

Enhancing Depth Accuracy on the Region of Interest in a Scene for Depth Image Based Rendering

Yongjoo Cho¹, Kiyoung Seo² and Kyoung Shin Park³

¹ Department of Media Software, Sangmyung University, Seoul, Korea 110-743
[e-mail: ycho@smu.ac.kr]

² Department of Computer Science, Dankook University, Yongin-si, Gyeonggi-do, Korea 448-701
[e-mail: windzard@empas.com]

³ Department of Applied Computer Engineering, Dankook University, Yongin-si, Gyeonggi-do, Korea 448-701
[e-mail: kpark@dankook.ac.kr]

*Corresponding author: Kyoung Shin Park

Received February 18, 2014; revised May 11, 2014; accepted June 18, 2014; published July 29, 2014

Abstract

This research proposed the domain division depth map quantization for multiview intermediate image generation using Depth Image-Based Rendering (DIBR). This technique used per-pixel depth quantization according to the percentage of depth bits assigned in domains of depth range. A comparative experiment was conducted to investigate the potential benefits of the proposed method against the linear depth quantization on DIBR multiview intermediate image generation. The experiment evaluated three quantization methods with computer-generated 3D scenes, which consisted of various scene complexities and backgrounds, under varying the depth resolution. The results showed that the proposed domain division depth quantization method outperformed the linear method on the 7-bit or lower depth map, especially in the scene with the large object.

Keywords: Depth Image-Based Rendering (DIBR), Linear depth quantization, Domain division depth quantization, Multiview intermediate image generation, 3D Display

1. Introduction

After a huge success of the movie Avatar, more 3D movies are produced for cinema given their greater sense of realism and entertainment value, and some existing films are re-released in 3D format. Most 3D movies released in theatres are also available on 3D Blu-ray disc so that people can enjoy 3D movies at home. Some companies such as Samsung and LG have sold 3D television sets for the home and various other display units are also being developed to support 3D formats. Major 3D television brands adopt the passive polarized, which uses light refraction, or the active shutter glass options, which alternately closes the left and right lenses. Although these methods produce realistic 3D images, users must wear special glasses to get the 3D effect, which can be a hassle for users. Hence, there has been much research on 3D display technologies that do not require the use of special glasses, such as holograms, integral imaging and multiview display. Each method has its share of advantages and disadvantages.

The multiview display provides multiple independent views of a scene taken from various viewing angles simultaneously. The most basic way of creating multiview contents is to setup the array of multiple cameras at each point of view and to take the picture at the same time [1]. However, not only is it very expensive to use several cameras at the same time, but also synchronizing each camera to take pictures at the same time is quite a difficult process. Furthermore, each camera needs to be calibrated because the intrinsic and extrinsic parameters of the cameras are all different. For these reasons, depth image based rendering (DIBR) is used instead to create multiview images that look like they have been filmed at various viewpoints [2,3]. In DIBR, center-view image and its associated depth image are used together to restore the 3D scene and synthesize the intermediate images from any point of view.

However, the intermediate images generated by DIBR cannot be completely restored to the original due to inadequate depth information and the occlusion problem. The reconstructed image may have holes because the needed information to restore the complete image does not exist. For example, a camera takes an image at the left-eye viewpoint where the object in front of the camera hides the object behind. When it takes the right-eye image, the occluded area is visible but no proper information is available from the left-eye image to fill the occlusion. Due to this occlusion problem, DIBR requires the precise depth information to properly restore the original image. Therefore, there have been many pre-processing depth image techniques [4,5,6,7] proposed to compensate for depth map or hole-filling algorithms [8] to produce better multiview intermediate images.

A depth map is an image with per-pixel depth information, indicating the distance of each three-dimensional data point from the center of the camera in the scene. Note that the depth map can only store a finite range of depth values. In computer graphics, the initial depth values are often quantized to 65536 scales (represented by a 16-bit depth image). However, more big holes are produced on DIBR intermediate images if the lower precision of depth map is used. In order to address this problem, it is important to utilize more detailed depth information from the depth map. In this research, we propose the domain division depth quantization method that provides the depth quantization with different specifications to a particular region of depth, to enhance the depth precision on the region of interest in a scene.

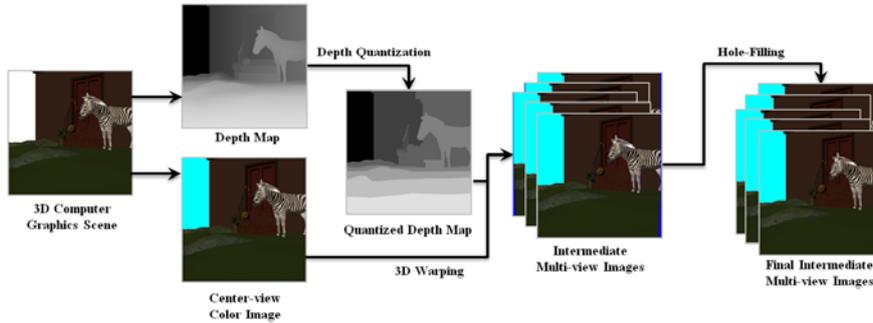


Fig. 1. The DIBR-based Multiview Intermediate Image Generation System using Depth Map Quantization for 3D Computer Graphics Scene

The domain division depth map quantization allows allocating more depth information in a particular region of the depth range to improve the quality of DIBR multiview intermediate images. This technique is developed to reduce the size of the depth data while minimizing the loss of image quality in DIBR. The size will be varied depending on the range of depth data that is allocated for the quantization. In general, people tend to focus more on interesting foreground objects than background. Let say if the objects are located in a particular region in the scene, the depth data for the rest of the objects is not so important. People may not feel difference even if the depth data is minimally allocated in the background area. If the objects are existed in 25% area of the whole depth range, it would be possible to allocate more depth data in the region of interest to reduce the amount of data needed for depth map representation.

Fig. 1 shows the multiview intermediate image generation system developed to evaluate the proposed depth map quantization method on DIBR. The system captures the center-view color image and corresponding depth map from 3D computer graphics scene. It then performs the depth map quantization where each pixel-depth value is rounded to the specific number of depth levels. Next, it processes the 3D image warping, which restores as a point cloud image by back-projecting the 2D points (per pixel from the center-view color image) into a 3D world depending on their depth value (using the quantized depth map). The virtual camera is then moved to different viewpoints to obtain the multiview intermediate color images. That is, 3D points are re-projected onto a 2D image plane of virtual camera. The intermediate image after the 3D warping contains holes, and the “fixed depth-weighted color average allocation” hole-filling algorithm [9] is applied to produce the final multiview intermediate color images.

In this research, we conducted an experiment to compare the linear depth quantization method against the domain division depth quantization with sixteen different quantization levels on the overall DIBR quality. In the experiment, each pixel depth of the scene was linearly (by equal distance intervals) or non-linearly (by domain specifications) quantized to the discrete number of pixel depth value, from the initial 16-bit (65536 scales) to 1-bit (2 scales) depth to generate the DIBR 9-view intermediate images. 3D computer-generated scenes selected in the experiment represent video with a variety of depth and foreground object complexity. The image quality was evaluated using Peak Signal to Noise Ratio (PSNR) of the original 9-view camera images directly taken from the scene and the 9-view intermediate images by DIBR.

This paper is structured as follows. Section II reviews related work on depth map quantization and coding in DIBR multiview intermediate image generation. Section III presents the proposed domain division depth quantization algorithm. Section IV describes a

comparative experiment on the proposed method evaluation. Section V will also discuss the experiment results. Section VI presents conclusions and future research directions.

2. Related Work

The recent studies in DIBR have investigated the depth coding, quantization and estimation to improve the quality of multiview image generation. Depth is the three-dimensional value indicating the distance between the camera and object. In general, the depth map captured from a depth camera is represented as an 8-bit image, 256 quantized values of distance. However, depth is linearly quantized in the disparity domain (i.e., pixel intensity), whereas it is non-linearly quantized in the depth domain (i.e. distance). In [10], the existing bi-directional prediction of the depth map cannot exactly predict the depth value for linearly moving objects. Thus, the non-linear bi-directional prediction is applied in the depth image for video coding to solve this problem due to the characteristics of non-linear depth quantization in the depth domain. This method performs 0.68 dB of the PSNR gain over the previous method.

There are a large variety of coding techniques for image compression, such as Discrete Cosine Transform (DCT) and wavelet-based JPEG-2000 standard, but they may not be optimal for encoding depth information [11]. Krishnamurthy et al [11] proposed a compression method of depth map for DIBR that considers region-of-interest (ROI) coding and reshapes the dynamic range of depth map to reflect the different importance of different depths. In this approach, ROI can be obtained from the depth estimation algorithm and reflect the accuracy of the depth map in that particular region or generated by an edge detector on the depth map. Then the depth map was passed through a function that expands or compresses the dynamic range for higher or lower depth. The experimental results show a performance improvement in coding, 1.1 dB of the PSNR gain, over the standard JPEG-2000 coder.

There are some studies to evaluate the effect of depth map quantization for DIBR. Kim et al conducted the systematic experiments to estimate the depth map quantization level that affords acceptable image quality on DIBR using the depth map from 16-bit to 1-bit depth [9]. The results indicates that a depth map above 7-bit is needed to maintain sufficient image quality for DIBR multiview intermediate image generation. Jung et al analysed the fundamental depth resolution and the factors for cardboard effect from DIBR process to perceived depth resolution on a multiview display [12]. Both numerical comparison with varying the depth resolution and subjective test using 9-view slanted lenticular display result shows that the depth resolution is saturated between 5- and 7-bit. The perceived depth resolution is lower than the fundamental depth resolution and it shows very similar distribution with the fundamental depth resolution.

Also, the minimum required depth accuracy on DIBR was studied with the navigation range of the virtual camera and the image resolution [13]. In this approach, a simple geometric problem formulation was developed to determine optimal depth scaling for a multiview video system and compared with the more general plenoptic sampling approach. Moreover, various elaborate depth map estimation algorithms were proposed for DIBR. Up-sampling the low-resolution depth from the depth camera image to the same size as the high resolution camera image was proposed to reduce the number of holes in 3D warping [5, 6]. A depth edge enhancement method coupling with a 3D propagation algorithm was proposed to efficiently correct and sharpen the depth edges in a low resolution depth image using high quality color images [7]. Segment-based propagation method composed of several non-overlapping color segments in a depth space was proposed to estimate more refined depth map [14].

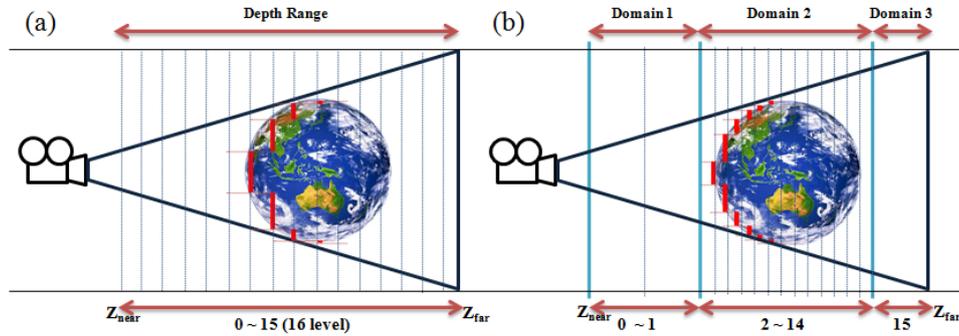


Fig. 2. Four-bit depth (sixteen scales) using the (a) linear and (b) the domain division depth quantization

Leon et al. studied the effects of depth image quality and compression on 3D video perception. The experiment was conducted on a 42 inch Philips WOWvx auto-stereoscopic display that used 9-view images created from a color image and a depth map (ranging from 0 to 255, i.e., 8-bit depth) [15]. Six videos were selected to represent video with a range of depth complexity. The video compression was done using H.264 depth coding at various qualities (i.e., 100-1000kbps). The experiment results showed that depth could be compressed significantly, to 0.0007 bits per pixel, without affecting the quality of 3D experience. However, the complexity of the depth image and motion had a strong influence on the acceptable levels of depth compression.

3. Depth Quantization

The quantization refers to the process of converting a continuous range of values into a finite range of discrete values, such as a function of analogue to digital signals. In this research, we define the depth quantization as a method of mapping the values from a fine-grained depth to a limited range of depth values. However, there is a loss of depth information during the quantization process which causes to make holes on DIBR intermediate view images. In this research, the domain division depth quantization is proposed to minimize the loss of critical depth information around the object of interest in the scene where accurate depth is most crucial. The quantized depth values are modified non-linearly by the domains. This technique aims to reduce the gap between quantization level intervals as a way to minimize the loss of depth information.

Fig. 2 shows 4-bit (16-scale) depth quantization where the vertical axis means the quantized value. All depth values of the object points will be replaced by the value of the reference point of their own quantization level. **Fig. 2** (a) shows the linear depth quantization where the depth is divided at regular intervals for the whole depth range. **Fig. 2** (b) shows the domain division depth quantization where 75 % of quantization depth bit is allocated to the object region (i.e., Domain 2) while 12.5% and 6.25% are assigned in the near or far regions (i.e., Domain 1 or 3) respectively. However, the number of depth bits assigned for the domain division depth quantization is the same as the linear quantization. It is assumed that the focused objects are located in a certain domain between the near and the far plane of camera so that a higher number of quantization scales can be allocated in this domain.

3.1 Linear Depth Quantization

The linear depth quantization is performed per pixel using the equation (1), where z is the depth value in the initial 16-bit depth map, 2^{16} is the maximum number of depth bits by computer graphics, and n is the quantization size. The quantization size, n , refers to how much quantized data is assigned in the depth map. If the quantization size is defined as 256, it can represent values from 0 to 255. This equation defines a simple linear mapping from the initial 16-bit depth range to the specific quantized depth range.

$$Q(z) = \frac{z}{2^{16}/n} \text{ with } z \in [0, \dots, 2^{16} - 1] \quad (1)$$

3.2 Domain Division Depth Quantization

The domain division depth quantization needs not only the quantization size but also the domain information. It also requires the ratio of the amount of depth information to be assigned for each domain. The domain information specifies how to divide the entire range of depths into the domains (i.e., regions of interest). The ratio of the amount of depth information represents as the percentage of allocating depth information for each domain, i.e., percentile for the full range of depth. This ratio must exist for all domains, and the sum of the ratios should be 100 %. The following steps describe the process of the domain division depth quantization.

First, the depth range of a scene is divided into domains by the relative proportions of distance from the view volume, in which each domain can apply different quantization depth bits using the ratio. Fig. 2 (b) shows the depth range divided into three domains; a higher percentage of quantization is assigned for the objects where people should pay more attention whereas a lower percentage is applied in the remaining areas. On the other hand, Fig. 2 (a) shows 100% quantization assigned in the entire depth range, which results in the linear depth quantization.

In computer graphics, each vertex of object is represented by a point (x, y, z) in a three-dimensional space. The depth value, z , in the depth buffer specifies the distance from pixel to the viewer. It is important to note that the relationship between the depth, z , and the distance, D , is non-linear when a perspective projection scheme is used. The non-linearity is affected by the near and far planes, z_{Near} and z_{Far} , of the camera perspective projection matrix. In practice, more depth precision is allocated for the region that is close to the near plane and less precision is used nearby the far plane.

Due to the non-linear characteristics of the depth value, the domains are partitioned using the actual distance information. To get the actual distance information, the depth is required to transform to the distance first. Using the equation (2), the depth value, z , is converted to the distance, D , by multiplying by the inverse of the camera perspective projection matrix. The difference between z_{Near} and z_{Far} are the depth range of a scene. This process makes to transform from the non-linear depth value proportional to the linear distance value. It is not necessary if the distance of a scene is already known.

$$D = \frac{z_{Near} * z_{Far}}{z_{Far} - z * (z_{Far} - z_{Near})} \quad (2)$$

Using the equation (3), the distance value is then converted to the normalized distance, \bar{D} , a value between 0 and 1. The normalized distance is obtained by dividing the distance by the maximum value of distance, T_D . T_D is represented by either the actual distance or the

percentile distance. This process helps to improve the calculation process. If a percentile is used for specifying the domain, an arbitrary actual distance can be calculated by multiplying the normalized distance with the maximum value of distance.

$$\bar{D} = \frac{D}{T_D} \quad (3)$$

Using the equation (4), the domain relative distance, rD , is calculated from the normalized distance. Note that the domain information is the accumulated data from the previous sections. To express this domain relative distance ranging from 0.0 to 1.0, it will have to change as the relative distance to the starting position of the domain. S_{min} and S_{max} mean the minimum and maximum value of the domain corresponding to the distance, respectively (i.e., the starting and ending position of the domain). That is, the relative distance value that belongs to a certain domain is calculated by subtracting the starting position of the domain from the distance and then dividing by the domain range from S_{min} and S_{max} .

$$rD = \frac{\bar{D} \cdot T_D - S_{min}}{S_{max} - S_{min}} \quad (4)$$

Using the equation (5), the quantized value is calculated from the domain relative distance. The number of domains is m , and i is the domain that the actual distance belongs to. $Spct_i$ means the percentage of depth bits allocated to the i -th domain. For example, if you set this domain ratio to 45%, $Spct_i$ will appear within the range of 0 to 0.45. The product of the domain relative distance and the quantization ratio for the corresponding section is added to the sum of the ratios assigned to the previous sections, $\sum_0^{i-1} Spct_i$, to interpolate the quantization values between domains. This result is then multiplied by the quantization size, n , to obtain the quantized value, Q , ranging from 0 to $n - 1$.

$$Q = Q(\bar{D}) = (rD \cdot Spct_i + \sum_0^{i-1} Spct_i) \cdot n \text{ where } 0 < i < m \quad (5)$$

Eventually, the quantized value obtained by the domain division depth quantization is different than the original depth information, which cannot be directly used as a depth for DIBR. Therefore, it requires the reverse quantization steps to return to the depth information. Using the equation (6), the quantized value, Q , is reverted back to the quantized normalized distance value, $Q\bar{D}$. Using the equation (7), the quantized distance value, QD , is then calculated by multiplying the quantized normalized distance value by the maximum value of distance. Using the equation (8), the quantized depth value, QZ , is finally obtained by multiplying the quantized distance value by the camera projection matrix.

$$Q\bar{D} = R(Q) = \left\{ \left(\frac{Q - \sum_0^{i-1} Spct_i}{Spct_i} \right) \cdot (Q \cdot S_{max} - Q \cdot S_{min}) + Q \cdot S_{min} \right\} \cdot \frac{1}{T_D} \quad (6)$$

$$QD = Q\bar{D} \cdot T_D \quad (7)$$

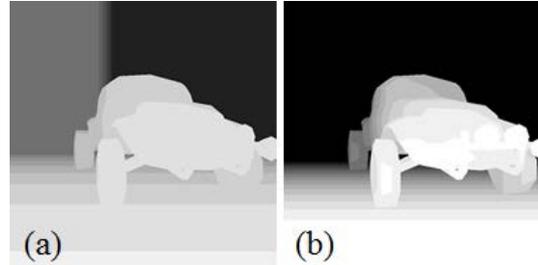


Fig. 3. Sixteen depth scales of the Car scene using the (a) linear and (b) domain division depth quantization

$$Q(z) = QZ = \frac{QD}{\frac{QD \cdot z_{Far}}{z_{Far} - z_{Near}} - \frac{z_{Far} \cdot z_{Near}}{z_{Far} - z_{Near}}} \quad (8)$$

In the domain division depth quantization, each domain assigns different quantization levels independently but the quantization per domain is actually the same as the linear depth quantization. If you assigns 100% quantization ratio for one entire section, the domain division depth quantization will be equal to the linear depth quantization.

Fig. 3 shows the gray-scale image of the depth map of the Car scene with 4-bit depth quantization using linear and domain division method. **Fig. 3** (a) presents the depth map linearly quantized to 16 depth scales using the equation (1). **Fig. 3** (b) presents the depth map non-linearly quantized using the equation (8) where 96% of depth bits were assigned in the Car object region. The pixel intensity of the depth image is the distance of the associated color image pixel to the viewer. In the 4-bit quantized depth map, the initial 16-bit per-pixel depth values are to be grouped into 16 scales ranging from 0 to 15, where 0 represents the closest to the viewer (foreground) and 15 represents the farthest from the viewer (background) in the scene. As shown in **Fig. 3**, more fine-grained depth scales are clearly appeared in the Car object region using the domain division method as compared to the linear method.

4. Experiments

The experiment was conducted to evaluate the effect of domain division depth quantization against the linear depth quantization. **Fig. 4** shows the center-view color image of the 3D computer graphics scenes used in this experiment: Car and Zebra scenes with different backgrounds (complex, simple and no texture background) and **varying** object sizes (large, medium and small). These scenes were chosen to find out whether the scene characteristics would affect on the quantization. The experimental data sets were the total of 18 different 3D scenes (2 scenes x 3 backgrounds x 3 object sizes). **Fig. 4** (a), (b), (c) shows the large Car object with no, simple and complex backgrounds. **Fig. 4** (c), (d), (e) shows the large, medium and small Car object with complex background. Similarly, **Fig. 4** (f), (g), (h) shows the large Zebra object with no, simple and complex backgrounds. **Fig. 4** (g), (i), (j) shows the large, medium and small Zebra object with simple background.

In this experiment, both color and depth image were rendered at an image resolution of 800x800 pixels. We assessed a 9-view camera array aligned in parallel. A parallel-aligned camera array looks at the depth of infinity, since each axis of multiple cameras is arranged in parallel where the center of a camera array is the fifth-view. The camera parameters were set to

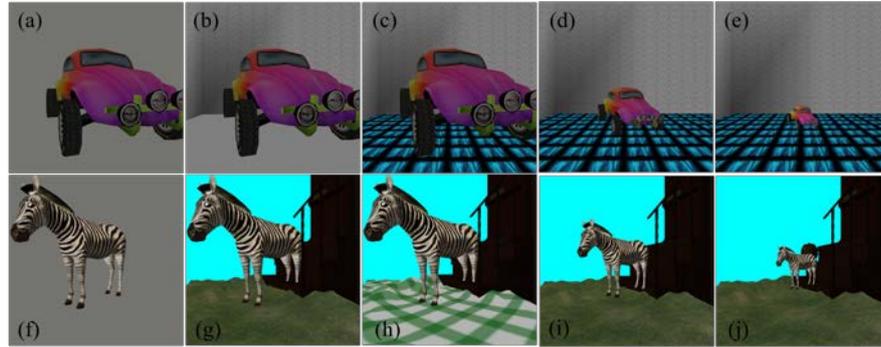


Fig. 4. The Car and Zebra scene with no/simple/complex background and large/medium/ small objects (a) Car-NL (b) Car-SL (c) Car-CL (d) Car-CM (e) Car-CS (f) Zebra-NL (g) Zebra-SL (h) Zebra-CL (i) Zebra-SM (j) Zebra-SS

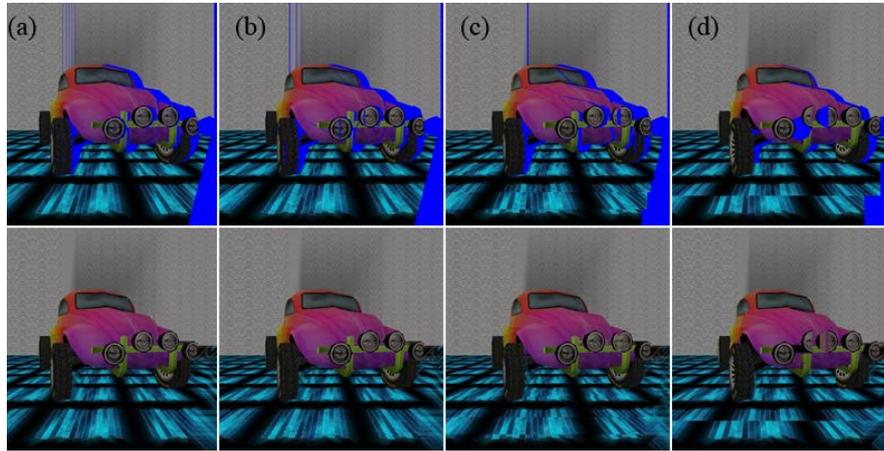


Fig. 5. The Car-CL scene taken at the 9th camera viewpoint by DIBR using domain division depth quantization with (a) 8-bit, (b) 6-bit (c) 4-bit, (d) 2-bit depth

the same ones used in our previous study. The gap between cameras was set to 32.5 mm, and the viewing distance from a virtual camera to the origin of the 3D scene was set to 6770 mm. The near and far clipping plane of the virtual camera view frustum was set to 500 mm and 10500 mm respectively.

The experiment evaluated the 9-view intermediate images generated by DIBR for each 3D scene, with the linear depth quantization and the domain division depth quantization, using various depth scales ranging from 16-bit depth (65536 scales) to 1-bit depth (2 scales). **Fig. 5** shows the intermediate images of the Car-CL scene taken at the ninth-view camera viewpoint generated by DIBR with the 8-bit, 6-bit, 4-bit and 2-bit domain division depth map quantization before and after hole-filling. As shown in **Fig. 5**, the locations and amounts of holes are slightly more increased when low-bit depth map quantization levels were used. In this case, the final intermediate image constructed with 8-bit or 6-bit depth map looks almost similar, whereas the contour artifacts at the borders increase slightly when the depth map quantization is 4-bit or below.

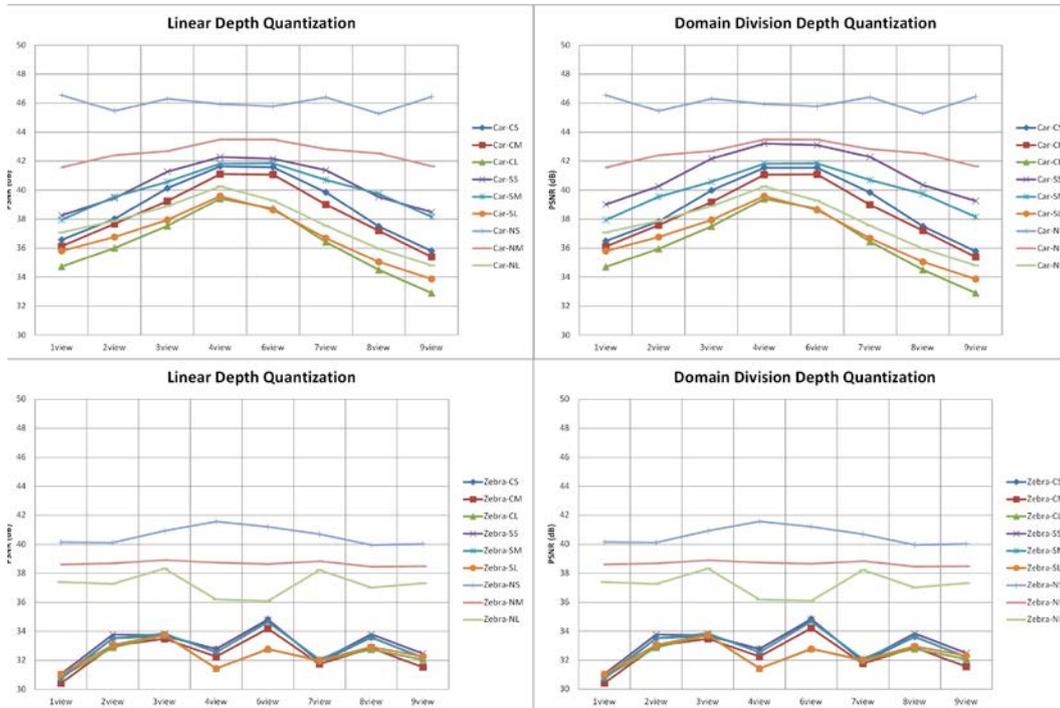


Fig. 6. Average PSNR at 9 camera viewpoints for 18 scenes with 16-bit depth map

The image quality was measured in the experiment using Peak Signal to Noise Ratio (PSNR) between the original 9-view camera images directly captured from the 3D scene and the 9-view intermediate images generated by DIBR with different depth quantization methods. The original camera captured images represent the ground truth view whereas the DIBR-based intermediate images represent the synthesized view. We calculated the average of PSNR values between the original camera images and the DIBR intermediate images over twenty sequences for each 3D scene captured from different camera view positions and angles in the scene.

5. Results and Discussions

The average PSNR value between the original 9-view camera captured images at each camera viewpoint and the DIBR 9-view intermediate images with the varying depth levels was measured for 20 frames of each 3D scene to evaluate the image quality degradation by the linear depth quantization versus the domain division depth quantization.

5.1 Linear Depth Quantization

First, we evaluated the effects of scene characteristics, such as background complexity and focused 3D object size, under two depth quantization methods. We evaluated the PSNR values between the 9-view camera images directly taken from the 3D scene and the 9-view intermediate images generated using DIBR with the initial 16-bit depth map for all 18 scenes. This result depicts the image quality degradation caused by DIBR multiview intermediate image generation. As shown in [Fig. 6](#), the average PSNR value for each camera viewpoint in

each scene was almost identical between linear and domain division depth quantization. The average PSNR was 36.26 ~ 46.02 dB (domain division) and 36.27 ~ 46.01 dB (linear) in the Car scenes, and 32.31 ~ 40.58 dB (domain division) and 32.30 ~ 40.57 dB (linear) in the Zebra

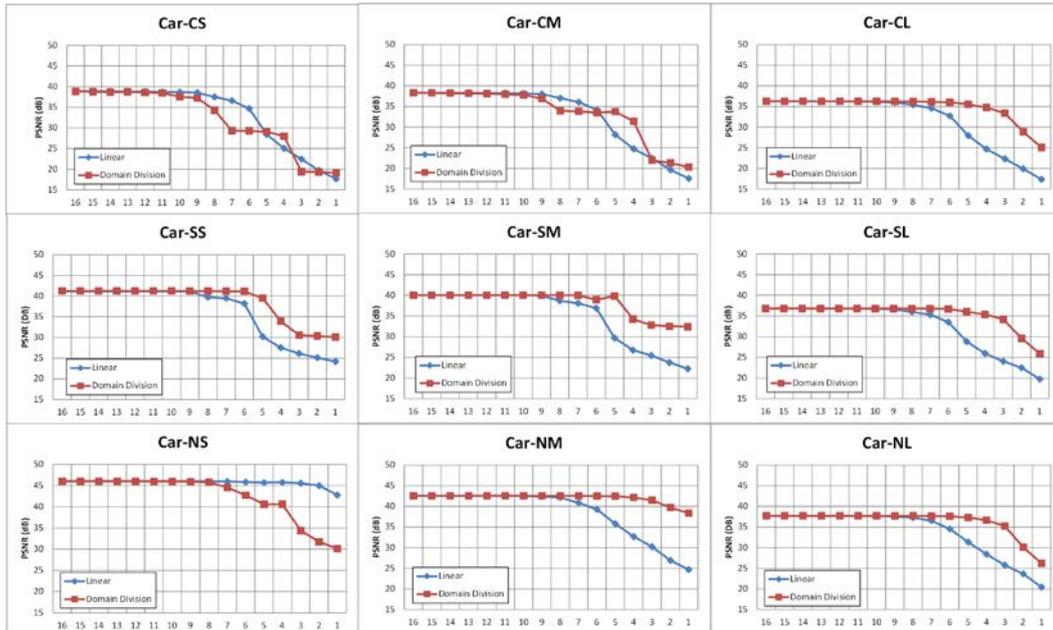


Fig. 7. Average PSNR between original 9-view images and DIBR 9-view intermediate images by 16 quantization levels for linear or domain division depth quantization for Car scenes

scenes. Overall the quality of DIBR multiview images was much lower for the scene with complex backgrounds than no backgrounds. Also, the PSNR was lower for the scenes with the large object under the same background, especially no background.

A three-way ANOVA test was conducted for the scenes with 16-bit depth to explore the impact of quantization method, scene characteristics and camera viewpoint on the PSNR. The main effect of quantization method was not significant ($p=0.715$ in Car and $p=0.62$ in Zebra). On the other hand, the main effect of scene characteristics and camera viewpoint was significant ($p=0.000$). The quantization method and scene characteristics interaction as well as the quantization method and viewpoint interaction was not significant ($p=1.000$). However, the scene characteristics and viewpoint interaction was significant ($p=0.000$). Intuitively, the less complex scenes (such as no background and small object) got higher PSNR. Also, PSNR was higher with no background in the Zebra scene. On the other hand, PSNR was lower with large object in the Car scene.

In addition, the image quality degraded when the camera viewpoint moved far from the center-view toward either of the sides. In the Car scenes, the image quality degradation was more clearly visible at the first view and the ninth view as compared to the fourth view and sixth view. In the Zebra scenes, the image quality was varied by the camera viewpoints.

5.2 Effects of Depth Quantization Methods with Varying Depth Levels

Fig. 7 shows the average PSNR value between the original 9-view color images directly taken from the 3D scene and the 9-view DIBR intermediate images constructed with 16 different

depth quantization levels (ranging from 16-bit to 1-bit) with two quantization methods (i.e., linear vs. domain division) for all 9 Car scenes (i.e., complex/simple/no background with small/medium/large car object). The bit indicates the bit of per-pixel depth on the depth map. That is, if the bit is 16 then the range of depth scales is 216, i.e., 65536. Overall the PSNR value was higher with no background but was degraded as the background was more complex.

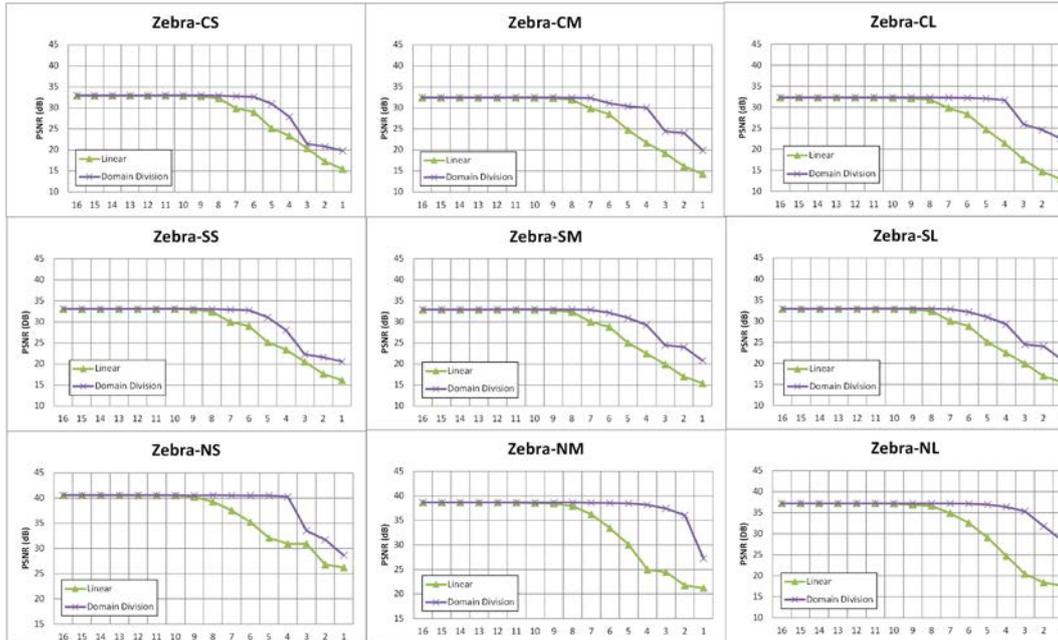


Fig. 8. Average PSNR between original 9-view images and DIBR 9-view intermediate images by 16 quantization levels for linear or domain division depth quantization for Zebra scenes

Also, the PSNR value was slightly higher with smaller object under the same background. The PSNR value with 8-bit or higher depth map was over 35 dB for all Car scenes regardless of quantization methods.

The effect of the depth quantization methods showed mixed results by the scenes. In the Car-CL, -SL, -NL, -SM, -NM, -SS scenes, the PSNR by domain division method was outperformed with 7-bit or lower depth. On contrast, the PSNR by linear method was better with 7-bit or lower depth in the Car-NS scene. For each scene, a two-way analysis of variance test was performed to investigate the effect of the depth quantization levels and the camera viewpoints on the PSNR. Regardless of the depth quantization methods, the PSNR was significantly degraded as the depth quantization level decreased and the camera viewpoint moved from the center to the side. Scheffe post-hoc comparisons revealed that 16- to 8-bit depth did not differ for the linear depth quantization method in all Car scenes except for Car-NS (that had no difference from 16- to 3-bit depth). In the domain division method, 16- to 6-bit depth did not differ in the Car-CL, -SS, -SM, -SL scenes and 16- to 4-bit depth did not differ in the Car-NM and -NL scenes.

Fig. 8 shows the average PSNR value between the original 9-view color images directly taken from the 3D scene and the 9-view DIBR intermediate images constructed with 16 different depth quantization levels (ranging from 16-bit to 1-bit) with two quantization methods (i.e., linear vs. domain division) for all 9 Zebra scenes (i.e., complex/simple/no background with small/medium/large zebra object). Similar to the Car scenes, the overall

PSNR value for the Zebra scenes was higher with no background than simple or complex background and it was slightly higher with smaller object under the same background. Also, there was no significant difference between two depth quantization methods with 8-bit or higher depth map for all Zebra scenes (over 32 dB PSNR). However, the PSNR by the domain division method was outperformed over the linear method with 7-bit or lower depth quantization levels regardless of the scene.

For each scene, a two-way analysis of variance test was performed to investigate the effect of the depth quantization levels and the camera viewpoints on the PSNR for each depth quantization method. In the linear method, the image quality was significantly degraded in DIBR as the depth quantization level decreased, especially below 7-bit depth for all Zebra scenes. Scheffe post-hoc comparisons also revealed that 16- to 8-bit depth did not differ for the linear method in all Zebra scenes. On the other hand, 16- to 6-bit depth did not differ for the domain division method in the Zebra-CS and -CS scenes. There was no difference from 16- to 5-bit depth in the Zebra-CL, -SS, -SM, -SL scenes and from 16- to 4-bit depth in the Zebra-NS, -NM and -NL scenes. The overall results showed that the domain division method outperformed the linear depth quantization method with the 7- to 4-bit depth map quantization levels where the average PSNR by domain division method was about 3 to 5 dB better.

6. Conclusion

The quality of DIBR multiview intermediate image generation highly depends on the accuracy of depth map to appropriately synthesize the views in dis-occlusion regions. Hence, there have been many pre-processing depth map approaches such as depth coding, quantization and estimation, or hole-filling algorithms proposed to compensate for depth map on DIBR. In this paper, we propose the domain division depth quantization method that provides the quantization with the specified number of depth bits assigned for a particular depth range in the scene. It allows allocating more depth information on focused 3D objects and makes the quality of the regions look better when DIBR intermediate images are generated. This domain division method is developed to reduce the amount of the depth data while minimizing the loss of image quality in DIBR. In this paper, this method is compared with the linear depth quantization method, a simple linear mapping from the initial depth to the specific quantized depth where the distance is divided at regular intervals for the whole depth range.

In the experiment, we evaluated the 9-view intermediate images for eighteen different 3D scenes (Car and Zebra scenes with no/simple/complex backgrounds and small/medium/large object) with the linear and domain division depth quantization, using various depth levels ranging from 16-bit depth (65536 scales) to 1-bit depth (2 scales). The 9-view camera array is aligned in parallel. When the domain division depth quantization is used, more fine-grained depth values are allocated in the region of interest (i.e., the area where objects are located) than the near or far regions of the objects. The image quality was measured using the average PSNR value between the ground-truth 9-view images directly taken from the 3D scene and the DIBR 9-view intermediate images over twenty sequences for each 3D scene taken at different camera view positions and angles.

The results showed that both methods did not produce the significant mean difference over an 8-bit or higher depth map (i.e., precise depth resolution). However, it also revealed that the domain division depth quantization method (assigning higher depth scales on the region of interest in the depth range) produced 1 to 15 dB improvements of DIBR multiview intermediate images over the linear depth quantization method under 7- or lower depth map although the PSNR gains varied depending on the characteristics of the scene. This

outperforming result by the domain division depth quantization method under lower depth bits was more clearly visible with the 3D object scenes with large object for both the Car and Zebra scenes. Since human has a characteristic of focusing on an important object in the scene, we expect on similar results with the subjective test. In the future, we plan to analyze the subjective evaluation of image quality about DIBR multiview intermediate images of 3D scenes displayed on a multiview 3D display.

References

- [1] Christoph Fehn, "Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV," in *Proc. of SPIE Stereoscopic Display and Virtual Reality Systems XI*, vol. 5291, no. 2, pp. 93-104, January 19-21, 2004. [Article \(CrossRef Link\)](#)
- [2] Christoph Fehn, "A 3D-TV Approach Using Depth-Image-Based Rendering (DIBR)," in *Proc. of Visualization, Imaging, and Image Processing*, pp.482-487, September 8-10, 2003. [Article \(CrossRef Link\)](#)
- [3] Svitlana Zinger, Luat Do, Peter H. N. de With, "Free-viewpoint depth image based rendering," *Journal of Visual Communication and Image Representation*, vol. 21, no. 5-6, July-August, pp. 533-541, 2010. [Article \(CrossRef Link\)](#)
- [4] Pascal Fua, "A Parallel Stereo Algorithm that Produces Dense Depth Maps and Preserves Image Features," *Machine Vision and Applications*, vol. 6, no. 1, pp. 35-49, 1993. [Article \(CrossRef Link\)](#)
- [5] Zefeng Ni, Dong Tian, Sitaram Bhagavathy, Joan Llach, B.S. Manjunath, "Improving the quality of depth image based rendering for 3D video systems," in *Proc. of the 16th IEEE International Conference on Image Processing*, pp. 513-516, November 7-10, 2009. [Article \(CrossRef Link\)](#)
- [6] Qingxiong Yang, Ruigang Yang, James Davis, David Nistér, "Spatial-Depth Super Resolution for Range Images," in *Proc. of IEEE Computer Vision and Pattern Recognition*, pp.1-8, June 18-23, 2007. [Article \(CrossRef Link\)](#)
- [7] Quang H. Nguyen, Minh N. Do, Sanjay J. Patel, "Depth Image-Based Rendering with Low Resolution Depth," in *Proc. of IEEE International Conference on Image Processing*, pp. 553-556, November 7-10, 2009. [Article \(CrossRef Link\)](#)
- [8] Luat Do, Svitlana Zinger, Peter H. N. de With, "Objective Quality Analysis for Free-Viewpoint DIBR," in *Proc. of IEEE International Conference on Image Processing*, pp. 2629-2632, September 26-29, 2010. [Article \(CrossRef Link\)](#)
- [9] Minyoung Kim, Yongjoo Cho, Hyon-Gon Choo, Jinwoong Kim, Kyoung Shin Park, "Effects of Depth Map Quantization for Computer-Generated Multiview Images using Depth Image-Based Rendering," *KSII Transactions on Internet and Information Systems*, pp.2175-2190, November 29, 2011. [Article \(CrossRef Link\)](#)
- [10] Kwan-Jung Oh, Sehoon Yea, Yo-Sung Ho, "Hole-Filling Method Using Depth Based In-Painting for View Synthesis in Free Viewpoint Television (FTV) and 3D Video," in *Proc. 27th Conference on Picture Coding Symposium*, pp. 233-236, May 6-8, 2009. [Article \(CrossRef Link\)](#)
- [11] Ravi Krishnamurthy, Bing-Bing Chai, Hai Tao, Sriram Sethuraman, "Compression and Transmission of Depth Maps for Image-Based Rendering," in *Proc of IEEE International Conference on Image Processing*, pp. 828-831, 2001. [Article \(CrossRef Link\)](#)
- [12] Jae-Hyun Jung, Jiwoon Yeom, Jisoo Hong, Keehoon Hong, Sung-Wook Min, Byoungho Lee, "Effect of fundamental depth resolution and cardboard effect to perceived depth resolution on multi-view display," *Optics Express*, pp. 20468-20482, Oct. 2011. [Article \(CrossRef Link\)](#)
- [13] Ingo Feldmann, Oliver Schreef, Peter Kauff, "Nonlinear Depth Scaling for Immersive Video Applications," in *Proc of 4th European Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2003)*, pp. 433-438, April 9-11, 2003. [Article \(CrossRef Link\)](#)
- [14] Sang-Beom Lee, Kwan-Jung Oh, Yo-Sung Ho, "Segment-based Multiview Depth Map Estimation Using Belief Propagation From Dense Multi-view Video," in *Proc. of IEEE, 3DTV Conference:*

The True Vision - Capture, Transmission and Display of 3D Video, pp. 193-196, May 28-30, 2008. [Article \(CrossRef Link\)](#)

- [15] Gustavo Leon, Hari Kalva, Borko Furht, “3D Video Quality Evaluation with Depth Quality Variations,” in *Proc. of IEEE 3DTV Conference: The True Vision – Capture, Transmission and Display of 3D Video*, pp. 301-304, May 28-30, 2008. [Article \(CrossRef Link\)](#)



Yongjoo Cho is currently an associate professor at the Department of Media Software, College of Software, Sangmyung University in Korea. He earned his Bachelor of Science in Computer Science at University of Illinois at Urbana-Champaign and Masters' of Science and Ph. D. degrees at University of Illinois at Chicago. He currently is the director of the Interactive Computing and Entertainment Laboratory at Sangmyung University. His research interest includes virtual reality, 3D display, tiled display system, computer supported cooperative work, and virtual learning environments.



Kiyoung Seo is currently working on a master degree at the Department of Computer Science, Graduate School, Dankook University, Korea. He got his Bachelor's degree at the department of Multimedia Engineering at Dankook University. His expertise is in 3D display, human-computer interface, tiled display systems and virtual reality.



Kyoung Shin Park is an associate professor at the Department of Applied Computer Engineering, Dankook University in Korea. She acquired her Ph. D. in Computer Science at the University of Illinois at Chicago (UIC) in 2003, with an emphasis in human computer interaction and graphics. She had also worked as a research professor at Digital Media Laboratory at Information & Communications University until she joined at Dankook University and was a visiting scholar at MIT Media Laboratory in 2004. Her research interest includes human computer interaction, virtual reality, augmented reality, collaborative environments, tiled display systems, multimedia systems and 3D imaging.