VHFp1-6L
Late-News Paper

3D CG Image Region of Interest Estimation and Visual Attention Based on Saliency Map

Norifumi Kawabata[1,2]

E-Mail: norifumi@mdsc.hokudai.ac.jp; norifumi@nagoya-u.jp
[1] Education and Research Center for Mathematical and Data Science, Hokkaido University, Sapporo, Japan
[2] Research Division of Computational Imaging, Computational Imaging Lab, Sapporo, Japan
Keywords: 3D CG Image, Region Of Interest (ROI), Saliency Map, Visual Attention, Stochastic Model

ABSTRACT

In this paper, we estimated that there is the region of interest (ROI) in which object or background region of 3D CG image from the view of stochastic model for saliency map, and then we evaluated and discussed quantitatively including visual attention.

1 Introduction

In the case of seeing images and videos, human judge by themselves whether these images and videos are importance region or not. To reproduce human's high-quality intelligence on computer, thus far, there were studies such as selecting image information using computer image processing according to level of importance [1],[2]. On the other hand, changing of high-definition and high-quality images and videos such as Quad Full HDTV (QFHDTV: 4K UHDTV), Super Hi-Vision (SHV: 8K UHDTV), as contents creators, it is very important to create contents after perceiving in advance that there are interests for users in which regions, in how patterns. Actually, in case contents creators compose on purpose, it is not always for users to require on purpose. And then, between subjective evaluation by human and objective evaluation by computer is the difference. Therefore, from the view of visual attention, we need to develop image information system enabled to acquire visual information such as focus points in advance, and verify this system after development. In this paper, we estimated from the view of stochastic model in saliency map that there is region of interest in which object or background regions of 3D CG image, and then, we evaluated these results quantitatively, finally discussed.

2 Related works

In this section, we describe the related works from the view of (1) saliency map, (2) visual attention, and (3) image quality assessment. For saliency map, Spectral Residual method is proposed [3], [4]. In this method, they extracted characteristics by processing in Fourier space such as frequency transformation and representation space for image data represented in two-dimension. For visual attention, stereoscopic visual attention guided disparity control for multi-view image, which is studied [5]. For image quality assessment, pre-attention and spatial dependency driven no-reference image quality assessment, which is studied [6],[7]. Overall, we describe including saliency map, visual attention, and image quality in this study.

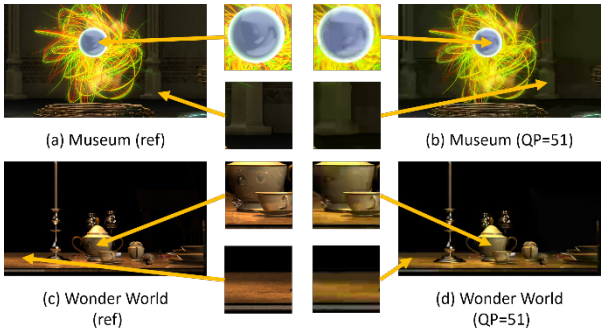

Fig. 1: 3D CG still images used in this study ("Museum" (upper), "WonderWorld" (lower))

3 Experimental setup

In this study, we used 3D CG contents (Museum (M), Wonder World (W)) provided by NICT for a free of charge as shown in Fig.1. For image generation, first we construct 8-viewpoints CG cameras, and then we processed of camera work, rendering. After that, we generated 8 viewpoints still images. Originally, this content is included multi-view 3D image. In this study, we used 1-viewpoint (single-viewpoint) image of 8-viewpoints still images. Saliency map, which is calculated saliency in each pixel of image or video that it is easy to gaze in which part of image or video from the view of visual attention. In saliency map, both representation by grayscale and heat map by color, which are used in general. In this study, we represent as saliency map by using grayscale. The implementation of saliency map is used by Python 3.6.13 and library embedded in OpenCV 3.3.1. Here, Spectral Residual Model is shown in Eq. (1), and to implement saliency map, StaticSaliencySpectralResidual_create() such as library embedded in OpenCV 3.3.1, which can be used.

$$H(Image) = H(Innovation) + H(Prior\ Knowledge) \quad (1)$$

Next, we consider transformation from spectral residual (SR) to saliency map as shown in the following Eq. (2).

$$H(R(f)) = H(L(f)|A(f)) \quad (2)$$

On the other hand, in OpenCV 3.3.1, it is possible to implement by using StaticSaliencyFineGrained_create() to compose intensity map. Human pay attention to somewhere or something on a daily basis, and this behavior is based on process of brain. For visual attention, it is not defined specifically, however, at present, it is known for a lot of interesting phenomenon related to attention. For an example, these are such as spatial attention, attention effect, objective attention, visual search, associative error, and physiological aspects etc. In this study, we are able to estimate that spatial attention, attention effect, and objective attention for something, which are appropriated to consider visual attention on an image. As experimental method in this study, we composed the saliency map algorithm by using the integrated development environment such as Anaconda, Python 3.6.13, and OpenCV 3.3.1, and then implemented as composing computer program originally. Loading image is used as 3D CG image encoded and decoded by H.265/HEVC. Quantization Parameter is set 7 types of $Q = ref, 20, 25, 30, 35, 40, 51$ in our experiment. As evaluation method in this study, we see whether we are able to measure saliency or not by observation to compare between original and evaluation images. The subjective assessment by carrying out evaluation experiment cooperated with assessors, which is a problem in the near future. In this study, we focused on how saliency can be measured as objective metrics.

4 Experimental results and discussion

As experimental results, result images after processing Spectral Residual are shown in Fig. 2. On the other hand, result images after processing Fine Grained are shown in Fig. 3. From experimental results, from the view of observation, for “Museum”, in “SpectralResidual (SR)”, saliency around objective border frame can be seen regardless to Q. On the other hand, for object region such as a ball, it is seen as jet black, therefore, we are not able to see saliency. In “FineGrained (FG)”, object region for “Museum” cannot be classified, however, background region can be seen as region segmentation. For “WonderWorld”, in “SR”, texture regions of an image can be seen as saliency regardless to Q. On the other hand, for background region, saliency cannot be seen. In “FG”, chair region of an image is lack by processing the coded degradation, however, classification in the case of “FG” can be seen than that in the case of “SR”. Thus, by observing saliency map from aspects both “SR” and “FG”, this is enabled to perceive the difference and characteristics of saliency score between object region and background region.

5 Conclusion

From results in this study, as knowledge and novelty in the case of this study, CG contents dependence is not hardly seen, and then we estimate enabled to specify of saliency by using both results of “SR” and “FG”. As the near future

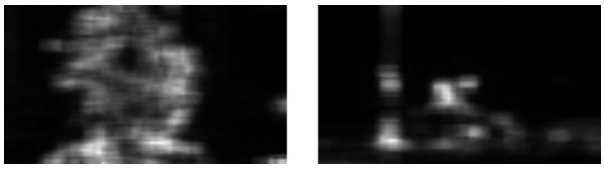

Fig. 2: Spectral Residual (SR), $Q = 30$, “Museum” (left side), “Wonder World” (right side)

Fig. 3: Fine Grained (FG), $Q = 30$, “Museum” (left side), “Wonder World” (right side)

work, if we carry out the localization by using bounding box, quantitative assessment from digitalization and visualization by using texture analysis, and discussion for general saliency map and visual attention considering the subjective evaluation and gaze evaluation, we will expect enabled to condition perception in more detail. This study is improved contents presented in the Forum on Information Technology (FIT2021).

References

[1]. N. Kawabata and M. Miyao: “3D CG Image Quality Metrics by Regions with 8 Viewpoints Parallax Barrier Method,” *IEICE Trans. on Fundamentals*, Vol.E98-A, No.08, pp.1696--1708, August 2015.

[2]. N. Kawabata and M. Miyao: “Multi-view 3D CG Image Quality Assessment for Contrast Enhancement Based on S-CIELAB Color Space,” *IEICE Trans. on Inf. and Syst.*, Vol.E100-D, No.07, pp.1448--1462, July 2017.

[3]. X. Hou and L. Zhang: “Saliency Detection: A Spectral Residual Approach,” *IEEE CVPR2007*, 8 pages, June 2007.

[4]. M. Cheng, G. Zhang, N. J. Mitra, X. Huang, and S. Hu: “Global contrast based salient region detection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.37, No.3, pp.569--582, March 2015.

[5]. J. Lei, S. Li, B. Wang, K. Fan, and C. Hou: “Stereoscopic Visual Attention Guided Disparity Control for Multiview Images,” *J. Display Technology*, Vol.10, No.5, pp.373--379, May 2014.

[6]. L. Liu, T. Wang, and H. Huang: “Pre-Attention and Spatial Dependency Driven No-Reference Image Quality Assessment,” *IEEE Trans. on Multimedia*, Vol.21, No.9, pp.2305--2318, September 2019.

[7]. D. Chen, Y. Wang, and W. Gao: “No-Reference Image Quality Assessment: An Attention Driven Approach,” *IEEE Trans. on Image Processing*, Vol.29, pp.6496--6506, May 2020.