively. The two reconstructed images are nearly identical in terms of quality. This fact is also indicated by the PSNR measurement. For this result, we consider empirically that the number of cameras used to perform SFS is sufficient to generate CGHs. The key finding is that, with the conventional method, we need to capture MVIs from 512 viewpoints, while with the proposed method, the required number of viewpoints is reduced to just 14.

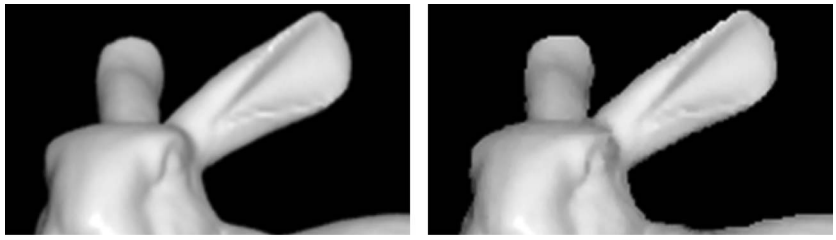
### B. Experiment II: Generating Holograms at Arbitrary Positions

We used the proposed method to generate holograms at multiple positions from a set of MVIs. In this experiment, we used 14 cameras to create a voxel model by SFS. The camera arrangement for SFS was the same as experiment I [shown in Fig. 6(b)].



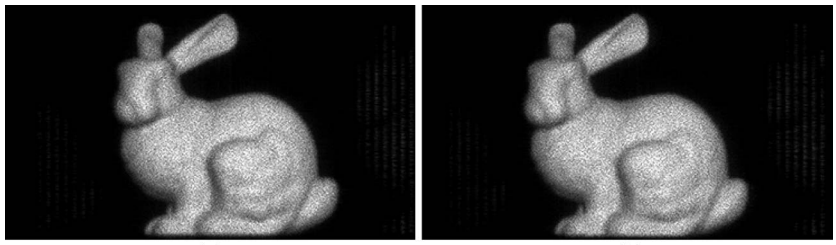
(a) (b)

Fig. 7. Selected elementary images captured of "Bunny": (a) conventional method, (b) proposed method.



(a) (b)

Fig. 8. Magnifications of elementary images of "Bunny": (a) conventional method, (b) proposed method.



(a) (b)

Fig. 9. Reconstructed images of "Bunny" by a computer simulation: (a) conventional method, (b) proposed method.

These 14 cameras were actually simulated by rotating the object using a single camera. We generated four holograms of a real existing object, "Crane." The distance and directions of the holograms are listed in Table 2. Here, $r_H$ denotes the distance between the object and the hologram plane, and $\theta_H$ and $\varphi_H$ denote the direction of the hologram plane (Fig. 10). The elementary images are rendered by 512 virtual camera positions on each hologram plane at the same intervals as experiment I.

To determine if an object had been reconstructed correctly, we performed an optical experiment by displaying CGHs on an LCD. The reconstructed images are shown in Fig. 11. The appearance of the reconstructed objects are different in terms of angle and size depending on the position and direction of each hologram. Figure 12 shows the captured images for comparison with the reconstructed images. These images are captured by cameras located 500 mm from the object and pointed in the same direction as each hologram plane. A comparison of Figs. 11 and 12 shows that the reconstructed images are approximately correct. Note that, in SFS, there are no cameras at the position of $r_H$ being 700 mm and $\theta_H$ being −30°. When the elementary images are rendered correctly, the hologram planes are arranged at arbitrary positions regardless of the presence or absence of cameras in SFS.

We also generated two holograms of a real existing object, "Knight," which includes specular reflection. Figures 13 and 14 show the reconstructed and captured images, respectively, for comparison. In the case of an object that includes specular refraction, the obtained reconstructed images are approximately correct. However, specular reflection was not completely accurate because we assumed all surfaces of the object were perfectly diffuse during the calculation of the brightness.

**Table 2. Distance and Directions of Holograms for Experiment II**

| | $r_H$ | $\theta_H$ | $\varphi_H$ |
|---|---|---|---|
| Hologram 1 | 700 mm | 60° | 15° |
| Hologram 2 | 500 mm | −30° | 15° |
| Hologram 3 | 500 mm | 60° | 0° |
| Hologram 4 | 700 mm | −30° | 0° |

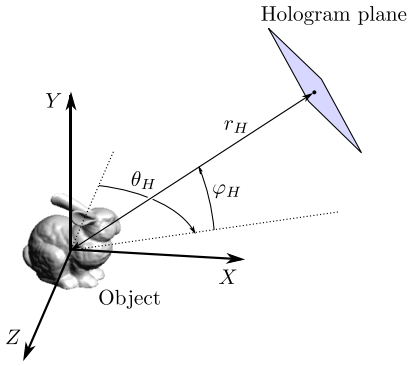



Fig. 10. (Color online) Distance and direction of hologram plane.

## C. Calculation Time

Table 3 shows the calculation time of the proposed method. We used an Intel Core i7 CPU 2600 K with 3.4 and 8 GB RAM. In camera calibration and SFS, the calculation times are the sum of the time spent on computation of 14 cameras, and in volume rendering, the calculation time is the sum of the time spent on the computation of 512 cameras. The calculation times are the average of five measurements, and each process is calculated in parallel on the CPU. In the proposed method, the calculation

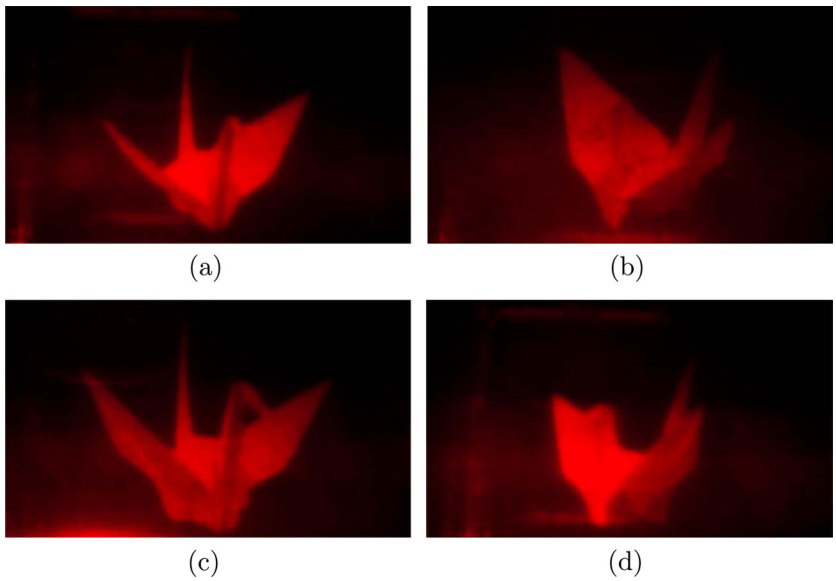



Fig. 11. (Color online) Reconstructed images of "Crane": (a) hologram 1, (b) hologram 2, (c) hologram 3, (d) hologram 4.

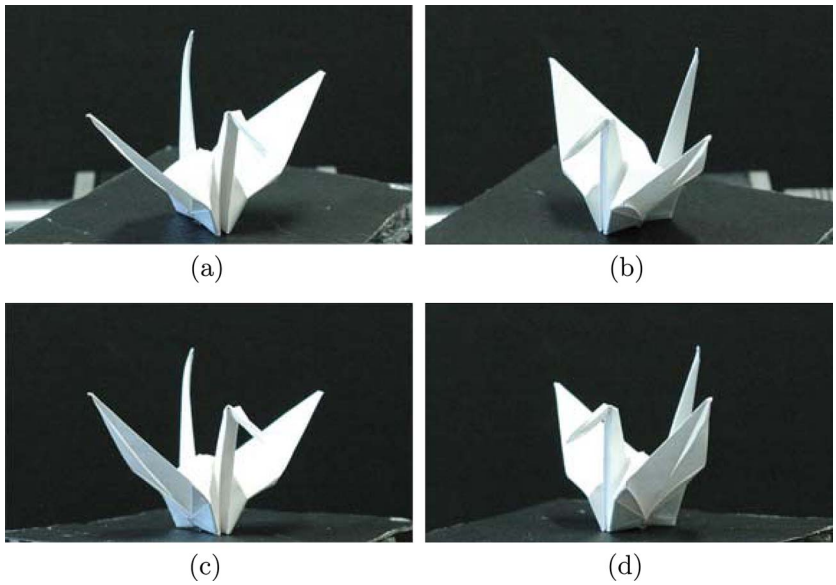



Fig. 12. (Color online) Captured images of "Crane" to compare: (a) with hologram 1, (b) with hologram 2, (c) with hologram 3, (d) with hologram 4.

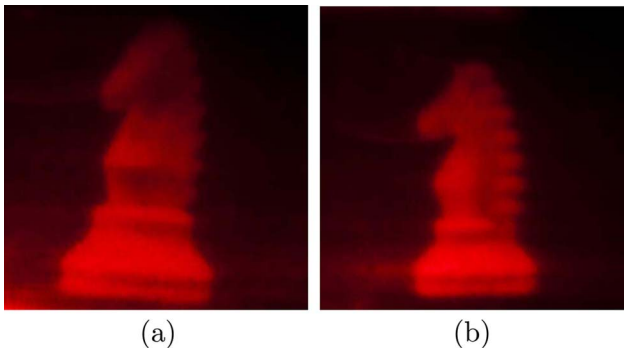



Fig. 13. (Color online) Reconstructed images of "Knight": (a) hologram 3, (b) hologram 4.

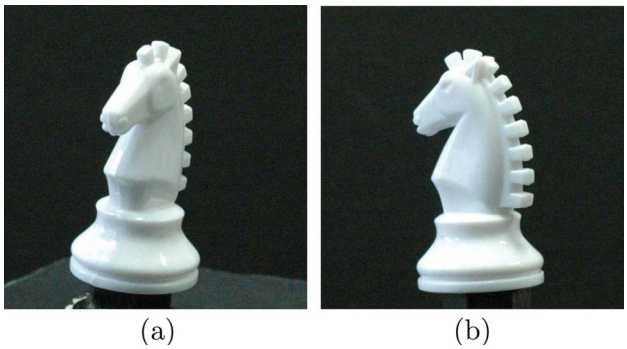



Fig. 14. (Color online) Captured images of "Knight" to compare: (a) with hologram 3, (b) with hologram 4.

**Table 3. Calculation Time of Each Process in Proposed Method**

| | Calculation Time |
|---|---|
| Camera calibration | 3.77 s |
| Shape-from-silhouette | 146 s |
| Volume rendering | 57.2 s |

time for generating CGH is longer than in the conventional method [5] due to the time needed for camera calibration, SFS, and volume rendering. Furthermore, we have to either increase the number of voxels or make the pitch of the voxels finer, depending on the size and shape of the object. In this case, the calculation times increase a lot. We need to develop a method for calculating more rapidly, or use hardware such as a GPU to make the calculations faster.

## 5. Conclusion

In this paper, we discussed an algorithm for generating CGHs using MVIs and proposed a method to generate CGHs using MVIs with a small number of sparsely arranged cameras. In the proposed method, elementary images are not captured by cameras but rather rendered from voxel models created by SFS, which is performed by a small number of cameras. A computer simulation and optical experiments were performed to determine the effectiveness of the proposed method. We generated a CGH from 14 images whose reconstructed images were nearly identical in quality with the reconstructed images of a CGH generated from 512 images by the conventional method. In addition, we demonstrated that the proposed method generate CGHs using MVIs at arbitrary positions. We were able to generate CGHs of real existing objects at multiple positions from just 14 images. This demonstrates that the hologram planes are arranged at arbitrary positions regardless of the presence or absence of cameras at that position when the elementary images are rendered correctly. In the future, we intend to research the relation between accuracy the number and position of objects and cameras, and how to calculate the correct specular reflection.

## References


1. Y. Li, D. Abookasis, and J. Rosen, "Computer-generated holograms of three dimensional realistic objects recorded without wave interference," Appl. Opt. **40**, 2864–2870 (2001).
2. Y. Sando, M. Itoh, and T. Yatagai, "Full-color computer-generated holograms using 3-D Fourier spectra," Opt. Mater. Express **12**, 6246–6251 (2004).
3. T. Mishina, M. Okui, and F. Okano, "Calculation of holograms from elemental images captured by integral photography," Appl. Opt. **45**, 4026–4036 (2006).
4. F. Okano, J. Arai, H. Hoshino, and I. Yuyama, "Three-dimensional video system based on integral photography," Opt. Eng. **38**, 1072–1077 (1999).
5. K. Kushimoto and Y. Sakamoto, "Computer-generated hologram calculated from multi-view images of real existing objects," in *Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/Information Photonics/Signal Recovery and Synthesis Topical Meetings on CD-ROM*, OSA Technical Digest (CD) (Optical Society of America, 2007), paper DWB6.
6. Y. Sakamoto, N. Hayashi, A. Kato, and M. Kinoshita, "3D holo TV system based on multi-view images," *IEEE/ACIS 9th International Conference on Computer and Information Science* (IEEE, 2010), pp. 831–835.
7. T. Yatagai, "Stereoscopic approach to 3-D display using computer-generated holograms," Appl. Opt. **15**, 2722–2729 (1976).
8. M. Kinoshita and Y. Sakamoto, "Computer-generated holograms at an arbitrary viewpoint synthesized from multi-view images," Proc. SPIE **7233**, 72330Z (2009).
9. N. Hayashi, Y. Sakamoto, and Y. Honda, "Improvement of camera arrangement in computer-generated holograms synthesized from multi-view images," Proc. SPIE **7957**, 795711 (2011).
10. B. Katz, N. T. Shaked, and J. Rosen, "Synthesizing computer generated holograms with reduced number of perspective projections," Opt. Express **15**, 13250–13255 (2007).
11. Y. Rivenson, A. Stern, and J. Rosen, "Compressive multiple view projection incoherent holography," Opt. Express **19**, 6109–6118 (2011).
12. H. Baker, "Three-dimensional modelling," Proc. IJCAI **2**, 649–655 (1977).
13. A. Laurentini, "The visual hull concept for silhouette based image understanding," IEEE Trans. Pattern Anal. Mach. Intell. **16**, 150–162 (1994).
14. O. Masatoshi and K. Takeo, "A multiple-baseline stereo," IEICE Trans. **75**, 1317–1327 (1992).
15. T. Hideyuki, I. Toshio, and K. Yukio, "Planes recognition method for polyhedrons from actual shading images using photometric stereo," IEICE Trans. **83**, 1895–1904 (2000).
16. W. Matusik, C. Buehler, R. Raskar, S. Gorlter, and L. McMillan, "Image-based visual hulls," in *Computer Graphics, SIGGRAPH Proceedings* (Association for Computer Machinery, 2000), pp. 369–374.
17. C. Rocchini, P. Cignoni, F. Ganovelli, C. Montani, P. Pingi, and R. Scopigno, "Marching intersections: an efficient resampling algorithm for surface management," in *International Conference on Shape Modeling and Applications* (IEEE, 2001), pp. 296–305.
18. S. M. Seitz and C. R. Dyer, "Photorealistic scene reconstruction by voxel coloring," Int. J. Comput. Vis. **35**, 151–173 (1999).
19. K. N. Kutulakos and S. M. Seitz, "A theory of shape by space carving," Int. J. Comput. Vis. **38**, 199–218 (2000).
20. M. Tarini, M. Callieri, C. Montani, and C. Rocchini, "Marching intersection: an efficient approach to shape-from-silhouette," in *Proceedings of the Vision, Modeling, and Visualization Conference 2002* (AKA GmbH, 2002), pp. 283–290.
21. M. Toyoura, M. Iiyama, K. Kakusho, and M. Minoh, "An accurate shape reconstruction method by integrating visual hulls in time sequences," IEICE Trans. **88**, 1549–1563 (2005).
22. Z. Y. Zhang, "An accurate shape reconstruction method by integrating visual hulls in time sequences," IEEE Trans. Pattern Anal. Machine Intell. **22**, 1330–1334 (2000).
23. Y. Sakamoto and Y. Aoki, "Autostereoscopic visualization of volume data using by computer-generated hologram," IEICE Trans. **86**, 302–309 (2003).
24. T. S. Chen, C. C. Chang, and M. S. Hwang, "A virtual image cryptosystem based upon vector quantization," IEEE Trans. Image Process. **7**, 1485–1488 (1998).