IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 1, 2020, accepted April 29, 2020, date of publication May 6, 2020, date of current version May 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2992511

A Real-Time Super Multiview Rendering Pipeline for Wide Viewing-Angle and High-Resolution 3D Displays Based on a Hybrid Rendering Technique

SHUJUN XING[1], XINZHU SANG[2], LIANGCAI CAO[1], YANXIN GUAN[2], AND YUANHANG LI[2]

[1]State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University, Beijing 100084, China
[2]State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China

Corresponding author: Xinzhu Sang (xzsang@bupt.edu.cn)

This work was supported in part by the National Key Research and Development Program under Grant 2017YFB1002900, in part by the National Natural Science Foundation of China (NSFC) under Grant 61575025, and in part by the Fund of the State Key Laboratory of Information Photonics and Optical Communications.

ABSTRACT Super multiview (SMV) 3D displays with a directional diffuser can provide dense viewpoints, wide viewing angles and high image quality. However, traditional SMV 3D rendering pipelines have many disadvantages, such as high time costs, low image quality, incorrect lighting effects and limited viewing angles. To overcome the abovementioned drawbacks, we propose a new SMV 3D rendering pipeline based on a hybrid rendering technique (HRT) that effectively combines and improves the Turing multiview rendering technique, the multiple view plus depth (MVD) technique and the deferred shading technique, introducing additional normal information, diffuseness information and shininess information into the SMV 3D rendering pipeline. The SMV HRT pipeline can generate SMV 3D images with accurate lighting in real time when the viewpoint number of an SMV 3D display is greater than 50, the viewing angle is greater than 100 degrees, the resolution of a single viewpoint image is more than 512*512 and the resolution of the LCD panel is 7680*4320; in particular, complex scenes can be generated in real time. The proposed HRT pipeline is the only viable real-time rendering pipeline with accurate lighting effects for super multiview 3D displays with wide viewing angle and high resolution that runs on a common PC with standard graphics hardware. An 8k SMV 3D display with an 80° viewing angle and 100 viewpoints is applied to demonstrate the validity of the HRT method in experiments. When the number of reference views is four, the experimental results show that the rendering frame rate is more than 35 fps on average for complex virtual scenes, and the peak signal-to-noise ratio (PSNR) value of the HRT is greater than 36 dB, illustrating its high image quality and accurate lighting. Compared with state-of-the-art MVD methods, HRT processes more lighting information and is more suitable for virtual scene rendering. Therefore, the HRT method has advantages regarding the image quality of the reconstructed 3D image and can easily achieve real-time rendering for SMV 3D displays with wide viewing angles.

INDEX TERMS 3D display, super multiview display, real-time rendering, wide viewing angle, computer-generated 3D image.

I. INTRODUCTION

Recently, there have been rapid advancements in 3D display techniques. Three-dimensional displays with a directional diffuser [1]–[4] can enhance the 3D image quality and effectively increase the viewing angle to 100 degrees. To provide a continuous 3D image, the viewpoint number of the 3D displays must be at least 50 [5]. With increasing viewpoint number and viewing angle, traditional real-time computer-generated image methods for SMV 3D displays are facing great challenges. The traditional methods for SMV rendering can be divided into two broad categories.

The associate editor coordinating the review of this manuscript and approving it for publication was Jenny Mahoney.

A. GEOMETRY-BASED RENDERING METHODS

The input data of the geometry-based rendering methods are 3D geometry data, texture, etc. The 3D image is generated by the methods without a complex process.

The simplest geometry-based rendering method is each camera viewpoint-independent rendering (ECVIR) [6], which sets up one camera for each viewpoint and renders all the viewpoint images in turn. Consequently, with increasing viewpoint number, the rendering efficiency of this method decreases quickly. The multiview rendering (MVR) [7] method first renders scenes to epipolar plane images and then transforms them into viewpoint images. MVR can promote rendering efficiency in theory, but it fails to be supported by standard computer graphics rendering engines. The backward ray tracing (BRT) method [8] for the SMV display is difficult to render in real time due to the many calculations for ray-object intersections. Although great progress has been made in the real-time ray tracing technique, it is impossible to achieve real-time ultrahigh resolution rendering for complex virtual scenes on a PC machine within the next decade according to an NVIDIA report [9].

B. IMAGE-BASED RENDERING METHODS

The input data of image-based rendering methods is an array of images. The common image-based rendering methods for SMV displays are volume rendering and light field rendering. Volume rendering based on the ray-casting technique [10] has the same problem as the BRT method. Light field rendering for SMV requires considerable graphics memory, so it is impractical for rendering a large virtual scene in real time. However, the depth image-based rendering (DIBR) technique [11]–[19] is a promising 3D rendering method only for traditional SMV 3D displays with viewing angles of less than 10 degrees. In addition, the DIBR method generates images with poor quality and has difficulties with filling the holes and dealing with light in complex scenes. Multiple view plus depth (MVD) 3D representation carries both multiple view color and partial geometry information of the scene (carried by the multiple view depth maps) that can be used in combination to render image data for an SMV 3D display.

Recently, many works on MVD have focused on removing holes and promoting image quality. The Gaussian mixture modeling (GMM) method to realize virtual view synthesis for MVD [20] is a valid approach for addressing the issue of image quality degradation. Emerging holes in a target virtual view can be greatly alleviated by making good use of other neighboring complementary views in addition to the two closest neighboring primary views [21]. Reference [22] presents the multiview video plus depth retargeting (MVDRT) technique for stereoscopic 3D displays, which takes shape preservation, line bending and visual comfort constraints into account and simultaneously optimizes the horizontal, vertical and depth coordinates in the display space. Reference [23] uses color correction of reference views and combines depth-based image fusion with direct color image fusion to decrease ghost effects. Additionally, the cracks are filled using depth filtering and inverse warping. To accelerate the generation of new viewpoint images, many references [24], [25] adopt a parallel computing scheme.

Although the MVD methods effectively remove holes and promote image quality, the lighting problem is not considered, so the newly generated viewpoint image has some color distortions and inaccurate lighting, which leads to degradation of the image quality. There are three different types of light sources in virtual scenes: directional lights, point lights and spotlights. A material in a virtual scene should contain four components: ambience, diffuseness, specularity and shininess. This does not strictly require that the scenes in the above MVD methods only consider the ambient light and diffuseness of material to illustrate higher PSNR than other algorithms. This may be suitable for the same special video sequences, but the image quality is unacceptable for virtual scene rendering.

In computer graphics, lighting is very important for rendering [26] and is based on a simple model of the interaction of materials and light sources. There are three familiar lighting models for real-time rendering – Phong [27], Blinn-Phong [28] and Cook-Torrance [29]. Forward shading is a straightforward approach where we render an object, light it according to all the light sources in a scene and then render the next object and repeat this sequence for each object in the scene. This process is quite computationally intensive as each rendered object has to iterate over each light source for every rendered fragment, which is a considerable number of steps. Deferred shading [30] overcomes this issue and is widely used in games and interactive 3D programs. This method consists of a geometry pass and lighting pass. The geometry pass is used to retrieve all kinds of geometric information from the objects that are stored in a collection of textures. The lighting pass calculates the lighting for each fragment using the geometric information stored in the textures. The lighting pass is an image-based rendering method. The great advantage of deferred shading in comparison with traditional rendering algorithms is that this method has the worst-case computational complexity $O(N_o + N_l)$, where N_o and N_l denote the number of objects and the number of light sources, respectively.

In computer graphics, hybrid rendering is a common and important technique to solve rendering problems, such as rendering quality and efficiency. For instance, a hybrid rendering method exists that combines a color-coded surface rendering method and a volume rendering method, exploits the advantages of both rendering methods, provides an excellent overview of the tracheobronchial system and allows clear depiction of the complex spatial relationships of anatomical and pathological features [31]. A multiview rendering hardware architecture consisting of hybrid parallel DBIR and pipeline interlacing is proposed to improve the performance in [32]. The proposed multiview rendering architecture can achieve 60 frames per second for processing full HD (1920×1080) video in a real-time processing system. A hybrid algorithm [33] is presented for accurately and efficiently rendering hard shadows that combines the strengths of shadow maps and shadow volumes. This approach simultaneously avoids the edge aliasing artifacts of standard shadow

maps and avoids the high fillrate consumption of standard shadow volumes.

The hybrid rendering technique, which combines rasterization and real-time ray tracing techniques, has made great progress since 2018 since the revolutionary NVIDIA Turing^{TM}architecture [34] was proposed. Barré-Brisebois *et al.* [35] proposed a hybrid rendering pipeline in which rasterization, computation, and ray tracing shaders work together to enable real-time visuals to approach the quality of offline path tracing in 2019.

Here, a whole new SMV rendering pipeline based on a hybrid rendering technique is presented to address the problems that exist in all the previous SMV rendering methods. The proposed method introduces additional normal information, diffuseness information and shininess information and exploits the advantages of ECVIR of superior quality 3D images without viewing angle limitations, the high rendering efficiency of the MVD technique and the perfect lighting effect of deferred shading and can be treated as a hybrid rendering technique.

The HRT rendering pipeline contains four steps. First, images of sparse reference viewpoints are generated. Then, we apply multiple view reprojection and hole-filling to generate images of new viewpoints. The target view of images (e.g., depth images, normal images, diffuseness images and specular images) can composite target color images with an accurate lighting effect using the deferred shading technique. Finally, the reconstructed 3D image is generated according to the viewpoint arrangement of the SMV 3D display.

The remainder of the paper is organized as follows. In section II, a new SMV rendering pipeline based on the HRT method is proposed. Then, the principles of generating SMV 3D images with large viewing angles are illustrated. In section III, we carry out experiments to demonstrate the validity of the HRT method. Finally, we conclude our work in Section IV.

II. THE HRT SMV RENDERING PIPELINE AND PRINCIPLES OF GENERATING SMV 3D IMAGES

The HRT SMV rendering pipeline is shown in Figure 1, which contains four stages to render a one frame SMV 3D image: sparse reference viewpoint image generation, dense viewpoint image generation, deferred shading and image synthesis. The former two stages increase the viewing angle and promote rendering efficiency. Deferred shading is used to generate an accurate lighting effect for every viewpoint image. The image synthesis stage generates the reconstructed 3D image according to the parameters of the SMV 3D display.

A. SPARSE REFERENCE VIEWPOINT IMAGE GENERATION

The first stage applies the render-to-texture technique [36] and programming shaders for generating sparse reference viewpoint images, including depth images, normal images, diffuseness images and specular images. The common render-to-texture techniques to create multiview images have two types: single pass stereo (SPS) and Turing multiview rendering (TMVR). TMVR is a new technique that can simultaneously generate four viewpoint images, and its rendering speed is 2-3 times that of the SPS technique [9]. Therefore, we choose the TMVR technique to generate reference viewpoint images and store them in GBuffers, as shown in Figure 1. The traditional usage of TMVR is only to generate two viewpoint color images for VR devices, but we used it to generate four kinds of images for sparse reference viewpoints.

There are two kinds of input data in this stage, including a 3D virtual scene and an N-view virtual camera array. As shown in Figure 2, each viewpoint image can be generated from a translational-offset virtual camera, which corresponds to different off-axis asymmetric sheared view frustums with parallel view directions. Every virtual camera property can be determined by its view matrix and perspective matrix, and the n-th virtual camera view matrix M_{vn} and projection matrix M_{pn} in the camera array can be determined by Equation (1) [37]. The view matrix is applied to set the position and direction of the virtual camera. The projection matrix is used to project 3D world objects in homogeneous coordinates into an image. M_{vc} and M_{pc} are the view matrix and projection matrix of the center virtual camera in the virtual camera array, respectively. Because the virtual cameras in different positions have the same direction, their view matrix can be obtained by multiplying the translation matrix M_T and

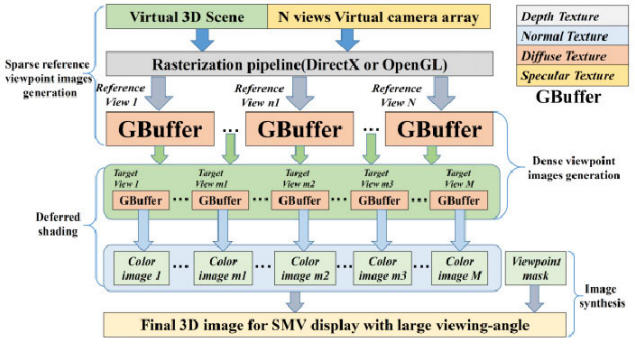


FIGURE 1. The SMV rendering pipeline based on the HRT method.

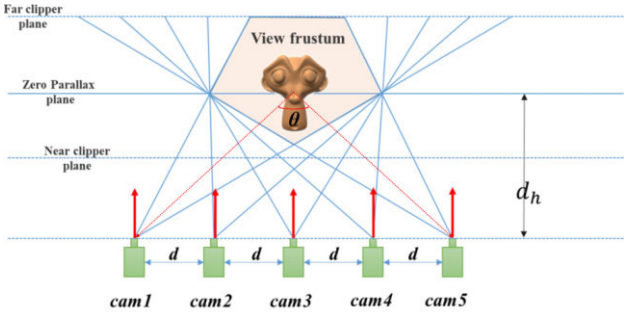


FIGURE 2. The arrangement of a virtual camera array.

the view matrix M_{vc}. The translation matrix is determined by the distance between adjacent virtual cameras and the index of the virtual camera. The projection matrix of each virtual camera can be obtained by multiplying M_{pc} and the shear matrix M_{shear}.

$$M_{vn} = M_T M_{vc} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ (n - N/2)d & 0 & 0 & 1 \end{pmatrix} M_{vc}$$

$$M_{pn} = M_{pc} M_{shear} = M_{PC} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ (n - N/2)d/d_h & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1}$$

where d represents the distance between adjacent cameras and d_h is the distance between the virtual camera array and the zero parallax plane.

Assuming that the viewing angle of the SMV 3D display is θ, d can be calculated by the following equation:

$$d = \begin{cases} 2d_h \tan\frac{\theta}{2}/(N-1) & N \text{ is odd} \\ 2d_h \tan\frac{\theta}{2}/N & N \text{ is even} \end{cases} \tag{2}$$

Four images as one group are generated by a virtual camera. By using these images as intermediate results, the viewpoint color image created with the deferred shading technique is illustrated in the fourth stage.

B. DENSE VIEWPOINT GENERATION

In this stage, dense target viewpoint images can be generated by sparse reference viewpoint images. We improved the MVD technique. The difference between the traditional MVD technique and the stage is that the latter has to process depth, normal, diffuseness and shininess information, whereas the former only has to process depth and color information. The precision of depth in the traditional MVD technique is 8 bits, while the precision of depth in the stage is 32 bits. The basic approach of generating new viewpoint images includes view reprojection and hole filling. In the first step, the points in the reference view are projected into a 3D space and then projected to a new view (reprojection view). Holes are introduced, which decreases the reprojection view image quality. However, a texture map from other reference views is developed to inpaint most of the holes and avoid degrading the 3D image quality. In the hole-filling step, the remaining holes are filled by linear interpolation.

Each pixel in the depth image contains one virtual point, and its position $\vec{p}$ can be determined by the following equation:

$$\vec{p} = (u, v, d, 1) M_{pn}^{-1} M_{vn}^{-1} \tag{3}$$

where (u, v, d, 1) represents the homogeneous coordinates under the normalized device coordinate system, (u, v) represents the texture coordinates, and d is the depth value.

The reprojection view can be obtained by shifting the image value in the horizontal direction according to the depth value. As illustrated in Figure 3(c), there are two virtual points (p_1,p_2) with the same projection point p_j. However, they have different depth values (d_1,d_2) in the reprojection view. The shift values of Δs_1 and Δs_2 can be calculated with the principles of similar triangles.

$$\begin{aligned} \Delta s_1 &= d_1 \Delta x/(d_h - d_1) \\ \Delta s_2 &= d_2 \Delta x/(d_h - d_2) \end{aligned} \tag{4}$$

where Δx is the distance from the reference camera to the target camera. Because the depth of p_1 is larger than p_2, and the shift value Δs_1 is larger than Δs_2, the final value of p_1 in the other images is saved at p_2. Therefore, occlusions can be solved effectively.

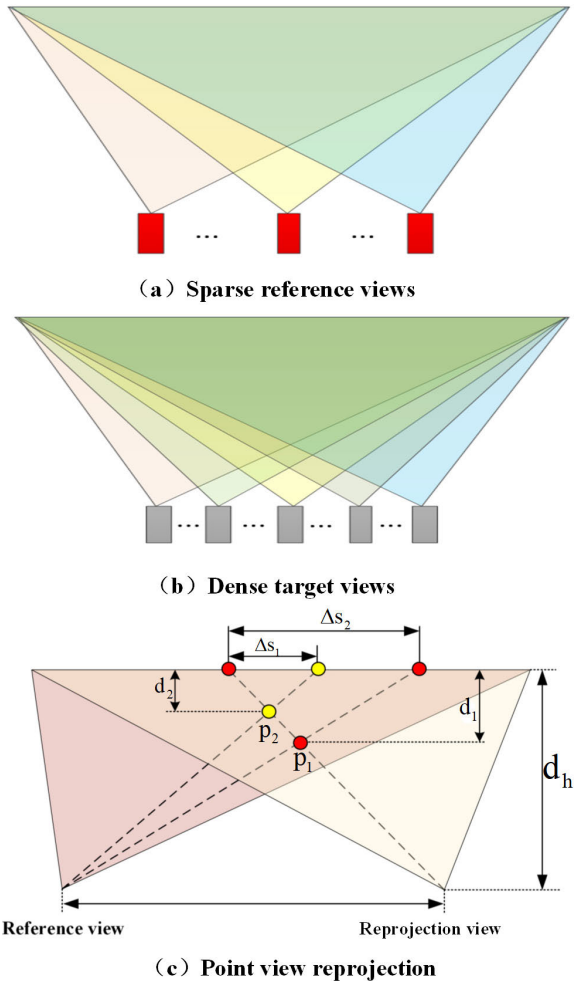


FIGURE 3. View reprojection.

There are N reference views in our display system, so each viewpoint image in the dense target views has N reprojection views. With increasing Δx, the holes caused by view reprojection become obvious. However, the target view can be synthesized by the N reprojection views, and most of the

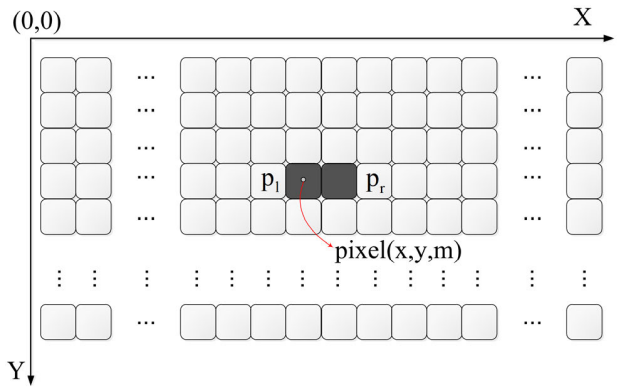


FIGURE 4. Hole-filling step, black part represents holes.

holes generated from one view reprojection image can be inpainted in by other view reprojection images.

Linear interpolation is the simplest approach for hole filling. Assuming that pixel (x,y) in target view **m** is a hole pixel, the nearest left valid pixel p_l and the nearest right valid pixel p_r are shown in Figure 4. Therefore, the filling value for pixel (x, y, m) can be calculated from the following equation:

$$V_{(x,y,m)} = (d_r V_l + d_l V_r)/(d_l + d_r) \tag{5}$$

where d_r is the distance between p_r and pixel (x, y, m), d_l is the distance between p_l and pixel (x, y, m), and V_l and V_r represent the values of pixels p_l and p_r, respectively. The values of the pixel can include depth, normal, diffuseness and specular values.

Parallel computing is implemented in this stage to provide the real-time performance of view reprojection and hole-filling. Given that the resolution of a target view is W×H and the number of reference views is M, M×W×H threads are set up in a computer shader in Figure 5. Thread (x, y, m) first reads the image value of the index (x, y) pixel from the n-th reference view. Then, the thread writes to the corresponding pixel in the m target view according to Equation (4). If there are simultaneously several threads for writing the same pixel

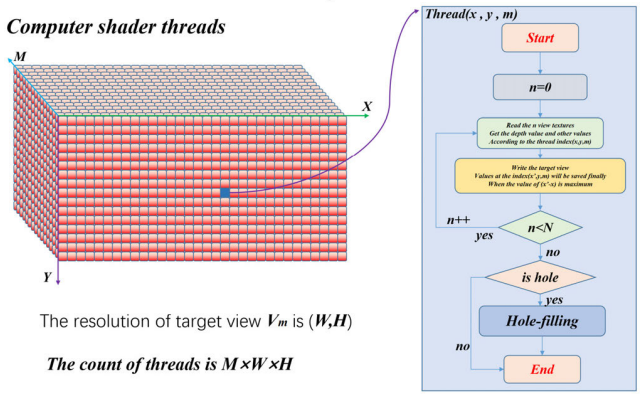


FIGURE 5. Super multiview image generation running in a computer shader.

FIGURE 6. Color texture is generated from the other four images with the deferred shading technique.

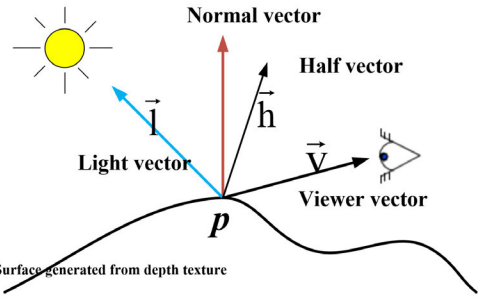


FIGURE 7. The parameters in the different shading stages.

in the m target view, the value with the minimum of d_n is ultimately saved. The value of the hole pixel can be calculated from Equation (5).

C. DEFERRED SHADING STAGE

In this stage, the color image of a viewpoint and the correct lighting effects can be generated from its four images, as shown in Figure 6. The lighting model used in this stage is the Blinn-Phong model, which is the most popular model used in interactive 3D programs. The following text is from [38], illustrating the detailed process of deferred shading.

For convenience, we only consider point lights, as shown in Figure 7. The total lighting equation of the Blinn-Phong model is shown in Equation (6).

$$I_{tot} = I_{amb} + I_{diff} + I_{spec} \tag{6}$$

The diffuseness term I_{diff}, for the diffuse components of both light sources and materials, can be computed as follows:

$$I_{diff} = MAX(0, \vec{l} \bullet \vec{n}) M_{diff} \otimes S_{diff} \tag{7}$$

where the operator ⊗ is used to perform component-wise multiplication, $\vec{l}$ represents the unit vector from light source position S_{pos} to the virtual point $\vec{p}$, $\vec{n}$ is the surface normal on $\vec{p}$, M_{diff} denotes the diffuseness of the material color, and S_{diff} is the light source color. The values of $\vec{p}$, M_{diff} and $\vec{n}$ can be obtained from depth images, diffuseness images, and normal images, respectively.

The specular term is a key component in determining the brightness of specular highlights, along with shininess to determine the size of the highlights. The specular term I_{spec}, which represents the specular parts of both light sources and materials can be computed as follows:

$$I_{spec} = MAX(0, \vec{h} \bullet \vec{n})^{M_{shi}} M_{spec} \otimes S_{spec} \tag{8}$$

where $\vec{h}$ is the unit half vector between $\vec{l}$ and $\vec{v}$ and $\vec{v}$ is the view vector from the point $\vec{p}$ to the viewer.

$$\vec{h} = \frac{\vec{l} + \vec{v}}{\left\| \vec{l} + \vec{v} \right\|} \tag{9}$$

The ambient term I_{amb} is represented by the following equation:

$$I_{amb} = M_{diff} \otimes Factor_{amb} \tag{10}$$

where $Factor_{amb}$ modulates the value of M_{diff} so that all the objects in the scene reflect a small part of the diffuseness contribution.

From Equations (6)- (10), the color images can be easily calculated in the compute shader.

D. IMAGE SYNTHESIS STAGE

The viewpoint mask as input data is a 2D buffer that records the viewpoint arrangement of the SMV 3D display. The relationship between the viewpoint number v_{kl} and the subpixel index (k, l) can be represented by the following equation [39], [40]:

$$v_{kl} = \frac{(k - 3l\tan\alpha) \bmod ((m+1)p_u/(mp_h\cos\alpha))}{((m+1)p_u/(mp_h\cos\alpha))} N_{tot} \tag{11}$$

where the microlens magnification $\boldsymbol{m}$ can be expressed in terms of the optimal viewing distance $\boldsymbol{D}$ and the lens focal length $\boldsymbol{f}$ as $m+1 = fD$. p_u is the lens pitch, p_h is the horizontal subpixel pitch, and N_{tot} is the total number of viewpoints. In graphic memory, there are dense viewpoint color images in the deferred shading stage that can be used to obtain all the subpixel values of the final 3D images. This stage works through parallel computing and only has assignment operations. Therefore, the time consumption is small enough to be ignored.

III. EXPERIMENTS AND RESULTS

The HRT method is implemented in the computer shader and fragment shader. The PC hardware includes an Intel ®I9 9980 XE (4.26 GHz) CPU with 16 GB of RAM and an NVIDIA GeForce 1080 GPU with 8 GB of RAM. The GPU is the main factor that affects the rendering frame rate. Six 3D models, including monkey, car, heart, buildings, Manhattan and furniture, are used to test the performance of the HRT method. Monkey, car, heart, and buildings are simple 3D models, while Manhattan and furniture are complex 3d models. The numbers of vertices, faces and triangles in the models are listed in Table 1. An 8k SMV 3D display with 80 degree viewing angle and 100 viewpoints is applied in the experiment, and other parameters of the display are shown in Table. 2.

In our experiment, PSNR is applied to measure the squared intensity differences of the synthesized and ideal view image pixels. The ideal view image can be obtained with the ECVIR or BRT methods. Then, based on the average PSNR performance, we compare the outcome of the HRT method with

TABLE 1. Specifications of 3D data in the experiment.

3D model data	Vertices	Faces	Triangles
Monkey	704	654	948
Car	809	778	1098
Heart	52,182	67,208	124,415
Buildings	0.8 million	1.7 million	1.5 million
Manhattan	542.6 million	764.3 million	764.3 million
Furniture	323.7 million	591.8 million	591.8 million

TABLE 2. Display configuration in the experiment.

SMV 3D display	Specifications	
LCD	Resolution	7680×4320
	Pixel pitch	136 μm
ECVIR Virtual camera array settings	Array size	100
	Horizontal field of view	80 degrees
	Neighboring distance	27.15 mm
	Distance between LCD panel and virtual camera array	2913 mm

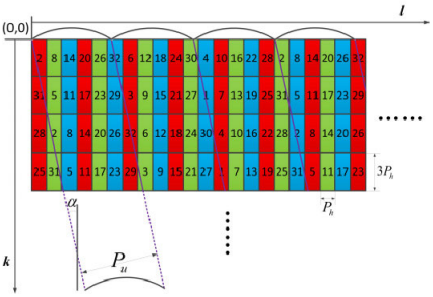


FIGURE 8. Subpixel-viewpoint arrangement of SMV 3D display and the corresponding parameters.

those of state-of-the-art methods, namely, GMMDIBR [20] and MVDRT [22]. We apply different hole-filling methods in GMMDIBR, MVDRT and HRT to refine the blended image. The input data of GMMDIBR and MVDRT are color images and depth images, while the input data of HRT are depth images, normal images, diffuseness images and specular images. Figure 10 shows that the proposed HRT method provides better performance than the existing state-of-the-art MVD methods. The number of reference views is four, the number of frames is 200, and the resolution of the reference views is 1024 × 768. The PSNR improvement range varies from 6.735 dB to 13.798 dB with an average improvement of 9.54 dB for GMMDIBR and 6.34 dB to 11.312 dB with an average improvement of 7.15 dB for MVDRT.

Because new viewpoint images should have different colors for the same virtual point in different reference views, the new viewpoint images generated by GMMDIBR and MVDRT have many error pixels in Figure 11. Color images and depth images cannot provide enough material

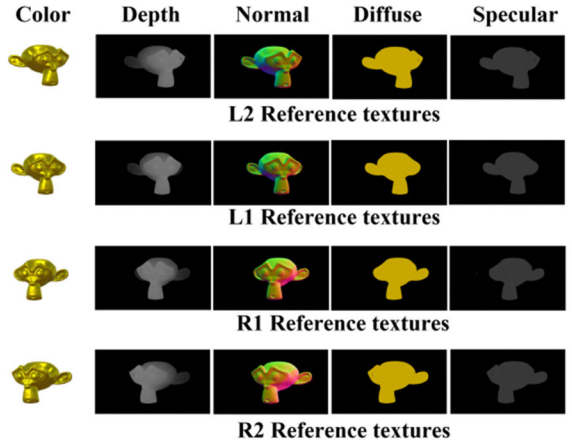


FIGURE 9. The textures of four reference views for the experiment.

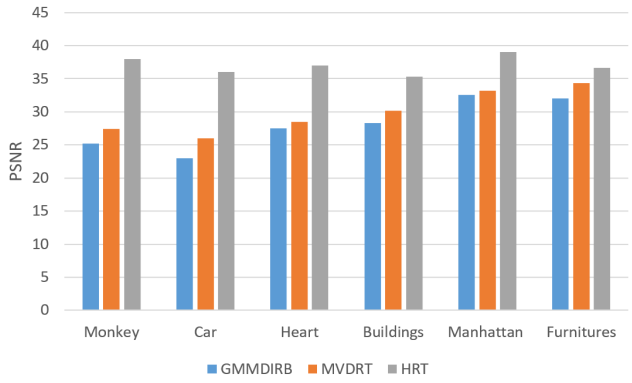


FIGURE 10. The PSNR values of the new viewpoint images generated by different methods.

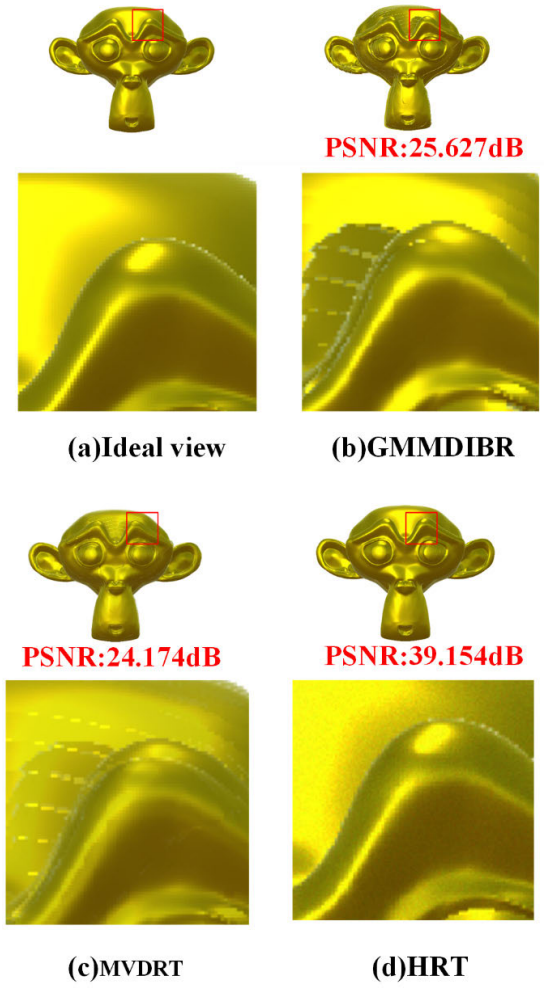


FIGURE 11. The new viewpoint images generated with different methods and their ideal view images. The HRT images have accurate lighting.

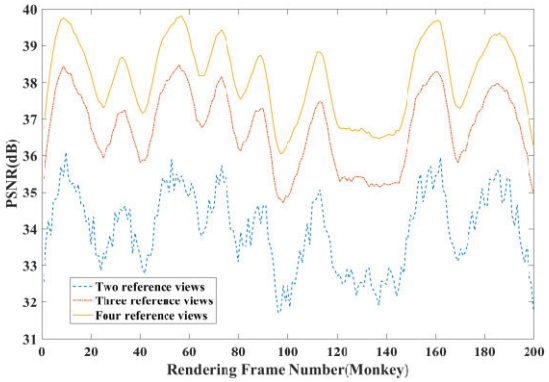


FIGURE 12. PSNR comparison for rendering monkey with 200 frames by using different numbers of reference views in the HRT pipeline.

information to generate new viewpoint images with accurate lighting, so the PSNR values of these methods are lower than that of the HRT method.

Figure 12 illustrates that with increasing reference views, our proposed HRT pipeline can provide better performance. The improvement ranges vary from 2.169 dB to 3.303 dB with an average improvement of 2.432 dB using three reference views and 3.585 dB to 4.630 dB with an average improvement of 3.726 dB using four reference views.

Observers standing at different positions should obtain different colors for one virtual point because of the different unit half vectors in Equation (8). The specular term I_{spec} can directly affect the image quality of the newly generated viewpoint, as shown in Figure 13. The values of M_{spec} and S_{spec} are (1.0, 1.0, 0.0) and (0.7, 0.7, 0.7), respectively. The monkey material is an ideal diffuse material when M_{shi} is zero. The PSNRs of GMMDIBR and MVDRT are higher than the PSNR of the HRT method. The hole-filling method of the former methods is better than that of the HRT method under these conditions. With increasing M_{shi}, the PSNR values of GMMDIBR and MVDRT decrease quickly, while the PSNR value of HRT increases slowly. Therefore, HRT is more suitable for virtual rendering than GMMDIBR and MVDRT for most virtual scenes.

The rendering times of the different methods and every stage of HRT are shown in Table 3 and Table 4. As shown

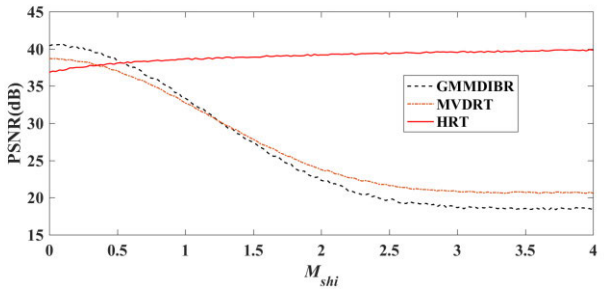


FIGURE 13. PSNR comparison for new viewpoint images with different specular materials by setting different M_{shi} (monkey).

TABLE 3. One frame 3D image rendering times of different methods.

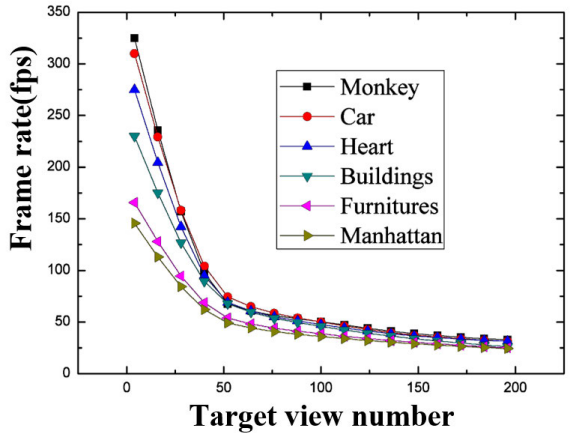
Model	HRT (ms)	GMMDIBR (ms)	MVDRT (ms)	ECVIR (ms)	BRT (ms)
Monkey	12.99	12.768	11.653	512.256	258.176
Car	14.961	13.263	13.143	732.44	321.342
Heart	17.523	15.367	14.748	654.117	876.341
Buildings	18.189	17.159	16.136	1354.284	1341.251
Furniture	23.116	21.269	21.673	7984.788	6854.632
Manhattan	25.768	22.174	21.876	7332.155	7823.322

TABLE 4. The time consumption at every stage in the HRT method.

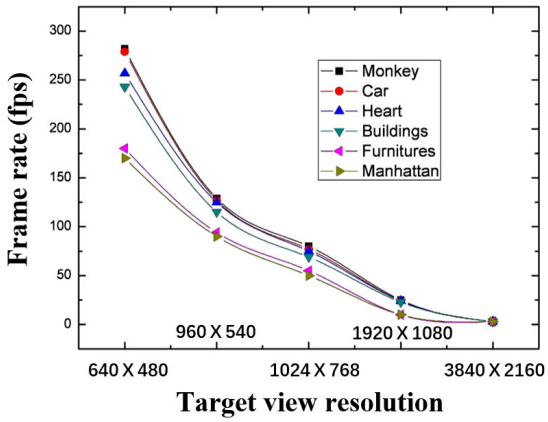
Model	Sparse reference viewpoint images generation (ms)	Dense viewpoints generation (ms)	Deferred shading stage (ms)	Image synthesis stage (ms)
Monkey	5.184	5.978	0.874	0.954
Car	6.856	6.015	0.916	1.174
Heart	9.117	6.154	1.003	1.249
Buildings	10.536	5.723	0.945	0.985
Furniture	14.581	6.073	0.922	1.540
Manhattan	17.661	5.946	1.167	0.994

in Table 3, the rendering time is related to the complexity of virtual scenes. The results also illustrate that our HRT method has an obvious advantage in rendering efficiency compared to those of the ECVIR and BRT methods. Because the HRT should process more input data and has more stages than GMMDIBR and MVDRT, it requires more time to render one frame of a 3D image, while the frame rate is still more than 35 fps on average. As depicted in Table 4, the time consumption of the last three stages has no relationship with the 3D models.

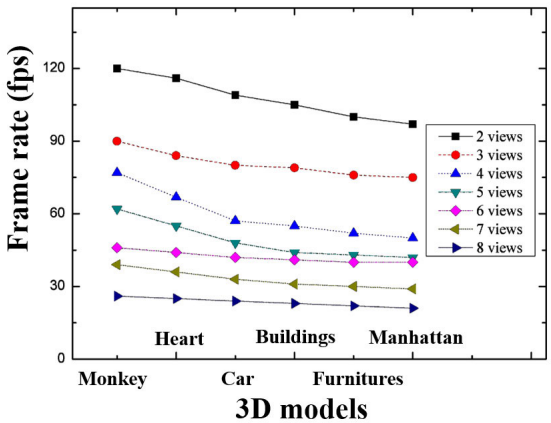
Figure 14 illustrates the main factors affecting rendering efficiency. The rendering frame rates of the six models decrease rapidly with increasing view numbers in Figure 14(a). However, the frame rate also remains above 20 fps when the viewpoint number reaches 200. As shown in Figure 14(b), with increasing target view resolution, the frame rate decreases rapidly because the number of computing units in the GPU is limited. In addition, as shown in Figure 14(c), increasing reference viewpoint number results in a decrease in frame rate because the first stage consumes more time.



(a) The relationships between rendering efficiency and number of target views



(b) The relationships between rendering efficiency and the target view resolution



(c) The relationships between rendering efficiency and different models

FIGURE 14. The factors affecting the rendering efficiency: (a) The number of target views. The resolution of the target view is 1024 × 768. The number of reference views is 3. (b) The resolution of the target views. The numbers of reference views and target views are 3 and 100, respectively. (c) The type of model. The resolution of the target view is 1024 × 768. The number of target views is 100.

The final 3D image that is displayed on the LCD panel of the SMV 3D display can be generated by the image synthesis stage, as shown in Figure 15. The proposed HRT algorithm is implemented on an 8k SMV 3D display. The real-time

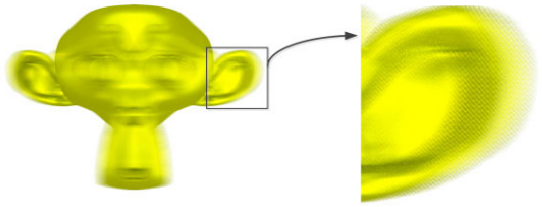


FIGURE 15. The final 3D image generated by the image synthesis stage.

reconstructed 3D images from different perspectives are shown in Figure 16.

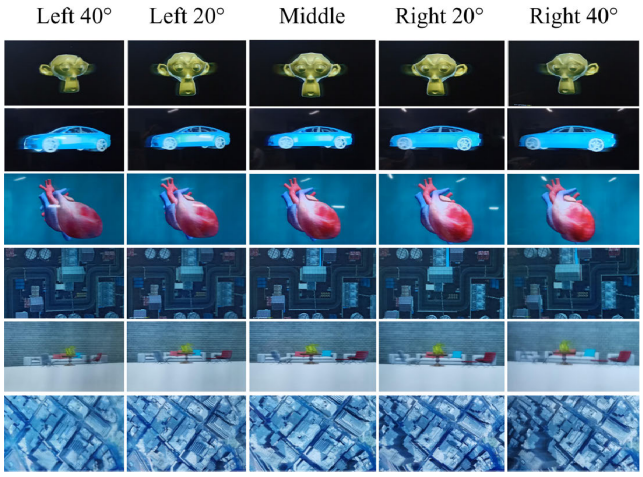


FIGURE 16. Different perspectives of the reconstructed 3D scene on the 8k SMV 3D display.

IV. CONCLUSION

In summary, a whole new SMV rendering pipeline based on the hybrid rendering technique (HRT) is constructed that can generate accurate lighting effects in real-time when the number of viewpoints of the SMV 3D display is greater than 50, the viewing angle is greater than 100 degrees, the resolution of a single viewpoint image is more than 512*512 and the resolution of the LCD panel is 7680*4320; in particular, complex scenes can be generated in real time. Real-time 3D optical reconstruction with accurate lighting effects is realized on an 8k SMV 3D display with an 80-degree viewing angle and 100 viewpoints. The main factors affecting the rendering efficiency are the number of target views, the number of reference views and the resolution of the target views. Experiments demonstrate that when the number of reference views is four and the resolution of the target view is 1024×768, the frame rate is more than 35 fps, the PSNR value of HRT is greater than 36 dB, and the rendering result has a good lighting effect. The HRT method has an obvious advantage in image quality and lighting effects for most virtual scenes in comparison to GMMDIBR and MVDRT.

REFERENCES

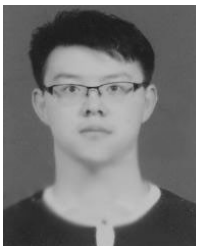
[1] S. Li, H. Li, Z. Zheng, Y. Peng, S. Wang, and X. Liu, "Full-parallax three-dimensional display using new directional diffuser," *Chin. Opt. Lett.*, vol. 9, no. 8, 2011, Art. no. 081202.
[2] Y. Takaki, "High-density directional display for generating natural three-dimensional images," *Proc. IEEE*, vol. 94, no. 3, pp. 654–663, Mar. 2006.
[3] F. C. Fan, "Three-dimensional display based on the holographic functional screen," *Opt. Eng.*, vol. 50, no. 9, Sep. 2011, Art. no. 091303.
[4] X. Sang, F. C. Fan, C. C. Jiang, S. Choi, W. Dou, C. Yu, and D. Xu, "Demonstration of a large-size real-time full-color three-dimensional display," *Opt. Lett.*, vol. 34, no. 24, pp. 3803–3805, Dec. 2009.
[5] F. C. F. Frank C. Fan, S. C. Sam Choi, A. C. C. J., and C. C. Jiang, "Demonstration of full-parallax three-dimensional holographic display on commercial 4 k flat-panel displayer," *Chin. Opt. Lett.*, vol. 14, no. 1, 2016, Art. no. 010007.
[6] D. Lin, Q. H. Nguyen, M. N. Do, and S. J. Patel, "Systems and methods for generating a virtual camera viewpoint for an image," Tech. Rep., 2015.
[7] M. Halle, "Multiple viewpoint rendering," in *Proc. 25th Annu. Conf. Comput. Graph. Interact. Techn.*, 1998, pp. 243–254.
[8] B. Pang, X. Sang, S. Xing, X. Yu, D. Chen, B. Yan, K. Wang, C. Yu, B. Liu, C. Cui, Y. Guan, W. Xiang, and L. Ge, "High-efficient rendering of the multi-view image for the three-dimensional display based on the backward ray-tracing technique," *Opt. Commun.*, vol. 405, pp. 306–311, Dec. 2017.
[9] *Turing Architecture White paper*, NVIDIA company, Santa Clara, CA, USA, pp. 68–69.
[10] B. Battin, G. Valette, J. Lehuraux, Y. Remion, and L. Lucas, "A premixed autostereoscopic OptiX-based volume rendering," in *Proc. Int. Conf. 3D Imag. (IC3D)*, Dec. 2015, pp. 1–5.
[11] W. Sun, L. Xu, O. C. Au, S. H. Chui, and C. W. Kwok, "An overview of free viewpoint depth-image-based rendering," in *Proc. APSIPA*, Dec. 2010, pp. 14–17.
[12] L. Do, G. Bravo, S. Zinger, and P. De With, "GPU-accelerated real-time free-viewpoint DIBR for 3DTV," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 633–640, May 2012.
[13] G. Bravo, L. Do, S. Zinger, and P. H. N. de With, "Real-time free-viewpoint DIBR on GPUs for 3DTV systems," in *Proc. IEEE Int. Conf. Consum. Electron. -Berlin (ICCE-Berlin)*, Sep. 2011, pp. 1–4.
[14] L. Wang, C. Hou, J. Lei, and W. Yan, "View generation with DIBR for 3D display system," *Multimedia Tools Appl.*, vol. 74, no. 21, pp. 9529–9545, Nov. 2015.
[15] Y. Yamauchi, S. Yanagawa, H. Kobayashi, K. Taira, and Y. Hirayama, "Real-time rendering for autostereoscopic 3D display systems," *ACM SIGGRAPH* vol. 5291, no. 53, p. 153, 2007.
[16] J. Lei, X. He, H. Yuan, F. Wu, N. Ling, and C. Hou, "Region adaptive R-λ model-based rate control for depth maps coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1390–1405, Jun. 2018.
[17] H. Yuan, S. Kwong, X. Wang, Y. Zhang, and F. Li, "A virtual view PSNR estimation method for 3-D videos," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 134–140, Mar. 2016.
[18] H. Yuan, J. Liu, H. Xu, Z. Li, and W. Liu, "Coding distortion elimination of virtual view synthesis for 3D video system: Theoretical analyses and implementation," *IEEE Trans. Broadcast.*, vol. 58, no. 4, pp. 558–568, Dec. 2012.
[19] H. Yuan, Y. Chang, J. Huo, F. Yang, and Z. Lu, "Model-based joint bit allocation between texture videos and depth maps for 3-D video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 4, pp. 485–497, Apr. 2011.
[20] M. D. Rahaman and M. Paul, "Virtual view synthesis for free viewpoint video and multiview video compression using Gaussian mixture modelling," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1190–1201, Oct. 2018.
[21] S. Li, C. Zhu, and M. T. Sun, "Hole filling with multiple reference views in DIBR view synthesis," *IEEE Trans. Multimedia*, vol. 20, no. 8, pp. 1948–1959, Mar. 2018.
[22] F. Shao, W. Lin, R. Fu, M. Yu, and G. Jiang, "Optimizing multiview video plus depth retargeting technique for stereoscopic 3D displays," *Opt. Express*, vol. 25, no. 11, p. 12478, 2017.
[23] L. Yao, Y. Han, and X. Li, "Fast and high-quality virtual view synthesis from multi-view plus depth videos," *Multimedia Tools Appl.*, vol. 78, no. 14, pp. 19325–19340, 2019.
[24] A. Morar, F. Moldoveanu, A. Moldoveanu, and V. A. O. Balan, "GPU accelerated 2D and 3D image processing," in *Proc. Federated Conf. Comput. Sci. Inf. Syst.*, 2017, pp. 653–656.

[25] L. Yao, Y. Han, and X. Li "Virtual viewpoint synthesis using CUDA acceleration," in *Proc. ACM Symp. Virtual Reality Softw. Technol.*, Nov. 2016, pp. 365–366.
[26] *Principles of Computer Graphics||Rendering: Shading and Lighting*, 2004, pp. 131–163, doi: 10.1007/b135398.
[27] G. Bishop and D. M. Weimer, "Fast phong shading," *ACM SIGGRAPH Comput. Graph.*, vol. 20, no. 4, pp. 103–106, Aug. 1986.
[28] A. Hast, T. Barrera, and E. Bengtsson, "Fast specular highlights by modifying the Phong-Blinn model," in *Proc. 30th Annu. Conf. Comput. Graph. Interact. Techn.*, 2003, pp. 1–3.
[29] Dimov, Ivan, T. Gurov, and Anton A. Penzov, "A Monte Carlo approach for the cook-torrance model," in *Proc. Int. Conf. Numer. Anal. Appl.*, 2004, pp. 257–265.
[30] C. Schied and C. Dachsbacher, "Deferred attribute interpolation for memory-efficient deferred shading," in *Proc. 7th Conf. High-Perform. Graph.*, 2015, pp. 43–49.
[31] M. D. Seemann and C. D. Claussen, "Hybrid 3D visualization of the chest and virtual endoscopy of the tracheobronchial system: Possibilities and limitations of clinical application," *Lung Cancer*, vol. 32, no. 3, pp. 237–246, Jun. 2001.
[32] H.-J. Chen, F.-H. Lo, F.-C. Jan, and S.-D. Wu, "Real-time multi-view rendering architecture for autostereoscopic displays," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2010, pp. 1165–1168.
[33] H. W. Jensen, "An efficient hybrid shadow rendering algorithm," in *Proc. 15th Eurograph. Conf. Rendering Techn. Eurograph. Assoc.*, 2004, pp. 1–3.
[34] Z. Jia, M. Maggioni, J. Smith, and D. P. Scarpazza, "Dissecting the Nvidia Turing t4 GPU via microbenchmarking," Tech. Rep.
[35] C. Barré-Brisebois, *Hybrid Rendering for Real-Time Ray Tracing*. Berkeley, CA, USA: Ray Tracing Gems Apress, 2019, pp. 437–473.
[36] C. Wynn, *Opengl Render-To-Texture*. Santa Clara, CA, USA: NVIDIA Corp, 2002)
[37] S. Xing, S. Liu, and X. Sang, "Multi-projector three-dimensional display for 3D geographic information system," *Optik*, vol. 139, pp. 385–396, Jun. 2017.
[38] F. Policarpo. (2005). *Deferred Shading Tutorial CheckMate Games*. [Online]. Available: https://nume.googlecode.com/svn-history/r12/trunk/monnezza/redsh/ogre1/Deferred_Shading_Tutorial_SBGAMES2005.pdf
[39] L. Do, S. Zinger, and P. H. N. de With, "Conversion of free-viewpoint 3D multi-view video for stereoscopic displays," in *Proc. IEEE Int. Conf. Multimedia Expo.*, Jul. 2010, pp. 1730–1734.
[40] S.-M. Jung, H. Kang, B.-Y. Lee, and I.-B. Kang, "Numerical simulation of the displayed image on the entire screen of autostereoscopic displays," *Opt. Express*, vol. 23, no. 6, p. 7842, Mar. 2015.

SHUJUN XING received the Ph.D. degree in electronic science and technology from the Institute of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing, China, in 2017. He is currently a Postdoctoral Researcher with the Department of Precision Instruments, Tsinghua University, Beijing. His major is in optical engineering and mainly studying the 3D rendering method for special 3D display. His research interests include a three-dimensional display, virtual reality, and multiview rendering.

XINZHU SANG received the dual bachelor's degrees in instrument science and management engineering from Tianjin University, Tianjin, China, in 1999, the M.S. degree from the Beijing Institute of Machinery, Beijing, China, in 2002, and the Ph.D. degree in physical electronics from the Beijing University of Posts and Telecommunications, Beijing, in 2005. From December 2003 to March 2005, he was a Research Assistant with the Optoelectronics Research Centre, Department of Electronics Engineering, City University of Hong Kong. From July 2007 to July 2008, he was a Postdoctoral Research Scholar with the University of California at Irvine, Irvine, CA, USA. He is currently a Full Professor with the Beijing University of Posts and Telecommunications. His research interests include a three-dimensional display, holography, and novel photonic devices. He is the Secretary-General of the Committee of Holography and Optical Information Processing, Chinese Optical Society, a Senior Member of the Chinese Institute of Communication, and a Senior Member of the Chinese Institute of Electronics.

LIANGCAI CAO received the B.S. degree from the Department of Physics, Harbin Institute of Technology, Harbin, China, in 1999, and the Ph.D. degree in volume holography from the Department of Precision Instruments, Tsinghua University, Beijing, China, in 2005. He is an Associate Professor with the Department of Precision Instruments, Tsinghua University. He has published more than 100 articles in reviewed journals and conferences. His research interests include imaging, storage, and display based on volume holography. He is a Senior Member of the OSA.



YANXIN GUAN received the B.S. degree in optical information science and technology from Qingdao University, in 2015. He is currently pursuing the Ph.D. degree with the Institute of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing, China. His research interests include 3D display technologies and computer graphics.

YUANHANG LI received the B.S. degree in communication engineering from Henan University, in 2016. He is currently pursuing the Ph.D. degree with the Institute of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing, China. His research interests include 3D display technologies, computer graphics, and integral photography and its applications in the 3D display.