Design and Implementation of High-Resolution Integral Imaging Display System using Expanded Depth Image

Min-Ho Song, Byung-Muk Lim, Ga-A Ryu
Department of Computer Science, Chungbuk National University, Cheongju, South Korea

Jong-Sung Ha
Department of Computer Engineering, Woosuk University, Chonbuk, South Korea

Kwan-Hee Yoo
Department of Computer Science, Chungbuk National University, Cheongju, South Korea

ABSTRACT

For 3D display applications, auto-stereoscopic display methods that can provide 3D images without glasses have been actively developed. This paper is concerned with developing a display system for elemental images of real space using integral imaging. Unlike the conventional method, which reduces a color image to the level as much as a generated depth image does, we have minimized original color image data loss by generating an enlarged depth image with interpolation methods. Our method was efficiently implemented by applying a GPU parallel processing technique with OpenCL to rapidly generate a large amount of elemental image data. We also obtained experimental results for displaying higher quality integral imaging rather than one generated by previous methods.

Key words: Integral Imaging, 3D Image Acquisition, Elemental Images, Depth Image, GPU Parallel Programming.

1. INTRODUCTION

Autostereoscopic display systems have focused on various applications due to the convenience. Many research has been conducted, such as [1]-[3]. Generally, in integral imaging system, a 3D scene from many different perspectives is captured as individual images obtained by the array of microlenses, called elemental images (EIs). In most of the previous research related to integral imaging, integral images have been generated by using virtual lens arrays [4]-[6]. This method lacks a sense of reality and is limited for developing interaction-based contents since it produces 3D images from virtual objects in a virtual space instead of real space. To resolve these problems, integral imaging can be produced from images taken with depth cameras that can store real space. This approach can be utilized in more applications [7], [8].

An advantage of integral imaging is its simplicity in obtaining 3D information without costly devices and demanding conditions. Our research has made the best use of Kinect V1 and Kinect V2 [15], which are the main devices for extracting information from real spaces at present, generating 640x480 color images / 320x240 depth images, and 1920x1080 color images / 512x424 depth images, respectively. As the first stage work, Li et al. [7] created elemental images from depth images and color images extracted by a Kinect V1 device. And Jeong et al. [8] developed a real-time system to create elemental images based on GPU parallel processing using the same depth camera. Particularly, Jeong et al. [8] made use of 320x240 reduced resolution instead of the original 640x480 resolution for color images in order to fit into the resolution of depth images, which reduces the quality of the final elemental images.

There are a number of attempts to increase the image quality of elemental images for a real scene [9]-[11]. Specially, Zhang, et.al [9] tried to generate elemental images with a lot of voxels based on KinectFusion, and Kwon, et.al [10] tried to enlarge elemental images themselves by interpolation method from a generated elemental image. Because of the low quality
of a given real scene, essentially, there is a limit to generate elemental images with good quality. To resolve the underlying problem, in this paper, we took a reverse approach to expand the resolution of depth images according to that of color images.

In section 2, we describe a method to generate elemental images for high-quality integral imaging. Section 3 presents experimental results that demonstrate that our method can improve the image quality in comparison to previous methods. Finally, we summarize our approach and suggest directions for further research in section 4.

2. INTEGRAL IMAGING DISPLAY SYSTEM USING EXPANDED DEPTH IMAGES

As illustrated in Fig. 1, high-resolution integral imaging display system using expanded depth images is composed of three stages. In the first stage, color images and depth images are captured from a real scene, and depth images are expanded by interpolation.

In the second stage, elemental images are generated from color images and expanded depth images. In the final stage, elemental images are displayed with the array of lenses in 3D autostereoscopic. In the first stage of capturing and expanding images, depth images and color images captured with a depth camera, Kinect V2, have the resolutions of 1920x1080 and 512x424, respectively. Since the depth resolution is lower than the color resolution, the elemental images generated then lose a lot of 3D information. To resolve this problem, we expand the depth images according to the size of the color images.

The simplest way to expand the depth images is to make empty depth images that are the same size as the color images, and then assign each of the depth images to them correspondingly based on the locations of the color images. This method cannot fully describe the real world because of many empty values in the expanded depth images.

Eq. 1 is modified to determine the pixel pitch as Eq. 2, where $RR = (RR_x, RR_y)$ is the resolution ratio. After that the size of color image is reduced to 512x424, Jeong et al. [8] gave the resolution ratio $RR$ of the reduced color images to $RR =$.
Since the enlarged 1570x1080 depth image is used in this paper, we define that of 1570x1080 color images as \( RR = \frac{512}{1570} \cdot \frac{424}{1080} \).

\[
P_T = RR \cdot \frac{Z_{\text{dp}}}{g} \cdot P_s \tag{2}
\]

Fig. 3 (a) and (b) show elemental images generated for the color and depth images with different sizes 512x424 and 1570x1080 for the same object space, respectively. Since the resolution of the 1570x1080 images is \( RR = \frac{512}{1570} \cdot \frac{424}{1080} \), their pixel pitch is about three times smaller than that of the 512x424 images as shown in Fig. 3.

The location of the \((u,v)\)th pixel of the elemental images to be generated is determined by using the color and depth information of the \((i,j)\)th pixel of the object space as in Eq. 3-5, which was presented by Jeong et al. [8]. In Eq. 3-5, \( P_L \) is the size of the elemental image and \( d_{(i,j)} \) is the depth of a transformed object, \( i_L \) and \( j_L \) are the index information of a lens array, respectively, and \( \min|d_i| \) and \( \max|d_i| \) are the minimum and maximum depths, respectively.

\[
u = P_k \cdot i_L - (i \cdot P_T - P_k \cdot i_L) \cdot \frac{g}{d_{(i,j)}} \tag{3}
\]

\[
v = P_k \cdot j_L - (j \cdot P_T - P_k \cdot j_L) \cdot \frac{g}{d_{(i,j)}} \tag{4}
\]

\[
d_{(i,j)} = \max\{d_i + \min|d_i|, Z_{\text{dp}}\}, \quad \frac{Z_{\text{dp}}}{2d_{(i,j)}} \cdot \frac{f \cdot g}{g - f} \tag{5}
\]

To generate elemental images based on GPU parallel programming techniques, a thread for each pixel of an object space is created to compute the color of the corresponding \((u,v)\)th pixel of the elemental images by using Equations (1-5). When we directly adopt the parallel algorithm presented by Jeong et al. [8], it often occurs that two different pixels \( P_1 \) and \( P_2 \) of the object space may correspond to the same elemental pixel, \( P_1 \) and \( P_2 \) depend on the order of thread operations, even though the pixel \( P_2 \), which is near to the elemental pixel \( DP \), should be chosen for the correction. We settled this nondeterministic phenomenon with an algorithm called Z-buffer [13]: the pixel near to the elemental images is selected by managing a buffer storing the depths from each pixel to the elemental images.

\[
\text{Fig. 4. Instance of two different object pixels corresponding to the same elemental pixel}
\]
Fig. 5. 1570x1080 original color image captured with Kinect V2 (a), 512x424 original depth image captured with Kinect V2 (b), 1570x1080 depth image expanded in brute-force (c), and 1570x1080 depth image expanded with bi-cubic interpolation (d)

To evaluate elemental images generated by the method proposed in this paper, we define three types, T1, T2, and T3 of test data for a 515x424 original depth image and a 1570x1080 original color image captured with Kinect V2 as follows:

T1: Reducing color image with the method of Jeong et al. [8]: 515x424 original depth image and 515x424 reduced color image.

T2: Expanding depth image with a simple non-interpolation method: 1570x1080 expanded depth image and 1570x1080 original color image.

T3: Expanding depth image with a bi-cubic interpolation method: 1570x1080 expanded depth image and 1570x1080 original color image.

We generated elemental images by the three cases under the same conditions; the size of the lens array was 5 mm, the number of lenses in the array was 30x30, the focus distance was 10 mm, the pixel distance of the display panel was 0.277 mm, and the elemental images were captured from 2 meters away. A few evaluation methods for integral images are proposed [14]. We evaluated the results of the three elemental images and the generated parallax images in terms of whether the whole image, background, human body, head, face, and arms are clearly displayed. The overall evaluation results are described in Table 2, and the details are shown in Fig. 6, Fig. 7 and Fig. 8.

Fig. 6 shows examples of generated elemental images, in which the parallax image size is 20x20 since the size of the elemental images is 600x600, and the size of the lens array is 30x30. To identify the evaluation of image quality in Table 2, we generated parallax images based on the elemental images from the viewpoints of top left, center, and bottom right, as respectively illustrated in Fig. 7. Generally, we can recognize more easily the shapes in Fig. 7 (b) than in Fig. 7 (a). It is verified that Fig. 7 (c) shows the least loss and distortion of images.

We tested the quality of auto-stereo image by visualizing elemental images with a 30x30 lens array without special glasses. Figure 8 shows examples of two photos produced with respect to the three cases, which were captured at 30 degrees to the left and from the front, respectively. The evaluation in Table 2 is supported by these results shown in Fig. 8. The images quality for T1 case as shown in Fig. 8(a) is low as a whole, and the head in particular is unrecognizable. Fig. 8(b) images for T2 case are cracked in the whole background and in
part of the right hand since there are ‘0’ values between depth pixels. As shown in Fig. 8(c) images for T3 case, we note that the overall shapes and the human can be more clearly recognized in comparison with the two other cases.

4. DISCUSSION

In this paper, we presented a method to increase the quality in integral imaging systems that provides auto-stereo images from the real world space. While previous methods reduce color images to fit into depth images to produce elemental images, we take a reverse approach of expanding the depth images to fit into the color images. The depth images are expanded by using an interpolation technique called the bi-cubic method, which results in an integral imaging system with higher image quality.

The computer technique of creating elemental images in real time for integral imaging by using depth cameras is still in the early stages, and more research is needed. For example, an effective method is needed to interpolate specific parts that are uncaptchaible in the left and right sides. In the future, we will extend our real-time integral imaging system to environments with high-density (HD) depth cameras.

ACKNOWLEDGEMENT

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

REFERENCES


Min-Ho Song
He received the B.E. in Computer Education at Chungbuk National University, Korea in 2014, and also received M.S. in Digital Informatics Convergence at Chungbuk National University in 2016. His research interests include computer graphics, and 3D digital contents.

Byung-Muk Lim
He received the B.E. in Computer Education at Chungbuk National University, Korea in 2015, and also received M.S. in Computer Science at Chungbuk National University in 2017. His research interests include computer graphics, and 3D digital contents.

Ga-A Ryu
She is a PhD student in Computer Science at Chungbuk National University. She received the B.E. in Cyber Security at Kwangju University, Korea in 2014, and also received M.S. in Digital Informatics Convergence at Chungbuk National University in 2016. Her research interests include computer graphics, and 3D digital contents.

Jon-Sung Ha
He is a professor who is working for department Game and Contents at Woosuk University, Korea. He received the B.E. in computer of computer engineering at Seoul National University, Korea in 1984, and also received M.S., Ph.D. in computer science from KAIST (Korea Advanced Institute of Science and Technology), Korea in 1986 and 1996, respectively. His research interests include computational geometry, computer graphics, CAD/CAM, and 3D digital contents.

Kwan-Hee Yoo
He is a professor who is working for Department of Computer Science at Chungbuk National University, Korea. He received the B.S. in Computer Science from Chonbuk National University, Korea in 1985, and also received M.S., Ph.D. in computer science from KAIST (Korea Advanced Institute of Science and Technology), Korea in 1988 and 1995, respectively. His research interests include a u-Learning system, computer graphics, 3D character animation, dental/medical applications.