# Image Generation Method Using Weight Maps for Subjective Quality Improvement in Two-Dimensional Image Synthetic Integral Three-Dimensional Display

**Hayato Watanabe[1,2], Jun Arai[1], Chihiro Tsutake[2], Keita Takahashi[2], Toshiaki Fujii[2]**

watanabe.h-fe@nhk.or.jp
[1]NHK, 1-10-11 Kinuta, Setagaya, Tokyo 1578510, Japan
[2]Nagoya University, Furo, Chikusa, Nagoya, Aichi 4648603, Japan
Keywords: integral three-dimensional display, two-dimensional image synthesis, image generation method, weight map

## ABSTRACT

*We propose an image generation method to display three-dimensional (3D) images with high maximum pixel density and improved subjective quality on a two-dimensional image synthetic integral 3D display. In addition to the target light field image, weight maps obtained from the depth information were used to generate the images.*

## 1 Introduction

Integral three-dimensional (3D) displays can display full-parallax 3D images with smooth motion parallax. However, the resolution characteristics of 3D images are low because the maximum number of pixels in a 3D image is limited by the number of elemental lenses that constitute the lens array [1]. Therefore, we are advancing the development of a two-dimensional (2D) image synthetic integral 3D display [2] that can display 3D images with improved resolution characteristics based on the principle of layered 3D displays [3], [4]. This method generates elemental images and 2D image in advance using optimization calculations. The elemental images are used to display an integral 3D image and are displayed on a display device comprising an integral 3D display. The 2D image is used to improve the pixel density of the integral 3D image and is displayed on a 2D display. Three-dimensional images with a maximum pixel density as high as the 2D image can be displayed by optically synthesizing the images displayed on each display using a half-mirror.

In a 2D image synthetic integral 3D display, the improvement effect of the resolution characteristics is larger when a 3D image is displayed at a shallower depth position relative to the display plane of the 2D image. When a 3D image is displayed at a deep depth position, the improvement in the resolution characteristics is less effective, and noise may be visible in the regions around the edges of the displayed image. Therefore, we propose an image generation method that uses weight maps obtained from the depth information of a 3D scene. In the proposed method, elemental images and 2D image are generated such that the resolution characteristics are improved at shallow depth positions relative to the display plane of the 2D image, and such that a general integral 3D image without noise is displayed at deep depth positions. A 3D image with a high maximum pixel density and improved subjective quality can be displayed instead of improving the resolution characteristics at deep depth positions. We verified the effectiveness of the proposed method through display experiments.

## 2 2D Image Synthetic Integral 3D Display

The basic principle of a 2D image synthetic integral 3D display is explained. The display system consists of an integral 3D display, 2D display, and half-mirror. The integral 3D display consists of a display device for displaying of elemental images and a lens array. The lens array is placed at a distance from the display device of $f$, which is the focal length of the lens array. The integral 3D display and 2D display are arranged such that each display surface is at an angle of 45° to the half-mirror surface. Figure 1 shows the configuration when the half-mirror optically synthesizes an integral 3D image and 2D image. $I_{s,t}$ and $V_{s,t}$ are the integral 3D image and target light field image from the viewpoint $s,t$, respectively, and $L$ is the 2D image. The depth positions of the display planes of the integral 3D image and target light field image are identical. The 2D image is spaced from them by a distance $d$. For simplicity, the pixel densities of the elemental images, 2D image, and target light field image are assumed to be identical. One pixel in the integral 3D image is an elemental lens that

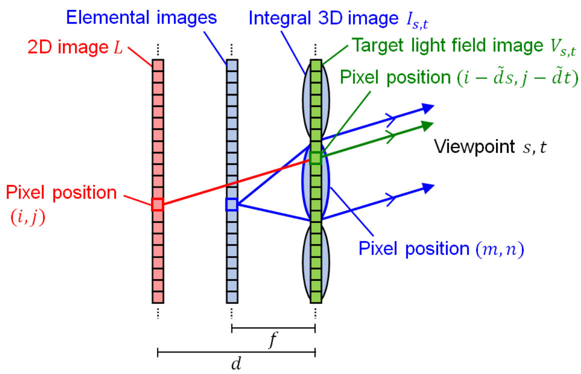


**Fig. 1 Principle of the 2D image synthetic integral 3D display**

constitutes the lens array. The luminance values of the pixels in the integral 3D image are determined by those of the corresponding pixels in the elemental images.

We consider a light ray emitted from the position $(i, j)$ in the 2D image and traveling in the direction of viewpoint $s, t$. This light ray passes through the position $(i - \tilde{d}s, j - \tilde{d}t)$ in the integral 3D image and target light field image. $\tilde{d}$ is the amount of disparity proportional to the distance $d$. Hereafter, the luminance value of the 2D image at the position $(i, j)$ is denoted as $L(i, j)$. The luminance values of the integral 3D image and target light field image at the position $(i - \tilde{d}s, j - \tilde{d}t)$ are denoted as $I_{s,t}(m, n)$ and $V_{s,t}(i - \tilde{d}s, j - \tilde{d}t)$, respectively. Based on the above configuration, the elemental images and 2D image are generated according to

$$\underset{L(i,j), I_{s,t}(m,n)}{\operatorname{argmin}} \left[ \sum_{s,t,i,j} \left\{ L(i,j) + I_{s,t}(m,n) - V_{s,t}(i - \tilde{d}s, j - \tilde{d}t) \right\}^2 \right]. \quad (1)$$

A 3D image with the luminance values that approximate the target light field image can be displayed by displaying the generated elemental images and 2D image on the integral 3D display and 2D display, respectively.

## 3 Image Generation Method Using Weight Maps

Figure 2 shows a flowchart of the proposed image generation method. The proposed method differs from the previous method in using weight maps for image generation. A target light field image and weight maps are generated from the known 3D models. Subsequently, the pixel values of the 2D image are set to zero. Afterward, the process becomes iterative. The elemental images are generated under the condition that the 2D image is known. Subsequently, the 2D image is generated under the condition that the elemental images are known. These processes are iterated for a predetermined number of cycles. The elemental images are generated according to the following equation obtained from Eq. (1):

$$I_{s,t}(m,n) = \frac{\sum_{s,t} V_{s,t}(i - \tilde{d}s, j - \tilde{d}t)}{N} - L(i,j), \quad (2)$$

where $N$ is the number of viewpoints. Subsequently, the 2D image is generated according to the following equation, which is the multiplication of the equation obtained from Eq. (1) with weights in the weight maps:

$$L(i,j) = \sum_{s,t} \Big\{ W_{s,t}(i - \tilde{d}s, j - \tilde{d}t) \cdot \frac{V_{s,t}(i - \tilde{d}s, j - \tilde{d}t) - I_{s,t}(m,n)}{N} \Big\}, \quad (3)$$

where $W_{s,t}$ is the weight map from the viewpoint $s, t$, and $W_{s,t}(i - \tilde{d}s, j - \tilde{d}t)$ is the weight at position $(i - \tilde{d}s, j - \tilde{d}t)$. The error in the luminance values between the displayed 3D image and target light field image is gradually reduced by iterating these image generation processes. The luminance must be positive. Therefore, the process of maintaining the luminance values of $I_{s,t}(m,n)$ and $L(i,j)$ within the range of 0 to the upper limit is performed each time the image is generated.

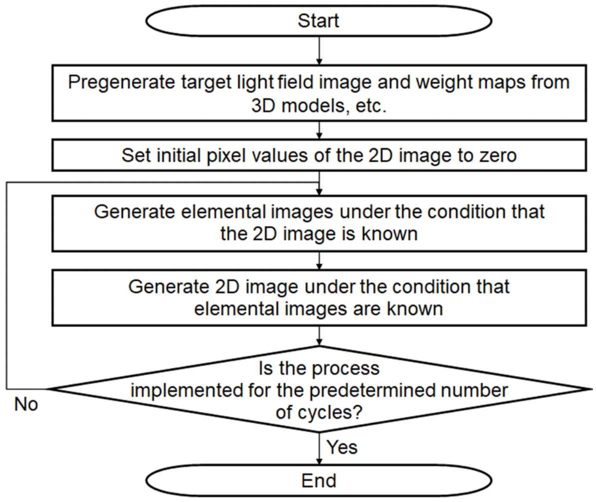


**Fig. 2 Flowchart of the image generation process**

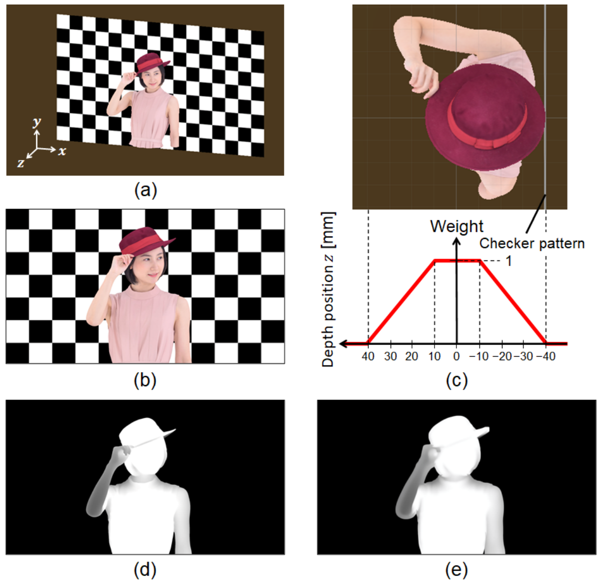


**Fig. 3 Generation of the target light image and weight maps: (a) 3D scene. (b) Target light field image from the center view. (c) Weight distribution. Weight map from the center view (d) before and (e) after the dilation process**

The generation method for the target light field image and weight maps is explained. Figure 3(a) shows an example of the 3D scene to be displayed. The plane on which the texture of the checker pattern is mapped is placed at a depth position of −40 mm. The 3D model of the woman is placed such that the eyes are at a depth position of 0 mm. For simplicity, the distance $d$ is set to 0 mm and the depth positions of the display planes of the integral 3D image, 2D image, and target light field image are the same as that of the $xy$-plane. Figure 3(b) shows the target light field image at the center viewpoint generated by orthographic projection. Similar images for all viewpoints $s, t$ are generated using oblique projection as the target light field image. Figure 3(c) shows an example of the weight distribution for generation of weight maps. The weight range is 0 – 1. The weight distribution should be set such that the weights are higher near the display plane of the 2D image. Figure 3(d) shows the weight map at the center viewpoint generated using the weight distribution. The weights of 0 – 1 were scaled to correspond to pixel values of 0 – 255. When this weight map is used to generate the elemental images and 2D image, the improvement effect of resolution characteristics on pixels of integral 3D images that straddle the contours of the 3D model of the woman becomes small. In this regard, it is necessary to dilate the regions of high weights by the size of few elemental lenses. We refer to this process as dilation process. Figure 3(e) shows the weight map after the dilation process. The dilation process can be performed, using the following spatial filter:

$$W_{\mathrm{out}}(x,y) = \frac{\sum_{b=-s}^{s}\sum_{a=-s}^{s}\left(W_{\mathrm{in}}(x,y)\cdot G(a,b)\cdot D(x,y,a,b)\right)}{\sum_{b=-s}^{s}\sum_{a=-s}^{s}\left(G(a,b)\cdot D(x,y,a,b)\right)},$$

$$G(a,b) = \exp\left(-\frac{a^2+b^2}{2{\sigma_1}^2}\right), \tag{4}$$

$$D(x,y,a,b) = \exp\left(-\frac{\{W_{\mathrm{in}}(x,y) - W_{\mathrm{in}}(x+a,y+b) + v_W\}^2}{2{\sigma_2}^2}\right),$$

where, $W_{\mathrm{in}}$ and $W_{\mathrm{out}}$ are the weight maps before and after the dilation process, respectively; $s$ is the filter size; $\sigma_1$ and $\sigma_2$ are the parameters to control the dilation process; and $v_W$ is the upper limit of the weight. For example, $s$, $\sigma_1$, and $\sigma_2$ are set to a size corresponding to several element lenses, $s/2$, and 0.25, respectively. Similar weight maps are generated for all viewpoints $s, t$. The elemental images and 2D image are generated using the target light field image and weight maps after the dilation process based on Eqs. (2) and (3).

## 4 Experiment

Image generation and display experiments were carried out to verify the effectiveness of the proposed method. The specifications of the prototype display system are listed in Table 1. In the experiments, the same 3D scene as in Fig. 3(a) was used to generate multi-viewpoint images with a resolution of approximately 4K. These images were used as the target light field image. The sampling positions of the weight maps were set to be the same as the pixel positions of the target light field image. $s$, $\sigma_1$, and $\sigma_2$ were set to 30, 15.0, and 0.25, respectively. Figure 4 shows the elemental images and 2D images generated by the previous method [2] and proposed method. These images were generated by applying the gamma correction process described in the literature of the previous method. The image generation process described in Section 3 was iterated 30 times to almost converge the error in the luminance values. In the 2D image generated by the previous method, the regions around the edges were brightened over the entire image. In contrast, in the 2D image generated by the proposed method, only the regions corresponding to the edges in the 3D model of the woman were brightened, while the regions corresponding to the checker pattern were a black image.

Figure 5 shows the results of displaying 3D images. First, 3D images were displayed in a virtual space using a simulation. In the general integral 3D image without synthesizing the 2D image, the resolution characteristics of both the checker pattern and 3D model of the woman were low. According to the results of the previous method, the resolution characteristics of the 3D image were improved in regions with shallow depth positions,

Table 1 Specifications of the prototype display system

| | | |
|---|---|---|
| 2D display | Number of pixels | 3840 × 2160 pixels |
| | Pixel pitch | 55.5 µm |
| | Depth distance from lens array $d$ | 0 mm |
| Integral 3D display | Number of pixels of elemental images | 3840 × 2160 pixels |
| | Pixel pitch of elemental Images | 55.5 µm |
| | Focal length of lens array $f$ | 8.846 mm |
| | Lens pitch | 1.38 mm |
| | Lens shape / arrangement | Hexagon / honeycomb |
| | Viewing angle | 8.9°(H) × 10.3°(V) |

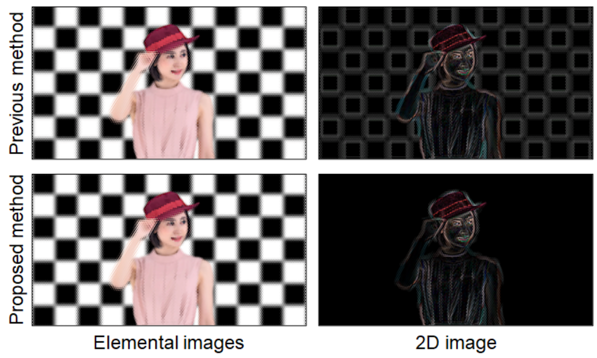


**Fig. 4 Results of generation of elemental images and 2D images**


**Fig. 5 Results of displaying 3D images**

such as the woman's face. The improvement in the resolution characteristics was small in regions with deep depth positions, such as the checker pattern. In addition, noise occurs in regions slightly away from the edges of the checker pattern instead of improving the resolution characteristics of the edges of the checker pattern, although this is challenging to notice in the figure. For example, the region indicated by the arrow in the figure was slightly darkened, even though it should be white. In this study, such unnatural luminance changes that occur in the regions around the edges in the 3D image displayed by the 2D image synthetic integral 3D display are described as noise. According to the results of the proposed method, the resolution characteristics were improved almost as much as those of the previous method at shallow depth positions. In regions with deep depth positions, such as the checker pattern, the general integral 3D image without noise was displayed. Subsequently, 3D images were displayed by the prototype display system, and a trend similar to that of the simulation results was obtained. However, the resolution characteristics were low at deep depth positions because a diffusing film was placed inside the integral 3D display to reduce the color moiré [5]. Therefore, the checker pattern was observed blurred compared with the simulation results.

As described above, we confirmed that 3D images with high maximum pixel density and improved subjective quality could be displayed with the proposed method. However, the improvement effect of the resolution characteristics at deep depth positions cannot be obtained. As a result, the quantitative error in the luminance values between the displayed 3D image and target light field image is larger with the proposed method than with the previous method. It is desirable to decide whether to generate images using the previous method or proposed method, depending on the specifications of the display system and configuration of 3D scenes. In the future, further quantitative and subjective evaluations of displayed 3D images will be carried out to optimize the image generation method for a 2D image synthetic integral 3D display.

## 5 Conclusions

We propose an image generation method for a 2D image synthetic integral 3D display. In the proposed method, elemental images and 2D image are generated using weight maps obtained from the depth information of a 3D scene, in addition to a target light field image. When a 3D image is displayed at a shallow depth position relative to the display plane of the 2D image, the resolution characteristics can be improved. In contrast, when a 3D image is displayed at a deep depth position, a general integral 3D image without noise can be displayed instead of improving the resolution characteristics. We confirmed that 3D images could be displayed according to the principle through display experiments. With the proposed method, the display quality of 3D images can be controlled by adjusting the weight distribution in the depth direction. We intend to improve and evaluate the image generation method for a 2D image synthetic integral 3D display to display high-quality 3D images.

## References


[1] H. Hoshino, F. Okano, H. Isono, and I. Yuyama, "Analysis of resolution limitation of integral photography", J. Opt. Soc. Am. A, vol. 15, no. 8, pp. 2059–2065 (1998).

[2] H. Watanabe, J. Arai, C. Tsutake, K. Takahashi, and T. Fujii, "Pixel-density enhanced integral three-dimensional display with two-dimensional image synthesis," Opt. Express, vol. 30, no. 20, pp. 36038–36054 (2022).

[3] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar, "Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting," ACM Trans. Graph., vol. 31, no. 4, pp. 1–11 (2012).

[4] S. Lee, C. Jang, S. Moon, J. Cho, and B. Lee, "Additive light field displays: realization of augmented reality with holographic optical elements," ACM Trans. Graph., vol. 35, no. 4, pp. 1–13 (2016).

[5] M. Okui, M. Kobayashi, J. Arai, and F. Okano, "Moire fringe reduction by optical filters in integral three-dimensional imaging on a color flat-panel display," Appl. Opt., vol. 44, no. 21, pp. 4475–4483 (2005).