DOI: 10.2352/ISSN.2470-1173.2016.21.3DIPM-035

3D Autostereoscopic Display Image Generation Framework using Direct Light Field Rendering

Young Ju Jeong, Yang Ho Cho, Hyoseok Hwang, Hyun Sung Chang, Dongkyung Nam, and C. -C Jay Kuo; Samsung Advanced Institute of Technology; Suwon, South Korea

Abstract

The rapid development of 3D display technologies allows consumers to enjoy the 3D world visually through different display systems such as the stereoscopic, multiview, and light field displays. Despite the hype of 3D display development, 3D contents for various display systems are very insufficient and the most of 3D contents are only stereo sequences for stereoscopic 3D. To handle with the lack of 3D contents for various display systems, various 3D display images are generated by stereo sequences. Howver, the conventional 3D display rendering algorithm suffers from increased computational complexity and memory usage with the increasing number of view in 3D displays to achieve a sufficient level of reality. This paper proposes an efficient method to generate 3D display images by 3D direct light field rendering. We propose new 3D image generation algorithm for hole filling, boundary matting, and postprocessing from common stereo input images.

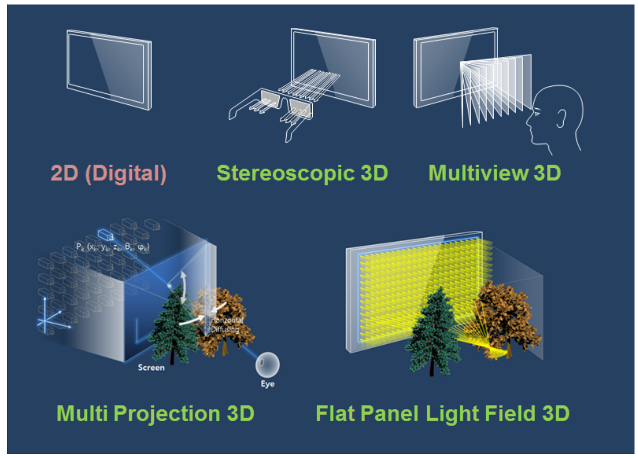

Figure 1. *Various 3D displays such as stereoscopic, multiview, multi-projection, and flat panel light field 3D display*

Introduction

3D display technologies have been rapidly developed. With 3D display, viewers experience the real three dimensional world that surpasses the experience provided by existing 2D display, which merely presents two dimensional images or movies. 3D display is categorized as either stereoscopic, for which glasses are required, or autostereoscopic, for which glasses are not required. Autostereoscopic 3D display is categorized as multiview display and light field display. The light field display can be implemented by either projection type 3D display or flat panel 3D display.

In human vision, the scenes shown by either eye are slightly different from each other because of the physical separation of the eyes. Because of these differences, the human brain interprets the information in 3D. In order to recreate the 3D world through a physical 2D display, this principle is applied in reverse. For a stereoscopic display, left and right images shown by left and right eyes are projected through a display. These projected left and right images for our either eyes enforce to interpret it 3D objects not just two view images. The right image from a stereoscopic display is filtered by glasses to left eye and the light image from a stereoscopic display is filtered by glasses to right eye to separate left and right images. However, the requirement of glasses induces inconvenience, and it causes viewing fatigue because motion parallax is not available unless a special motion tracking sensor is used [1].

Beyond stereoscopic display, which provides only two viewpoints 3D binocular vision, multiview stereoscopic display provides multiple viewpoint 3D. The multiview 3D display provides unique images to several viewpoints, offering limited motion parallax, compared with only one viewpoint in the case of stereoscopic 3D display. Multiview 3D display is implemented by lenticular sheet lens overlaid on a liquid crystal display (LCD) [2] or by parallax barrier display. Under a N multiview display system, viewers sense the first 3D view when the left eye sees the first virtual view and the right eye sees the second virtual view. When the viewpoint is moved, the left eye sees the second virtual view and the right eye sees the third virtual view, and the viewer perceives the second 3D view. In this way, $(N-1)$ virtual 3D scenes are changed according to $(N-1)$ viewpoints. The changing scenes are followed by viewpoint changes, giving us the illusion that we are looking around the 3D world as a real-life world.

Even though multiview display provides motion parallax with multiple viewpoints, these separate limited viewpoints cannot offers natural motion parallax as real world but discrete motion parallax. Moreover limited motion parallax can only be experienced when viewed from certain area called sweet spot.

Beside multiview display which provides limited stepwise viewpoints, light field display generates fluent viewpoints as real world. If the number of light rays increase so 3D display provide not limited viewpoints but countless viewpoints like real world, then we can enjoy real like 3D world. To develop 3D display which projects continuous viewpoints, multi-projection 3D display and flat panel display are used. Multi-projection 3D displays create light rays using a large number of projectors. Flat panel light field display is designed to provide possibly many numbers of viewpoints to reduce stepwise parallax artifacts.

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.21.3DIPM-035

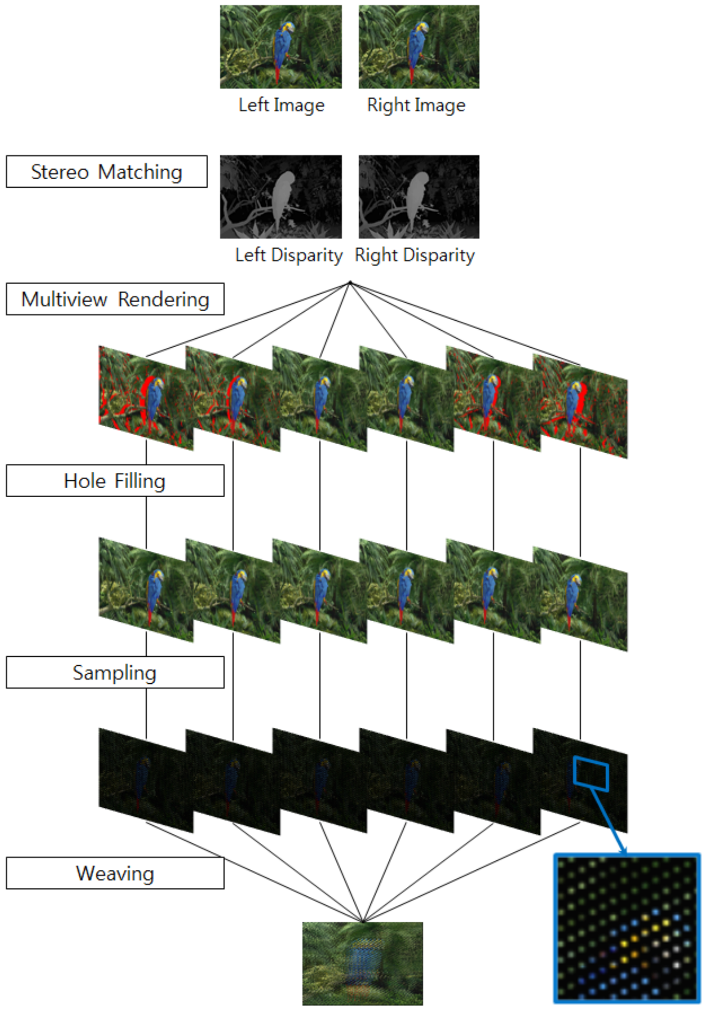

Figure 2. *MVR algorithm process: from stereo contents disparity maps are estimated by stereo matching algorithm. Multiview images are generated by pixel shift algorithm by depth image based rendering. Since exterior view has disocclusion areas, hole filling process is applied to the multiivew image. Finally, light rays from multiview images are sampled ans weaved according to the addressing point of 3D display pixels.*

3D Rendering Algorithms and Its Challenges

To generate multiview images for the multiview displays, multi-camema capturing system is required. However to generate multiview images by multi-capturing system has difficulties. To handle with these difficulties, Verkel introduced a Multiview Rendering (MVR) algorithm, which renders each group of rays from the same viewpoint using a perspective projection model [3, 4].

MVR algorithm uses the principle that light rays from multiview display converge to certain viewpoints. If we rearrange light rays from multiview display or light field display, we can get light rays for multiview images.

Multiview image can be generated by mapping the original pixel of the input image to a new virtual view image at the new location, which is shifted by the disparity pixels from the original image and this processing is Depth Image Based Rendering (DIBR).

In the processing of DIBR hole areas are created. Disoccluded areas are parts of holes that are occluded by object in the input image but are disoccluded by object in the virtual view images. Due to the changes in viewpoints, occlusion regions of the original image become disoccluded, resulting in problems related to the restoration of output image information that is not contained in the input image. Hole areas in the multiview images are restored by hole filling algorithm.

Once multiview images are generated, subsequently the MVR algorithm weaves multiview images in the 3D panel image by using information from the display light ray pattern [3]. The direction for pixels of 3D multiview display are designed, we can know the pixel of 3D display is addressed to which view image among multiview images. With this address, multiview images are sampled. Selected pixels are light rays for 3D display and others are not used. Accumulating such sampled pixels of multiview images to one panel images, then 3D panel image for the 3D multiview display is generated.

Figure 2 shows the process for DIBR algorithm. From the conventional stereo contents, depth information is estimated by stereo matching algorithm. Once depth information is known, then we can generate images for the other viewpoint. It is a DIBR

DOI: 10.2352/ISSN.2470-1173.2016.21.3DIPM-035

algorithm. If the generated multiview images express wider viewpoint compared with input stereo images, multiview images have disoccluded area. In the multiview synthesis process, red area of multiview images are these disoccluded hole areas. These hole areas are restored by hole filling algorithms in the image. Once hole filled multiview images are synthesized, then we have all light rays for 3D panel image.

The MVR algorithm is intuitive and easy to implement, but it has crucial difficulties in terms of computational complexity and memory usage. This may hinder the future commercial viability of 3D display products. The computational complexity and memory usage increase with increase in the number of viewpoints and the resolution of the 3D display because MVR needs to generate more multiview images with a higher resolution.

On the other hand, Direct Light Field Rendering (DLFR) algorithm reduces complexity and memory usage to build 3D panel image. It composes the display 3D panel image without reconstructing multiview images. Since it generate light rays for 3D display directly, general 2D image based processing cannot be applied. In this paper we introduce total framework which contains 2D image based processing.

Direct Light Field Rendering Framework

We introduce 3D display rendering framework in the light field domain. To generate high quality 3D images, display calibration process is applied [6]. Using calibrated display parameters, more accurate 3D display modeling function can be acquired. For the multiview rendering case, the calibration parameter offers accurate weaving pattern. Since the most common 3D content is stereo images, stereo matching algorithm restore 3D structure form 3D contents.

In the case that he amount of depth for 3D display is less than the amount of depth for stereo content, there is barely disocclusion area. Interpolated views of stereo content barely have disocclusion area, but extrapolated views of stereo contents have disocclusion areas. In the case of less volume than stereo contents, direct light field rendering algorithm sufficiently generates 3D panel image [5]. However, if the display volume is larger than stereo contents, disocclusion area is occurred as extrapolated views. To handle hole area in the light field image, we suggest hole filling algorithm in the light field domain. Apart from hole filling, there are 2D post processings such as boundary matting and view smoothing. We also introduce to handle these 2D image based post processing in the light field domain.

Figure 3 illustate the direct light field rendering framework. From stereo contents disparity maps are estimated by stereo matching algorithm. Hole region and information is estimated by hole filling algorithm. Direct light field rendering algorithm generate 3D panel image for input stereo contents and hole area separately. Finally 3D panel image is obtained by combining input stereo and hole area.

3D Display Calibration

The autostereoscopic display cannot be manufactured ideally, calibration process play important role to generate accurate high quality 3D. Display calibration estimate the manufactured 3D parameter for the 3D display such as gap between panel and 3D optical components, slant angle of pallax barrier or lenticular lens, and starting points of 3D optical components. The calibrated parameter for 3D display provides accurate weaving pattern for the DIBR algorithm and accurated 3D display function for the direct light field rendering algorithms. Display calibration is conducted by automatically by camera captured image. By calibration pattern image captured by camera, 3D display parameters are estimated as shows in Fig. 4.

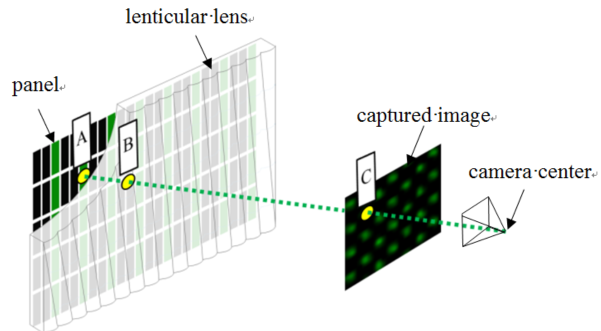

Figure 4. *3D display calibration: 3D display parameters are estimated by 3D calibration pattern image caputred by a camera [6].*

Stereo Matching

Stereo matching, which restores the 3D structure from stereo images is an essential technology for recreating the 3D contents for different 3D displays. If the stereo contents are well calibrated and rectified having horizontal disparities, than general stereo matching algorithm can be used. Global dynamic programming optimization algorithm is applied on the tree structured stereo images [7]. It simply assign high confidence region disparity using stereo matching dynamic programming algorithm, and then disparity fitting algorithm is used for the low confidence region.

Hole Area Estimation and Hole Filling

Hole area is the disoccluded area which does not contained in the input stereo images. The disoccluded area can be estimated by disparity value obtained by stereo matching algorithm and its absolute depth amount which is represent in a 3D display. The color and depth information of hole area is restored by depth guided inpainting algorithm.

Hole area is calculate by disparity difference of each image pixel included in the 2D color image, for example, a left difference and a right difference, based on the 2D color image and the depth image, and estimate one of or both a hole region in a horizontal. The hole region in the horizontal direction refers to a hole region occurring when the foreground moves in the horizontal direction.

Hole color and hole disparity images are generated based on a result of the restoring of the hole region. Hole color image is restored by filling the hole region with a color value having a similar attribute in the 2D color image or a color value of the background. Hole depth image is similarly generated by filling the hole region with a depth value having a similar attribute in the depth image.

DOI: 10.2352/ISSN.2470-1173.2016.21.3DIPM-035

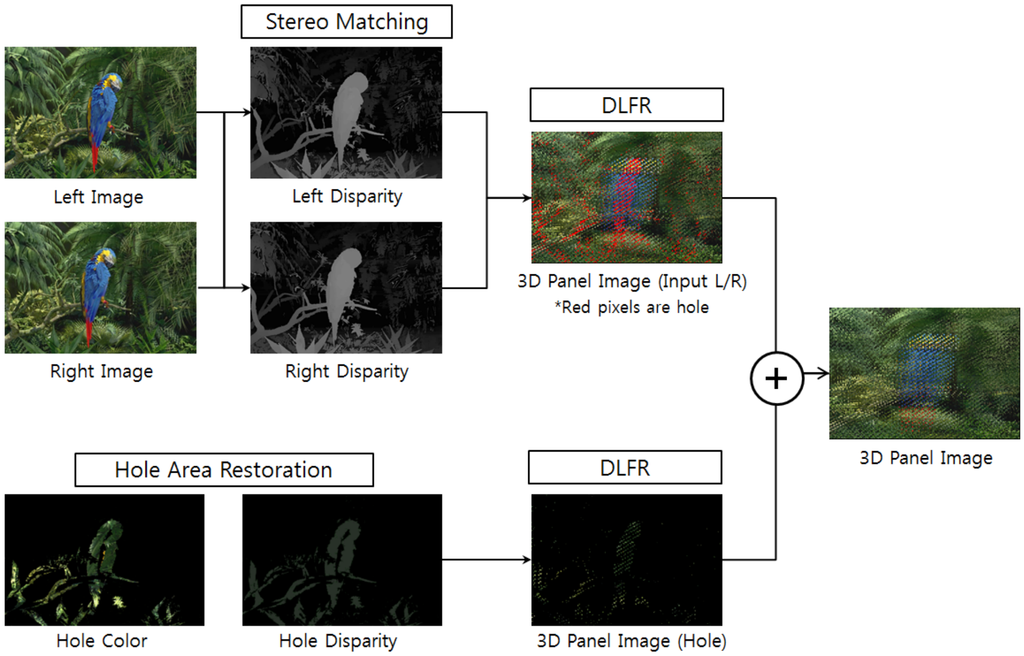

Figure 3. *Direct light field rendering framework: from stereo contents disparity maps are estimated by stereo matching algorithm. Hole region and information is estimated by hole filling algorithm. Direct light field rendering algorithm generate 3D panel image for input stereo contents and hole area separately. Finally 3D panel image is obtained by combining input stereo and hole area.*

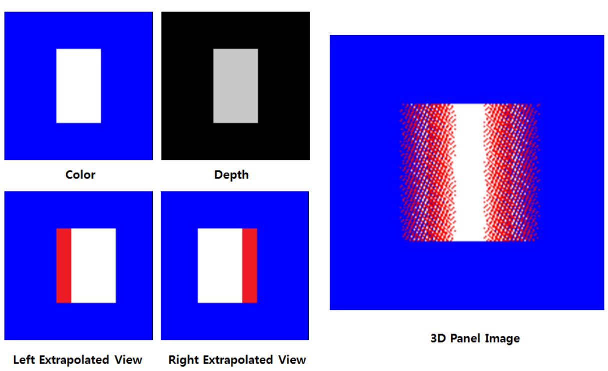

Figure 5. *Input color depth, extrapolated left and right images, and 3D panel image: Hole areas of extrapolated left and right image has 2D structure. However hole area of 3D panel image rendered by direct light field rendering algorithm does not have 2D structure.*

In the multiview rendering algorithm case, the hole area can be simply restored because it contain image structure. However for the direct light field rendering case, it is very difficult to restore in the 3D panel image. Figure 5 shows panel image of direct light field rendering, and the red area means not assigned pixel value due to the disocclusion. The not assigned pixel area is separated by pixel by pixel position, and the neighborhood area does not have any image structure information.

To handle hole area with the direct light field rendering algorithm, we propose layered depth and color representation. The color and depth information which is restored by earlier process is rendered to the 3D panel image first by the same direct light field rendering algorithm. Since disoccluded area is background, hole area can be overlapped when the foreground pixels area rendered. Then original input stereo images are rendered by direct light field rendering algorithm.

Figure 3 illustrate direct light field rendering process when there is occlusion areas. 3D panel image by input stereo images has occlusion areas as shows in upper 3D panel images. The hole area does not have 2D image structure to restore its information. Therefore it is impossible to fill the information in the 3D panel image. The lower process shows the hole area direct light field rendering. Its light rays does not have 2D image structure, too. The intergrated image of input stereo images and hole area images does not have any disoccluded region.

Light Field Rendering

Light rays from the 3D display and stereo contents with depth image can be represent by light field domain [8]. Light rays from 3D display can be modeled by display light ray function below,

$$s^A(x,y,p) = a(x-bp)+d(y) \tag{1}$$

where a denotes the slope of multiline, p the pth multiline, b the x interval between neighboring lines, and $d(y)$ the lens offset due to the slanted lens.

$$a = \frac{s^A[c+n_l,m]-s^A[c,m]}{n_l} \tag{2}$$

DOI: 10.2352/ISSN.2470-1173.2016.21.3DIPM-035

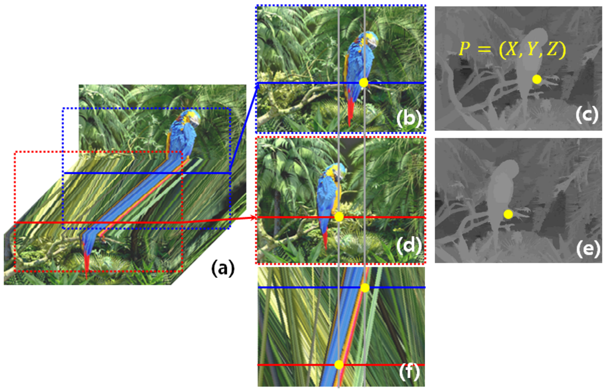

Figure 6. *Light field of 3D Point **P**: (a) Light field volume when t is fixed, (b) left image, (c) disparity in left image, (d) right image, (e) disparity in right image, and (f) epipolar plane image when y is given [5].*

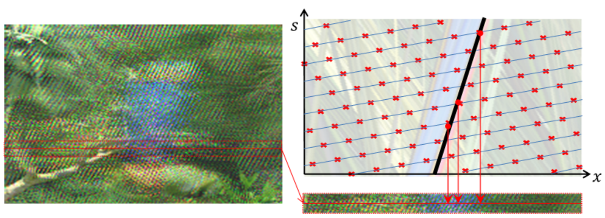

Figure 7. *Direct Light Field Rendering(DLFR): The 3D panel image is composed by the assignment of 3D point **P** to the intersecting point between the DLF and PLF [5]*

$$b = \frac{an_i - (s^A[c+n_i,m] - s^A[c,m])}{a} \quad (3)$$

$$d(y) = l \tan\theta y \quad (4)$$

where c denotes a constant, l the horizontal lens pitch, and θ the slanted angle of the lenticular lens.

Stereo input images can be interpreted as a light field. 3D point is represented as a line as shows in Fig. [?] (f). The point light field function is

$$s^{\mathbf{P}}(x) = \frac{s_R - s_L}{\Delta x}(x - x_L) + s_L \quad (5)$$

where x_L denotes x in the left image coordinate of the 3D point **P**.

The intersecting point of display light field function and point light field function represents the correct destinations among all possible destinations of the 3D point. Figure 7 illustrates how the 3D point is rendered as a 3D panel image upon solving the linear systems of the two variables x and s.

The x coordinate of the 3D panel image of the 3D point **P** is calculated as

$$x(p) = \frac{ab\Delta xp + (s_L - d(x))\Delta x + (s_R - s_L)x_L}{a\Delta x - (s_R - s_L)}, \quad (6)$$

which is the solution to Eq. (1) and Eq. (5). The y coordinate of the 3D panel image is the same as that of the 3D point **P**, because the slanted multiview display provides only horizontal parallax. Since Eq. (6) has variable p, x can take on multiple values; for example, there are three intersecting points in Fig. 3.

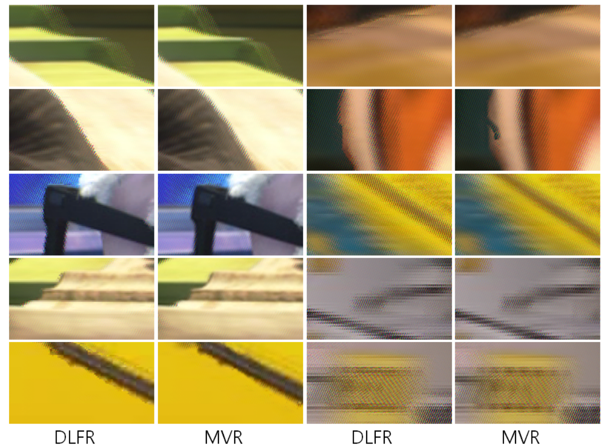

Figure 8. *Partition images of 3D panel image: 3D panel image by DLFR framework and MVR framework.*

Post Processing

Boundary matting algorithm in the light field domain is applied by detecting a boundary region between a foreground and a background in the 2D color image. Conventional boundary matting algorithm blends a color value of the foreground and a color value of the background, which are adjacent to the boundary region, based on a weighted value, and allocate the blended color value to a location in the 3D image corresponding to the boundary region.

Boundary matting in the light field domain allocates the blended color value to a location of a display pixel corresponding to the boundary region in the processing of direct light field rendering. First boundary region in the stereo contents are detected, then render and blend boundary foreground color in the 3D panel image.

For another example, when a location in the 3D image corresponding to a view direction of an image pixel of the 2D color image is included in a smoothing region. It can be implemented by allocating a smoothed color value, in lieu of an original color value, to the location in the 3D image. Through such an image post-processing process, a quality of the 3D image may be improved.

Experimental Results

Figure 8 illustrates partition images of 3D panel image. Figure 8 shows the comparison result of our proposed DLFR framework and MVR algorithm, resulting in almost same quality in the 3D panel image.

We also conduct the psychophysical the darkroom. 10 observers are attended. During the experiments, two generated images for each test image are shown in sequence and the observers are asked to rank the imagers according to their preference. There is no time limit. The observers are allowed to compare the imagers repeatedly. The experimental result shows that the proposed framework and MVR framework perform similar. That means the proposed framework can achieve 3D quality of MVR with much light computation time and memory usage.

DOI: 10.2352/ISSN.2470-1173.2016.21.3DIPM-035

Conclusion

We developed a direct light field rendering framework for 3D displays in the light field domain that provides an efficient rendering approach. The proposed framework deal with boundary artifacts, post processing, and hole area which is occurred when 3D displays present more volume than input stereo contents by retorting in the 3D panel image through direct light field rendering algorithm. Experimental results show that the panel image by direct light field rendering algorithm is similar as multiview rendering but performs better than multiview image in terms of computational complexity and memory usage.

References

[1] L. McMillan, G. Bishop, "Head-tracked stereoscopic display using image warping", Proceedings of SPIE, 21–30 (1995).

[2] C. Berkel, "Image preparation for 3D LCD", Proceedings of SPIE, 84–91 (1999).

[3] C. Verkel, "Image preparation for 3D LCD", Proceedings of SPIE - IS&T Electronic Imaging, SPIE Vol. 3639, 84–91 (1999).

[4] M. Halle, "Multiple viewpoint rendering", SIGGRAPH'98, Proceedings of the 25th annual conference on Computer graphics and interactive techniques, 243–254 (1998).

[5] Y. J. Jeong, H. S. Chang, D. Nam, C. -C. Kuo, "Efficient direct light-field rendering for autostereoscopic 3D displays", in SID symp. Dig. Tech. Papers, **46**(1), 155–159 (2015).

[6] H. Hwang, J. Park, H. S. Chang, Y. J. Jeong, D. Nam, I. S. Kweon, "Lenticular lens parameter estimation using single image for crosstalk reduction of three-dimensional multi-view display", in SID symp. Dig. Tech. Papers, **46**(1), 1417–1420 (2015).

[7] Y. J. Jeong, J. Kim, H. H. Lee, D. Park, "Confidence stereo matching using complementary tree structures and global depth-color fitting", In Consumer Electronics (ICCE), 468–469 (2013).

[8] M. Levoy and P. Hanrahan, "Light field rendering", SIGGRAPH'96, Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 31–41 (1996).

Author Biography

Young Ju Jeong received B.S. degree in Electrical Engineering at Korea University, South Korea, in 2005 and M.S. degree in Electrical Engineering at University of Southern California, U.S., in 2007. She is currently pursuing the Ph.D. degree in electrical engineering at University of Southern California, U.S. She is currently a research member at the Samsung Advanced Institute of Technology. She has been worked on computer vision and image processing for autostereoscopic displays. Her research interests include 3D vision and 3D lightfield rendering.