Adaptive Segmentation and Stitching on 8K UHD Video

Jihee Yang[1], Youngae Kim[2], Jaehyeon Lee[2], Myeongsuk Yoon[2] and Gooman Park[2]
[1]*Dept.of Information Technology and Media Engineering, Seoul National University of Science and Technology, Seoul, South Korea.*
[2]*Dept. of Media IT Engineering, Seoul National University of Science and Technology, Seoul, South Korea.*
Jihee8929@seoultech.ac.kr

***Abstract*— Transmission technology is necessary to display contents stored on the Cloud server. When excessive compression is performed to the transmission of ultra-high resolution image, tangibility is reduced. Therefore, in this paper, we proposed an algorithm that divides the image into a number of sub-images. The sub-images were restored to the original one by stitching at the receiver. In the existing study, important objects were located at the center of the images, but an exception occurred when they are at the edge of the images. The saliency map was used to detect their main part so that the region of interest will not be divided. The images were divided depending on the position of saliency map. It is expected that users will be able to provide realistic signage by displaying ultra-high resolution images with a large screen.**

***Index Terms*— Image Segmentation; Image Stitching; Panorama; Transmission.**

I. INTRODUCTION

Expected as the next generation broadcast service, the 8K UHD video has the resolution of 7680*4320, which is 16 times larger than conventional HD video. The UHDTV provides viewers with clearer image quality and richer sound quality [1,2].

In videos, the amount of data increases exponentially as the resolution, color space, color depth, or frame rate rises. Although the 4K UHD channels have already operated in IPTV, the transmission technology of videos such as 360 degree or 8K has yet to be proven due to the large amount of data. Therefore, it is suggested that the transmission bandwidth should be expanded. In this case, users can watch video in high definition, when there is an increase in the channel bandwidth that causes a lower ratio of the compression However, the bandwidth size is limited as the performance of an excessive compression to transmit an ultra-high resolution image may result in data loss, which eventually makes it difficult to restore the image at the receiver part. Hence, it is necessary to solve the problem of the reality reduction from low quality contents and the difficulty of massive distribution of high quality VR contents through IP networks. It is also necessary to divide the ultra-high definition images into sub-images, while maintaining bandwidth [3,4,5,6,7]. At the receiver, the image restoration part through the stitching algorithm is required in order to provide users with the original image.

In this paper, we proposed a method of segmenting and transmitting 8K images based on the region of interest. The important parts of the input images are setting-up the region of interest and dividing them according to the region so that no losses occur within the region at the transmission part. At the receiver part, we stitch the sub-images to provide users with ultra-high definition.

In Section II, the related research techniques are explained, and the significance of this research is presented. An overall system is described and the algorithm of the transmitter and the receiver is presented in Section III. The experimental procedure and the results are presented in Section IV, and finally the conclusion of the proposed system is presented in Section V.

II. RELATED WORK

A. Image Division Transmission

There are many ways to split and transmit video. FHD image or video is generally transmitted by dividing it into four sub-images with the same size. In the case of 360 degree images, the storage and the transmit format of omnidirectional VR images are adapted with the OMAF (Omnidirectional Media Application Format) at the MPEG-I (MPEG Immersive). As users' receiving terminals become diversified, various types of storage, compression and transmission technologies are being studied.

B. The Necessity of Development

Result images should be divided differently according to the size and resolution of each immersive signage display. The segmentation could occur through the region of important object in the image, although the size and resolution of the display are just a consideration. Recently, the size of the bezel has been decreased, but the large-sized signage displays are not seamless at its boundary. Therefore, it is necessary to divide the image into a plurality of sub-images by considering the contents so that the object is not located at the boundary of the display.

For the purpose of this study, the conduct of the experiment considered the condition of the objects or materials since the objects or materials are displayed at the center of the image. The image preservation rate was higher than the conventional division method because the lines did not pass through their center portion. However, the condition of the object was not established if the important object was not positioned at the center of the image.

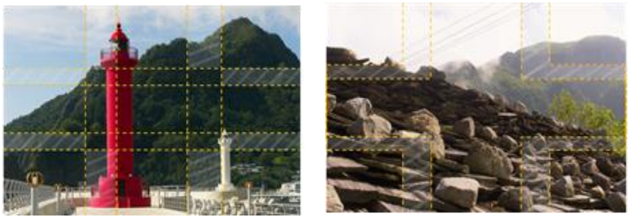


Figure 1: The necessity of development [9]

III. SYSTEMS CONFIGURATION

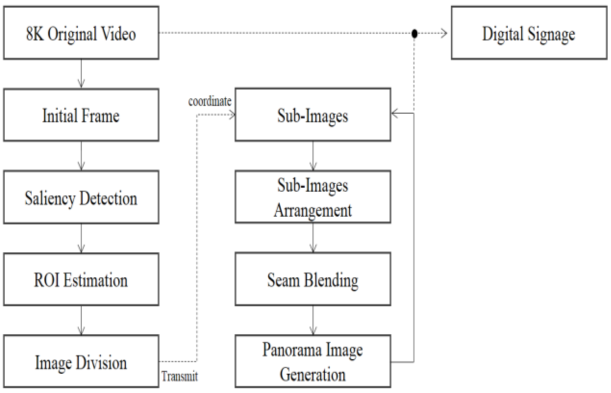


Figure 2: The overall proposed system

Figure 2 shows the overall configuration of the proposed system. In order to display a large amount of realistic contents such as 8K UHD images on a large display, the system was divided into a transmission part and a receiver part. At the transmission part, the area of the main object in the image was set-up as the region of interest. The image was divided and transmitted so that the division lines did not pass through the corresponding area. At the receiver part, the segmented images were stitched in regular sequence based on the coordinate values. The sub-images matched to original image were displayed through a large-scaled screen. In this paper, we analyzed only a single object image with a monotonous background to study the basic technology for system implementation.

A. Transmission

In the transmission, the position of interesting objects' area need to be detected to establish a criterion for image division. The location information was determined as an important region so that the image division lines did not pass through the region. In this case, the loss of the main part was prevented during the transmission of the image to the receiver.

In this paper, we used a context-aware saliency detection algorithm to analyze the image contents rather than dividing the important objects. In order to extract the saliency map, the images were characterized by three channels of color, intensity, and orientation. Further, the results were combined with a linear combination. The context-aware saliency detection algorithm was based on four principles.

- Local low-level considerations, including factors such as contrast and color.
- Global considerations, which suppress frequently occurring features while maintaining features that deviate from the norm.
- Visual organization rules, which state that visual forms may possess one or several centers of gravity, which the form is organized.
- High-level factors such as priors on the salient object location and object detection.

The principles are local low-level considerations, global considerations, visual organizational rules, and high-level factors. The algorithm distinguishes between foreground and background according to the scales values. Additionally, the final saliency map was generated. The context-aware saliency detection algorithm is shown in Figure 3[10].

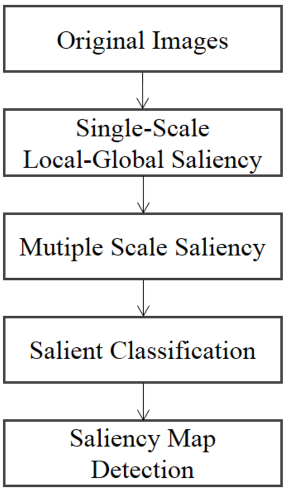


Figure 3: Saliency map detection algorithm

Drawn from this algorithm, we detected a major part in the image and set the area as a region of interest. The region of interest was set-up as $(x_{R\_left}, y_{R\_left}, x_{R\_right}, y_{R\_right})$ within the entire image area$(0, 0, x_{img}, y_{img})$. The pseudo-code for setting the division area based on the coordinate value is as follows.

Table 1
The Pseudo-code for Image Segmentation

Step1. Blob Labeling

$R_{roi}(x_{R\_left}, y_{R\_left}, x_{R\_right}, y_{R\_right});$

Step2. Image Segmentation

$\text{Rect } R_1 = \text{ new Rect}(0, 0, x_{R\_left}, 4320);$

$\text{Rect } R_2 = \text{ new Rect}(x_{R\_right}, 0, 7680, 4320);$

$\text{Rect } R_{roi} =$

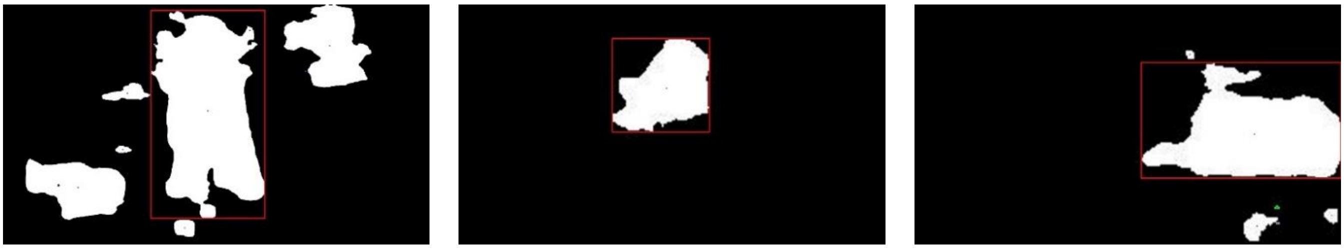
$\text{new Rect}(x_{R\_left}, y_{R\_left}, x_{R\_width}, y_{R\_height});$

※ R_{roi}: region of interest

In order to stitch the sub-images based on the coordinate values at the receiver, the coordinate information of the division line was included with a metadata type. Figure 4 shows the result of recognition using the saliency map to prevent the important object region, which is located in the sub-images. (a) is the input image, (b) is the saliency map image, and (c) is the result of detecting the region of interest labeling based on the saliency map. Figure 5 shows the sub-image segmented based on the setting region of interest.

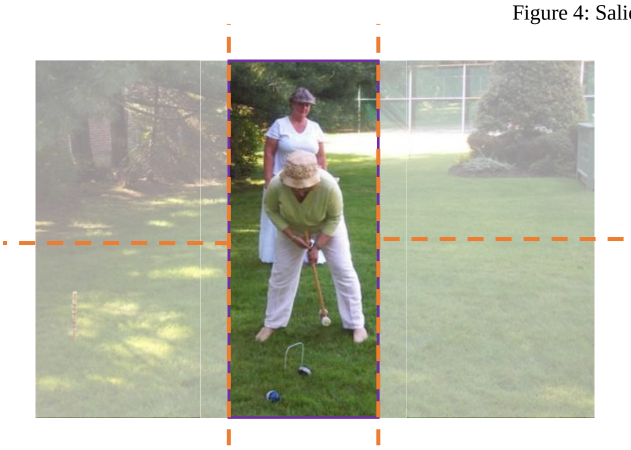
(a) Input image

(b) Saliency map

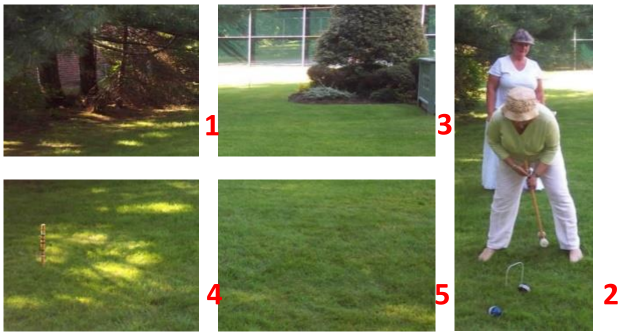


(c) Region of Interest

Figure 4: Saliency map detection



(a) Guide line



(b) Result

Figure 5: Result of sub-images

B. Receiver

In the receiver, the stitching algorithm was performed to display, divide and transmit the sub-images from the transmission. For image matching, the images were arranged in regular sequence. All of the size of the original one, the coordinate values of the division line, and the coordinate values of the ROI must be considered for generating overlapping areas when arranging the images. The images were segmented into sub-images as shown in Figure 6.

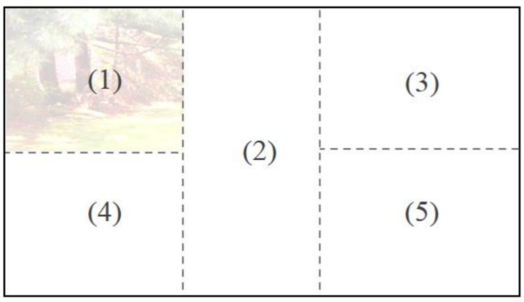
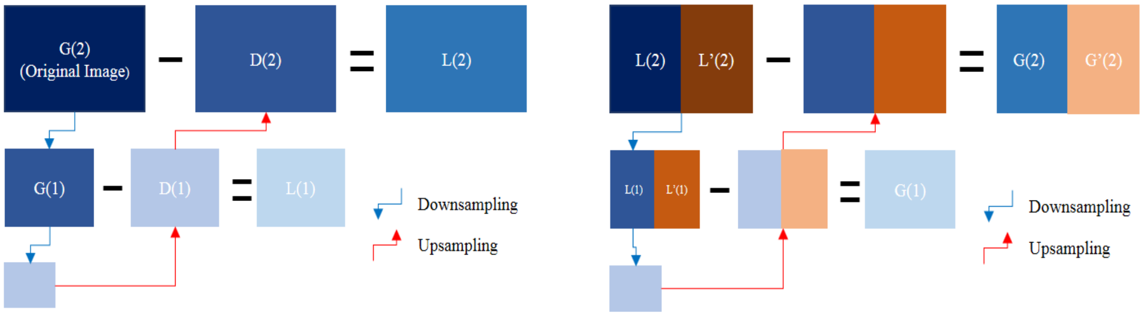


Figure 6: Placement of segmented image'

The seam was generated due to the image segmentation. Further, there was a loss of image during the transmission of the divided sub-images. Therefore, the correction was performed using the blending technique for seam. In this paper, we used a multi-band blending technique. The algorithm basically applies the method of pyramid image analysis and synthesis. It divides the image into Laplace and Gaussian. The Gaussian blur was performed on the input images to be stitched. The down-sampling was applied to the blurred image to create a Gaussian pyramid. Figure 7 shows an example of a multi-band blending technique [11].

Since the original ultra-high resolution image was divided based on the image, the probability of occurrence about color or distortion correction may be small. However, all possible defects were considered during the transmission of the image to improve accuracy. One ultra-high resolution image generated through the stitching process was displayed through a large-sized signage display.



(a) Pyramid image analysis course (b) Pyramid image synthesis process

Figure 7: Multi-band blending method

(a) Left input (b) Right input (c) Result

Figure 8: Multi-band blending result

IV. EXPERIMENT RESULTS

A. Experimental Environment

In this paper, the experimental environment was created using OpenCV 2.4.13 library based on c and c ++ language in Visual Studio 13.0. The performance of the system implemented was studied using ultra-high resolution image of 8K UHD from PASSTA dataset and MIT Saliency Benchmark dataset. From an arbitrary 8K image, an interesting object was detected to set a region of interest and eventually divided based on it. In order to assume the process of transmitting and receiving from a virtual server, Gaussian noise was applied to the whole images by image processing technique and image matching process was performed.

B. Comparison of Experimental Results

Table 2 shows the results of detecting the region of interest for the 8K images using the saliency map. The images used three non-noise images and two noise images.

Table 2
Coordinate Values About Region of Interest

	Images	Region of Interest Coordinate Values
Non-Noise (7680*4320)	gate ball	(2675, 115, 4706, 3836)
	fishing	(3991, 1593, 7680, 4320)
	snowboarding	(2763, 612, 4517, 2299)
Noise (7680*4320)	house	(2763, 612, 4517, 2299)
	beach	(1073, 2389, 4272, 3800)

Figure 9 and 10 show the result of dividing the image into sub-images, and stitching them through the proposed algorithm. In this paper, stitching was performed on the assumption that the sub-images are transmitted to the receiver. Stitching was performed using the coordinate information included in the sub images, and a single original image was generated by performing multi-blending. We confirmed that the original image was generated through the proposed algorithm.

(a) Segmented images

(b) Result

(c) Segmented images

(d) Result

(e) Segmented images

(f) Result

Figure 9: Non-Noise Ultra High Resolution(8K) Image Stitching

(a) Segmented imag

(b) Result

(c) Segmented image

(d) Result

Figure 10: Noise UHD(8K) image stitching

V. CONCLUSIONS

In this paper, an image segmentation method at the transmitter to minimize the loss of the important part and an image stitching method at the receiver was proposed. This work was motivated by the existence of bandwith limitation of the ultra-high resolution 8K UHD images in real time that use existing network. However, if the compression ratio of the image is increased, super-high resolution image cannot be provided to the user as it is. Therefore, in this paper, we use a method of dividing the image at the transmitter and re-matching it at the receiver. In the previous research, we assumed that the important objects or materials in the images are located at their centers. Since the division line is located not to pass through the center part of the image, the image preservation rate at the receiver is higher than before. However, when the important object is not at the center of the image, the condition was not established and an exception occurred. In the transmitter, an important part of the image is set as the region of interest, and the image segmentation is performed so that the division lines do not pass through the relevant portion. Therefore, it has adaptively prevented significant loss of contents due to data loss that may occur during image compression or transmission. In addition, in the receiver, the seamless image was generated by using the image matching method and the blending technique in order to remove the seam between the divided images. Future research will develop a limitation of the input image experiment environment to a single object image with a monotonous background to multiple object images with a complex background. Furthermore, a study on methods applicable to image segmentation and matching, and moving images will be processed.

ACKNOWLEDGEMENT

This work was partly supported by Institute for Information & Communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (No.2016-0-00099, Personal Broadcast Technology Development for Production Convenience and Maximum Viewing Experience)

REFERENCES

.

[1] M. Miyazaki, M. Abe, T. Sakiyama and N. Kurita, “State of the art OB vans for 8K-UHDTV and 22.2 multi-channel audio productions,” in Proc. of NAB Broadcast Engineering Conference (NAB Show), pp. 283-288, 2016.

[2] H. Takizuka, T. Torikai, A. Mitsui et al., “A proposal of novel optical interface to transmit 8K-UHDTV for consumer applications,” Microoptics Conference (MOC), 2013.

[3] B.C. Kim, G.H. Lee, I.J. Lee, K.H. Kim, "An Efficient Transmission Method of Panoramic Multimedia Contents in a Limited Bandwidth Environment", Journal of Broadcast Engineering, Vol. 16, No. 5, 2011, pp. 811-823

[4] M. Miyazaki, M. Abe, T. Sakiyama and N. Kurita, “State of the art OB vans for 8K-UHDTV and 22.2 multi-channel audio productions,” Proceedings of NAB Broadcast Engineering Conference (NAB Show), 2016, pp. 283-288.

[5] H. Takizuka, T. Torikai, A. Mitsui et al., “A proposal of novel optical interface to transmit 8K-UHDTV for consumer applications,” Microoptics Conference (MOC), 2013.

[6] S. Fujihara, “FLAPE : An essential storage technology in 4K and 8K business,” Proceedings of NAB Broadcast Engineering Conference (NAB Show), 2016, pp. 217-223.

[7] M. Naccari, A. Gabriellini, M. Mrak et al., “HEVC coding optimisation for ultra high definition television services,” Picture Coding Symposium (PCS), 2015, pp. 20-24.

[8] S. Koshy, M.N.S. Swamy, “Intellectual Signage: A New Method of Data Capture”, International Research Journal of Engineering and Technology(IRJET), Vol. 03, No. 08, 2016

[9] Y.A. Kim, J.H. Yang, G.M. Park, “8K UHD Video Segmentation and Stitching”, Proceedings of Asia Pacific International Conference on Information Science and Technology(APIC-IST), 2017.

[10] S. Goferman, L. Zelnik-Manor, and A. Tal, “Context-Aware Saliency Detection”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 10, 2012, pp. 1915-1926.

[11] M. Brown, D. G. Lowe, “Automatic panoramic image stitching using invariant features,” International Journal of Computer Vision, Vol. 74, NO. 1, 2007, pp. 59-73.