Enhancement of 3D Point Cloud Contents Using 2D Image Super Resolution Network

Seonghwan Park, Junsik Kim, Yonghae Hwang,
Doug Young Suh and Kyuheon Kim*

Kyung Hee University, Korea
E-mail: gocheenee@khu.ac.kr; junsik@khu.ac.kr; hyh717@khu.ac.kr;
suh@khu.ac.kr; kyuheonkim@khu.ac.kr
∗Corresponding Author

Abstract

Media technology has been developed to give users a sense of immersion. Recent media using 3D spatial data, such as augmented reality and virtual reality, has attracted attention. A point cloud is a data format that consists of a number of points, and thus can express 3D media using coordinates and color information for each point. Since a point cloud has a larger capacity than 2D images, a technology to compress the point cloud is required, i.e., standardized in the international standard organization MPEG as a video-based point cloud compression (V-PCC). V-PCC decomposes 3D point cloud data into 2D patches along orthogonal directions, and those patches are placed into a 2D image sequence, and then compressed using existing 2D video codecs. However, data loss may occur while converting a 3D point cloud into a 2D image sequence and encoding this sequence using a legacy video codec. This data loss can cause deterioration in the quality of a reconstructed point cloud.

This paper proposed a method of enhancing a reconstructed point cloud by applying a super resolution network to the 2D patch image sequence of a 3D point cloud.

Keywords: Point cloud, super resolution, deep learning network, 3D data.

1 Introduction

Media technology has been developed to provide users with immersive content. Additionally, transmission technology has been developed for active consumption of content, and compression technology for transmission has also been developed. Research, such as analyzing and reproducing media using deep learning networks, has been recently conducted.

Immersive media has been developed in 3D formats, such as Augmented Reality (AR) and Virtual Reality (VR) [1, 2], which are expressed using 3D data formats such as voxel, mesh, and point cloud [3]. Of these 3D data formats, the point cloud represents an object as a number of points and each point has coordinates and color information. Unlike 2D images that have a fixed size according to resolution, the size of the point cloud data is flexible according to the number of points and it is large. Because of the size of the point cloud, there is an issue with consuming point cloud data through transmission, which requires compressing the point cloud data. The Moving Picture Experts Group(MPEG), an international standardization group, has established an MPEG-I (Immersive) next-generation project group for developing a 3D point cloud data compression standard technology. Part 5 of MPEG-I, Video-based Point Cloud Compression (V-PCC) [4], is to propose a method of compressing the point cloud using the existing 2D video codec. In V-PCC, 3D point cloud data is projected along orthogonal directions into 2D image patches [5], which are put into a 2D image [6, 7]. This 2D image sequence is then compressed using a legacy 2D video codec [8]. Since V-PCC uses the conversion of 3D content into 2D image sequences and the lossy video codec, information loss is inevitable, which causes quality degradation of the reconstructed 3D point cloud. To improve the visual quality of the 3D point cloud, this paper proposes the use of a super resolution network [9–11] as a reconstructed 2D patch image sequence out of a legacy 2D video decoder. The super resolution network is designed for enhancing 2D images into high quality images in terms of subjective visual quality. The deep learning based super resolution network shows higher performance in subjective visual quality compared with the existing traditional methods,

and various networks are still studied. In this paper, Chapter 2 describes the underlying technology used for enhancement of point cloud data, and Chapter 3 describes the process of the point cloud enhancement technology using the super resolution network proposed in this paper. Chapter 4 presents and analyzes the results of the proposed method in terms of subjective and objective quality, and finally future work is suggested in Chapter 5.

2 Related Works

2.1 Video Based Point Cloud Compression

V-PCC is a method of compressing 3D point cloud using a legacy 2D video codec. The encoding process of V-PCC is explained in Figure 1, which begins with the 3D patch generation because the point cloud sequence has to be decomposed into patches to be compressed with a legacy 2D video codec. For 3D patch generation, points that have similar normal vectors are grouped based on six reference planes that are parallel to the XY, XZ, and YZ planes. In the patch packing process, patches that are projected into each projection plane are placed on a 2D grid. Additionally, an occupancy map is generated to indicate whether the patches are occupied on a 2D grid. The patches projected on the 2D grid are comprised of two pieces of information, such as coordinate based geometry and color based attribute, and then individual information is converted into a 2D image [5]. Among the coordinate information, the distance value to the projection axis is stored in the Y value of the geometry
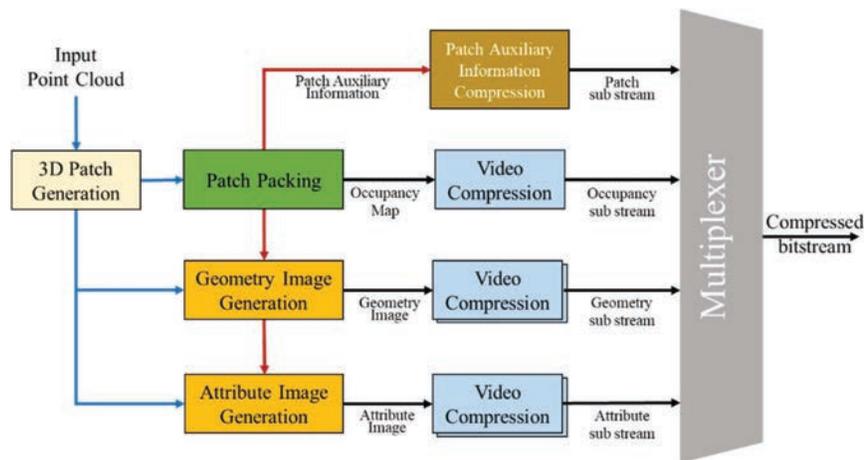

**Figure 1** Encoding process of V-PCC.
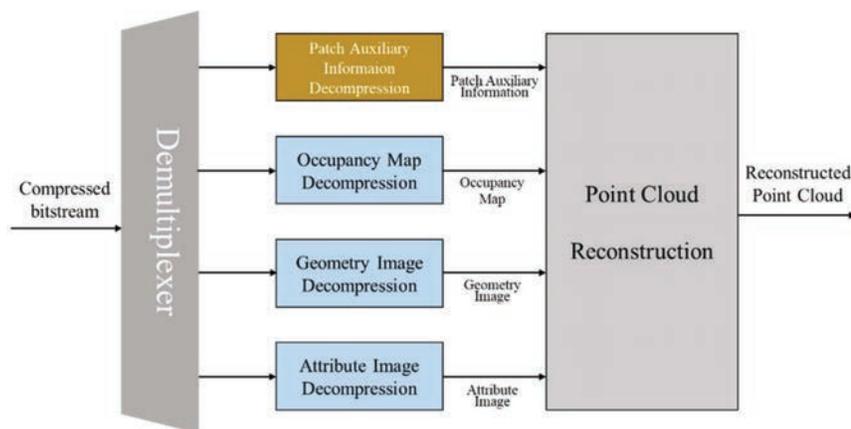
Figure 2   Decoding process of V-PCC.

image, and the attribute information is stored in the YUV value of the attribute image. When converting a 3D point cloud to a 2D image, overlapped points are lost in the projection process because multiple points cannot be stored in one pixel. To reduce this loss, one point cloud frame is stored in two geometry images [12]; for example, a near image that is comprised of points closest to the projection plane and a far image comprised of the furthest points. Accordingly, the attribute image is also comprised of two frames in one point cloud frame. Finally, the geometry image, attribute image, and occupancy map are compressed using a legacy video codec and delivered to the decoder. In addition, to reconstruct the patches located in the 2D grid to the point cloud, auxiliary patch information, such as the projection plane and position on the 3D, are required; thus, they are compressed through the auxiliary patch information compression module and also delivered to the decoder.

The decoding process of V-PCC proceeds in the reverse order of the previously explained encoding process, as described in Figure 2, where the geometry image, attribute image, and occupancy map are decoded using a legacy video codec, and patch information is acquired from geometry image and attribute image based on an occupancy map. Consequently, the generated 2D patch is reconstructed into a 3D point cloud using auxiliary patch information.

However, the reconstructed 3D point cloud has quality deterioration because of the loss that occurs when creating the patch and compressing the geometry and attribute images by the legacy video codec. In particular, the compression loss greatly influences the visual quality of the reconstructed

Figure 3   Original 3D point cloud (left) and Reconstructed 3D point cloud (right).

point cloud because it reduces the high frequency component, as shown in Figure 3. Thus, to solve the problem of visual quality degradation caused by compression loss, this paper proposed the method of applying a 2D super resolution network to the converted 2D image.

2.2 Neural Network Based 2D Super Resolution Networks

Single image super resolution (SISR) [13] is the process of upscaling and or improving the details within an image. Before the emergence of neural networks, the example based SISR [14] and sparse coding based SISR [15] were the most representative SISR methods. The example based SISR is a method of storing a pair of low/high resolution image patches in a dictionary, and the sparse coding based SISR is a method of encoding the input low resolution image with a sparse coefficient and then reconstructing it using a dictionary. After the emergence of the neural network, SRCNN [9] was proposed to improve the objective performance over traditional methods such as an example based method and sparse coding based method. Because SRCNN has only three layers, better performance is only achieved by stacking deeper layers.

When training a deep network, the gradient gradually decreases or increases in the backpropagation process, which is called a gradient vanishing/exploding problem. Parameters of the network may not be properly trained because of this gradient vanishing/exploding problem. Therefore, it is difficult to expect an increase in network performance by simply designing the deep network by adding layers. Thus, ResNet [10] proposed a new learning method called residual learning, so that even in deep networks, the
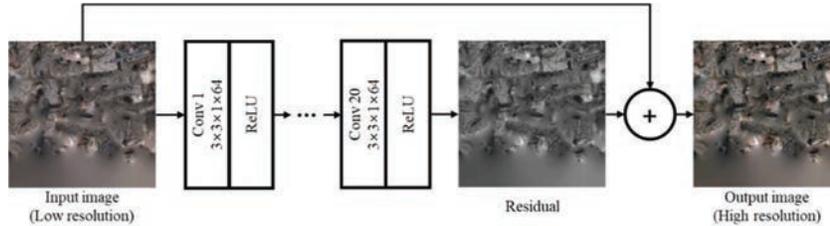
Figure 4   Structure of VDSR.

gradient vanishing/exploding problem does not occur, which improved the performance. Residual learning is a method of adding the input image to the final output image and learning the difference between the two images. Because both input and output images are similar, the difference is generally very small or zero, which makes it possible to avoid the gradient vanishing/exploding problem. VDSR [11] was able to build a deep artificial neural network by applying the residual learning method, which introduced a higher performance in the super resolution process. Thus, VDSR was defined and intensively overcame the limitations of SRCNN.

As shown in Figure 4, the VDSR proposed a deep 20-layer network using residual learning. The input is a upscaled, low-resolution image, and the output was a high-frequency signal used for reconstructing a high-resolution image. Thus, VDSR learns the high-frequency signal by residual learning between both the low-resolution and high-resolution images, and restores the super-resolution image by the sum of the low- resolution image and output high-frequency signal. On the basis of this analysis, this paper determined that VDSR is suitable for deep learning based super resolution.

3 The Proposed 3D Point Cloud Enhancement Using a 2D Super Resolution Network

In Chapter 2, Section 1 described the structure of V-PCC and the quality degradation caused by the loss in the compression process, and Section 2 introduced the neural network based 2D super resolution network for improving 3D point cloud. On the basis of the analysis in Chapter 2, this paper proposed a method of enhancing the visual quality of the 3D point cloud by applying a 2D super resolution network to a patch image and reconstructing a 3D point cloud using the enhanced image as shown in Figure 5.

As shown in Figure 5, the input original point cloud is decomposed using both patch generation and packing, and creates occupancy map, geometry
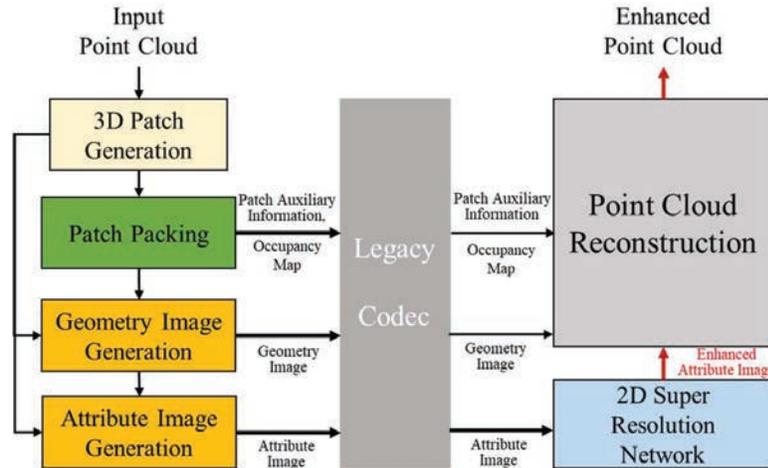
Figure 5 Architecture of the proposed 3D point cloud enhancement using a 2D super resolution network method.

image, and attribute image using the patch information. The generated 2D images and patch auxiliary information, required for point cloud reconstruction, are compressed and decompressed using a legacy encoder, and thus the compression loss occurs through this process. Among the decompressed 2D images, the attribute image with color information is enhanced with a 2D super resolution network, and the enhanced point cloud data is obtained by reconstructing the 3D point cloud along with the occupancy map, geometry image, and patch auxiliary information.

The detailed process of the proposed enhancement method in Figure 5 is described in the subsequent sections as follows. Section 1 describes the patch segmentation process that converts 3D point clouds into 2D images, and Section 2 describes how to enhance the attribute image which is generated through patch segmentation by using 2D super resolution network. Finally, the reconstruction process of a 3D point cloud is explained in Section 3 by combining the enhanced 2D attribute image with geometry, occupancy map, and patch auxiliary information.

3.1 3D to 2D Transformation

In this paper, the patch segmentation process of V-PCC was used to convert 3D point clouds into 2D images. The patch segmentation process combines the points into patches, which decomposes the point cloud into patches
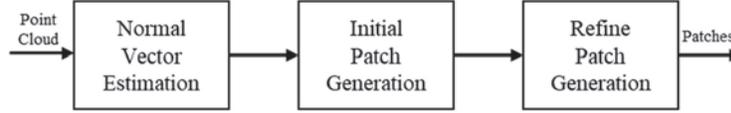
Figure 6   Patch segmentation process.

through three stages: normal vector estimation, initial patch generation, refine patch generation as shown in Figure 6.

Normal vector estimation is a process of calculating the normal vector of each point. The normal vector of each point can be obtained by estimating a plane with the nearest points. Thus, the nearest points of the point p=[x,y,z] have to be calculated for searching the normal vector. For searching nearest points, various nearest search techniques [16–19] such as KD-tree can be used. The normal vector of p is calculated using p and the nearest points set P:

$$\bar{p} = \frac{1}{m} \sum_{i=1}^{m} P_i \tag{1}$$

where, p_i means the elements of P, and the barycenter $\bar{p}$ of P is obtained with Equation (1). The normal vector $\vec{n}_p$ of point p is calculated from the eigen decomposition of $\bar{p}$ and p_i as follows:

$$\vec{n}_p = \sum_{i=}^{m} (p_i - \bar{p})(p_i - \bar{p})^T \tag{2}$$

Since $\vec{n}_p$ is calculated from all points in the point cloud, each point has one $\vec{n}_p$. In the Initial Patch Generation, the $\vec{n}_p$ of each calculated point is approximated by 6 unit vectors $\vec{r}_i$ which is comprised of (1,0,0), (0,1,0), (0,0,1), (−1,0,0), (0,−1,0), (0,0,−1) and then segmented into a patch with six reference planes. The approximation of each point $\vec{n}_p$ is determined by the vector closest to $\vec{r}_i$ as follows:

$$\max_i \{\vec{n}_p \cdot \vec{r}_i\} \tag{3}$$

The unit vector of each point is the same as the projection direction, and the projection direction of each point is determined according to the unit vector. Neighboring points that have the same projection direction are created as a single patch. However, because the number of patches is large and the patches are small, high frequency components are generated. These high frequency components reduces the compression efficiency of the 2D

video codec. Thus, the process of increasing the size and reducing the number of patches is performed using Refine Patch Generation which determines the optimal projection plane by considering the projection direction of the neighboring patches and $\overrightarrow{r}_i$, and integrates the patches that have the same projection direction into one patch. The patch created through the above process is a group of neighboring points that have the same projection direction. By placing them on the 2D Grid, the 3D point cloud can be converted into a compressible 2D image with a legacy video codec.

3.2 2D Patch Enhancement Using a Super Resolution Network

To enhance the 2D patch image, this paper proposed to apply the VDSR, which was introduced in Chapter 2. In super resolution, the resolution of the output image was determined according to the upscaling multiple. For example, when the upscaling multiple was 2 and the input image was 1280×1280, the output image is 2560×2560. As explained in Chapter 2, converting a 3D point cloud into a 2D patch image has produced three images such as an attribute image, geometry image, and occupancy map. When reconstructing a 3D point cloud, the attribute and geometry image refer to the occupancy map, and thus the resolutions of those three images must be the same. Consequently, the upscaling multiple for super resolution is applied to the attribute image and the geometry image and occupancy map.

The results of the super resolution network are shown in Figure 7, where a point cloud was not well reconstructed because the network was trained for color data with three channels. However, a geometry image is comprised of depth information with one channel, which is not suitable to a trained network with three channels. Thus, in this paper, a super resolution network is applied only to the attribute image, not to the geometry image. Additionally, the upscaling multiple was set to 1 to maintain the resolution of the geometry and attribute images as above.

The result of applying a super resolution network to the attribute image of patches is shown in Figure 8, indicating the super resolution network enhanced the subjective visual quality compared with the original image.

3.3 3D Point Cloud Reconstruction Based on Enhanced 2D Patch Images

VPCC performs a reconstruction process that converts a 2D image that is decoded from a video codec into a 3D point cloud. First, the decoded geometry image is converted to 3D points, and by merging attribute values

Figure 7   Reconstructed 3D point cloud with Super resolution network.

Figure 8   Original attribute image (red) and Enhanced attribute image (blue).

of the decoded attribute image to these 3D points, the 3D point cloud f_t^* was created. When the decoded geometry image is $G_t^*(u, v)$, the reconstructed 3D point $p^* = [x, y, z, r, g, b], (p^* \in f_t^*)$ is calculated as follows:

$$x_{p^*} = \text{s0} + u - \text{u0} \tag{4}$$

$$y_{p^*} = \text{r0} + v - \text{v0} \tag{5}$$

$$z_{p^*} = 0 + Y_{G^*_t(u,v)} \tag{6}$$

$$z_{p^*} = 0 + Y_{G^*_{t+1}(u,v)} \tag{7}$$

where the values x_{p^*}, y_{p^*}, and z_{p^*} are the x, y, and z coordinates of the reconstructed point p^*. Additionally, s0, r0, and 0 are the minimum coordinate values of the 3D patch which was decomposed from the original 3D point cloud, and u0 and v0 were the minimum coordinate values of the patch projected into the 2D plane. These s0, r0, 0, u0 and v0 values are described in the auxiliary patch information that was created in the patch generation process. Equations (4), (5), (6) and (7) are examples of 3D point reconstruction in a patch projected on the XY plane, and thus the Z position value was calculated using the Y value of the YUV values of $G^*_t(u, v)$. Because a 3D point cloud frame is converted into two 2D geometry images, $G^*_t(u, v)$ and $G^*_{t+1}(u, v)$ are used to reconstruct the point of the f^*_t as shown in Equations (6) and (7). The 3D reconstructed points which only have coordinate information get attribute values using the YUV value of the decoded attribute image $A^*_t(u, v)$. Since the attribute value of f^*_t corresponds to two attribute images $A^*_t(u, v)$ and $A^*_{t+1}(u, v)$, $A^*_t(u, v)$ is applied to the reconstructed point based on $G^*_t(u, v)$, and $A^*_{t+1}(u, v)$ is applied to the reconstructed point based on $G^*_{t+1}(u, v)$. Finally, the 3D point cloud reconstruction is finished by converting the YUV value of each attribute image into RGB values and storing the values in each point.

4 Experimental Results

In this paper, a 3D point cloud enhancement using a 2D super resolution network is proposed to compensate information loss that occurs in the V-PCC process converting 3D point cloud data into 2D images and video encoding using a legacy codec. To verify the proposed method, both the objective 3D PSNR and subjective visual quality were measured.

As shown in Table 1, PSNR is measured for four point cloud sequences, where 'V-PCC' and 'Proposed' show the PSNR of the reconstructed point cloud through the V-PCC process and the proposed enhancement method with the original point cloud, respectively. As explained in Chapter 3, the proposed enhancement method was not applied to the geometry image, and thus the geometry PSNR was not changed. Even the proposed enhancement method is applied to the attribute image; however, the attribute PSNR of the proposed enhancement method is less than that of the V-PCC. However, when

Table 1   3D PSNR comparison of the V-PCC and proposed methods

Sequence		Geometry PSNR [dB]		End-to-end PSNR [dB]		
		D1	D2	Luma	Chroma Cb	Chroma Cr
longdress	V-PCC	71.44	75.29	36.93	39.89	39.08
	Proposed	71.44	75.29	26.44	35.36	34.51
redandblack	V-PCC	70.69	74.67	40.07	43.24	39.31
	Proposed	70.69	74.67	28.27	34.74	33.82
loot	V-PCC	71.97	75.51	41.95	49.50	49.59
	Proposed	71.97	75.51	28.61	45.41	45.06
soldier	V-PCC	71.24	75.14	39.25	48.08	48.71
	Proposed	71.24	75.14	28.68	43.35	45.87

(a) Original   (b) V-PCC   (c) Proposed
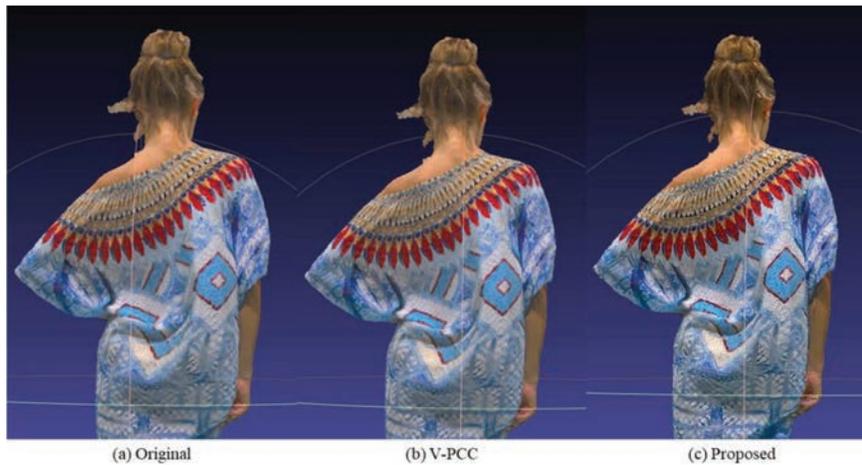
Figure 9   Result of longrdress content.

the reconstructed point clouds were compared in terms of subjective quality, as shown in Figures 9, 10, 11, and 12, the reconstructed clouds by the proposed enhancement method provide the remarkable subjective visual quality enhancement of the texture corresponding to the high-frequency component. On the basis of those objective and subjective results, the proposed enhancement method of using the 2D super resolution network can increase subjective visual quality, though it has difficulty in obtaining an attribute value closer to an original point cloud. The neural network generates values on the basis of similar patterns, which produces better visual quality [20–22].
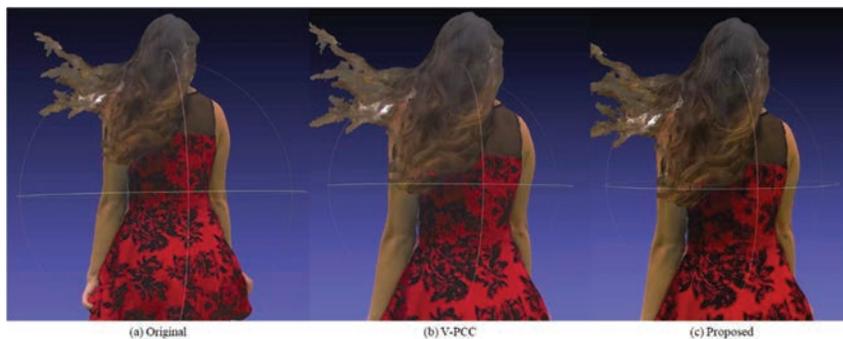
Figure 10   Result of redandblack content.
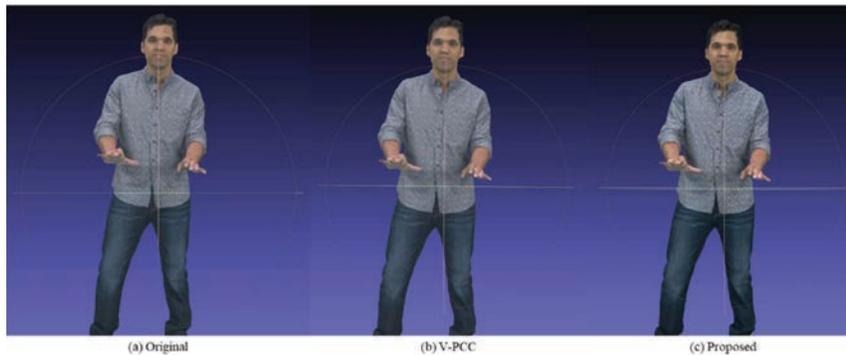
Figure 11   Result of loot content.

Figure 12   Result of soldier content.

5 Conclusion

Media technology has been developed to give users a sense of immersion. Recently, media data using 3D spatial data, such as AR and VR, has attracted attention. Point cloud is a data format that consists of a number of points for expressing 3D media data using 3D coordinates and color information for each point. Because the point cloud has a larger capacity than 2D images, a technology to compress it is required, and thus the international standard organization MPEG has developed V-PCC. V-PCC is a technology that decomposes 3D point cloud data into patches and project the patches orthogonally into a 2D image, followed by compression using legacy 2D video codecs. Data loss may occur while converting a 3D point cloud into a 2D image and encoding the 2D image using a legacy video codec. This data loss can cause quality deterioration of the reconstructed point cloud. Thus, in this paper, we proposed a method of enhancing point cloud data by applying a super resolution network to the 2D image of point cloud patch. When a super resolution network is applied to the attribute image, the objective result of PSNR was less; however, the subjective visual quality improved. It is necessary to research how to enhance geometry information by training separately geometry as the resolution of the 2D image has been enhanced.

Acknowledgement

References

[1] "Ubiquitous AR to dominate focused VR by 2022," https://techcrunch.com/2018/01/25/ubiquitous-ar-to-dominatefocused-vr-by-2022/, accessed: 2019.

[2] "Creating 3D Content with Reality Composer," https://developer.apple.com/documentation/realitykit/creating_3d_content_with_reality_composer, accessed: 2019.

[3] Emin Zerman, Cagri Ozcinar, Pan Gao, and Aljosa Smolic, "Textured Mesh vs Coloured Point Cloud: A Subjective Study for Volumetric Video Compression", 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX).

[4] "Text of ISO/IEC CD 23090-5 Video-based Point Cloud Compression," ISO/IEC JTC1/SC29/WG11 MPEG2019/N18670, Gothenburg, Sweden, Oct. 2019.

[5] H. Hoppe, T. DeRose, T. Duchamp, J. A. McDonald, W. Stuetzle, "Surface reconstruction from unorganized points," Proc. SIGGRAPH, pp. 71–78, 1992.

[6] Schwarz Sebastian, Miska M. Hannuksela, FakourSevom Vida, Sheikhi-Pour Nahid, "2d video coding of volumetric video data", Picture Coding Symposium (PCS) 2018 – Proceedings, pp. 61–65, 2018.

[7] S. Schwarz, et al., "Emerging mpeg standards for point cloud compression," in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 9, no. 1, pp. 133–148, Mar. 2019.

[8] "High efficiency video coding test model, HM16.18+SCM8.7," https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.18+SCM-8.7/, accessed: 2019.

[9] "Image Super-Resolution Using Deep Convolutional Networks", Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang, ECCV, 2014.

[10] "Deep Residual Learning for Image Recognition", Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[11] "Accurate Image Super-Resolution Using Very Deep Convolutional Networks", Jiwon Kim, Jung Kwon Lee and Kyoung Mu Lee, CVPR, 2016.

[12] E. S. Jang et al., "Video-Based Point-Cloud-Compression Standard in MPEG: From Evidence Collection to Committee Draft [Standards in a Nutshell]," in IEEE Signal Processing Magazine, vol. 36, no. 3, pp. 118–123, May 2019.

[13] "Super-resolution from a single image", Daniel Glasner; Shai Bagon; Michal Irani, IEEE 12th International Conference on Computer Vision, 2009.

[14] "Example-based super-resolution", W.T. Freeman; T.R. Jones; E.C. Pasztor, IEEE Computer Graphics and Applications, Vol. 22, Issue: 2, Mar/Apr 2002.

[15] "Image super-resolution as sparse representation of raw image patches", Jianchao Yang; John Wright; Thomas Huang; Yi Ma, IEEE Conference on Computer Vision and Pattern Recognition, 2008.

[16] S. A. Nene and S. K. Nayar, "A simple algorithm for nearest neighbor search in high dimensions," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 9, pp. 989–1003, Sept. 1997.

[17] H. Samet, "K-Nearest Neighbor Finding Using MaxNearestDist," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 243–252, Feb. 2008.

[18] K. C. K. Lee, B. Zheng and W. Lee, "Ranked Reverse Nearest Neighbor Search," in IEEE Transactions on Knowledge and Data Engineering, vol. 20, no. 7, pp. 894–910, July 2008.

[19] E. Agrell, T. Eriksson, A. Vardy and K. Zeger, "Closest point search in lattices," in IEEE Transactions on Information Theory, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[20] "Deep Networks for Image Super-Resolution With Sparse Prior", Zhaowen Wang, Ding Liu, Jianchao Yang, Wei Han, Thomas Huang, Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015, pp. 370–378.

[21] A Convolutional Neural Network Approach for Objective Video Quality Assessment, Patrick Le Callet, Member, IEEE, Christian Viard-Gaudin, and Dominique Barba, IEEE Transactions on Neural Networks, vol. 17, No. 5, September 2006

[22] Deep Learning of Human Visual Sensitivity in Image Quality Assessment Framework, Jongyoo Kim, Sanghoon Lee, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1676–1684.

Biographies

Seonghwan Park received B.S. in digital information engineering from Hankuk University of Foreign Studies, Yongin, South Korea, in 2013 and received M.S. in electronics engineering from Kyung Hee University, Yongin, South Korea, in 2015, where he is currently pursuing the Ph.D. degree in electronics engineering. His current research interests include MPEG systems, point cloud, super resolution and image processing.

Junsik Kim received the B.S. and M.S. degrees in electronics engineering from Kyung Hee University, Yongin, South Korea, in 2017 and 2019, respectively, where he is currently pursuing the Ph.D. degree in electronics engineering. His current research interests include point cloud compression, MPEG systems, digital broadcasting technologies, and image processing.

Yonghae Hwang received B.S. in astronomy & space science and electronics engineering from Kyung Hee University, Yongin, South Korea, in 2020, where he is currently pursuing M.S. in electronics and information convergence engineering. His current research interests include MPEG 3DG, V-PCC, point cloud and video codec.

Doug Young Suh (Member, IEEE) received the B.S. degree in nuclear engineering from Seoul National University, Seoul, South Korea, in 1980,

and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 1990. In 1990, he joined the Korea Academy of Industry and Technology, and conducted research on HDTV until 1992. Since 1992, he has been a Professor with the College of Electronics and Information, Kyung Hee University, Seoul, South Korea. He has been a Korean Delegate for the ISO/IEC MPEG Forum since 1996. His research interests include networked video and computer games.

Kyuheon Kim received the B.S. degree in electronic engineering from Hanyang University, Seoul, South Korea, in 1989, and the M.Phil. and Ph.D. degrees in electrical and electronic engineering from The University of Newcastle, Newcastle upon Tyne, U.K., in 1996. From 1996 to 1997, he was with Sheffield University, U.K., as a Research Fellow. From 1997 to 2006, he was with the Electronics and Telecommunications Research Institute, South Korea, as the Head of the Interactive Media Research Team, where he standardized and developed T-DMB specification, and conducted the Head of Korean delegates for MPEG standard body, from 2001 to 2005. Since 2006, he has conducted research at Kyung Hee University, Seoul. He has published numerous technical articles. His current research interests include interactive media processing, digital signal processing, and digital broadcasting technologies. He was a recipient of the Ministry Award from the Ministry of Information and Communication, in 2003 and the Prime Minister Award, in 2005.