Journal of Artificial Intelligence and Data Science (JAIDA)
e-ISSN: 2791-8335
https://dergipark.org.tr/pub/jaida

Vol. 3, No. 1, pp. 50-53, 2023
Research Article
Received: 10/05/2023 Accepted: 26/06/2023

Edge Detection for 3 Dimensional Video Quality Assessment

Gökçe NUR YILMAZ *, Yücel ÇİMTAY

Abstract
The impact of 3 Dimensional (3D) video technologies on user perception of 3D videos can only be highlighted by considering the impact of these technologies on user perception of 3D videos. It is true that enabling this can be accomplished by taking into account the crucial criteria of defining the nature of a 3D video. Under the light of this fact, an edge detection algorithm relying on the Canny edge detector, is used to derive preliminary results to develop a No Reference (NR) model for the VQA of a 3D video. The performance evaluation results obtained by using the proposed algorithm prove that proposed study is quite effective in the VQA of a 3D video.
Keywords: 3D video, edge detection, video quality assessment.

1. Introduction

Improving the performance of three-dimensional (3D) multimedia services is heavily dependent on meeting two goals. One of these goals is to achieve maximum efficiency during the transmission of 3D video contents, while the other is to improve the viewing experience at the user end. These two goals can only be addressed by developing costeffective and Human Visual System (HVS)-friendly 3D Video Quality Assessment (VQA) metrics [1][2][3][4].

The VQA is essentially accomplished in two ways in the literature. One of these ways is subjective testing, in which real viewers are used for VQA. Another approach is to use computer algorithms to perform objective evaluation. Because real viewers are employed during the VQA processes, the data provided by subjective tests is assumed to be correct. However, subjective tests are both costly and time consuming. As a result, objective measures are preferred by researchers attempting to accomplish the goal of optimum transmission efficiency due to their low cost and speedy processes [5][6].

In the literature, the objective metrics are classified into three groups based on the use of the original video at the receiver side during the VQA process. One of these categories is the Full Reference (FR) metric, which requires the original video to be present at the receiver in order to compare the quality of the compressed (distorted) video to that of the original. Another category makes use of basic data derived from the original video rather than entire information about the original video. This metric type is known as Reduced Reference (RR). The No-Reference (NR) metric type is the third VQA category in which the original video sequence is not used at all during the VQA process. Because the original video must be transferred to the receiver side for the VQA process, the FR and RR kinds cause problems in multimedia applications. Although there are FR metrics such as PSNR (Peak-Signal-to-Noise-Ratio), VQM (Video Quality Metric) [7] and SSIM (Structural Similarity Index) [8] that are widely used to estimate video quality, it is difficult to provide an equally recognized example of an NR metric. As a result, for the VQA, the NR metric type is more practicable and efficient [5][6]. In light of this fact, an NR metric is considered in this study.

Due to the nature of HVS, it is quite significant to determine effective factors for developing 3D NR VQA metrics.

Given these facts, the Canny, edge detection operator, which is an excellent approach for gathering edge information of objects in an image, is used in this work to provide preliminary findings utilizing color videos for establishing a NR VQA metric.

The rest of the paper is organized as follows. The edge information measurement algorithm is introduced in Section 2. In Section 3, the performance assessment results and discussions are presented. Finally, Section 4 concludes this study and points to the future work.

*Corresponding author

GÖKÇE NUR YILMAZ; Ted University, Department of Computer Engineering, Türkiye; e-mail: gokce.yilmaz@tedu.edu.tr; 0000-0002-0015-9519

YÜCEL ÇİMTAY; Ted University, Department of Computer Engineering, Türkiye; e-mail: yucel.cimtay@tedu.edu.tr; 0000-0003-2980-9228

2. Edge Information Measurement Algorithm

In this study, the methodology of edge detection for video sequences is based on the Canny edge detection algorithm, which is a mature and efficient method for calculating structural information. The edge detection optimization operator Canny is determined by the variation of pixel sizes in an image [9] [10]. The Canny algorithm has five phases that can be implemented. The first stage involves smoothing image noise with a Gaussian filter. The second step is to determine the slope force and direction of each pixel using Gradient, the first order partial derivative. A base Gradient operator can be described mathematically as in (1.1).

$$\nabla f \equiv grad(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} \equiv \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (1.1)$$

Using the numerical optimization of the one-way edge step gives the conclusion that Gauss is the first derivative, which is represented with the following mathematical equation (1.2).

$$\frac{d}{dx} e^{-\frac{x^2}{2\sigma^2}} = \frac{-x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \quad (1.2)$$

A two-dimensional circular Gaussian function softens the image so that it can determine all possible distances in all conceivable directions. Gradient is then calculated. The edges associated to the available coordinates in the image are then calculated using the maximum non-suppression to refine the smoothing image Gradient in the third stage.

Gradient amplitude and direction are then used to estimate the edge size and direction at each point.

If f(x,y) is the input image and G(x,y) is given as a Gaussian function, equation 1.3 can be written.

$$G(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1.3)$$

As a result of the convolution of G and f, a smooth image is obtained as formulated in 1.4.

$$f_s(x, y) = G(x, y) * f(x, y) \quad (1.4)$$

This process is continued by calculating the gradient amplitude (M) and direction (α) as shown in equations 1.5 and 1.6.

$$M(x, y) = \sqrt{g_x^2 + g_y^2} \quad (1.5)$$

$$\alpha(x, y) = \tan^{-1} = \begin{bmatrix} \frac{g_x}{g_y} \end{bmatrix} \quad (1.6)$$

In the equations, gx and gy are partial derivatives.

Since it is obtained using gradient, it typically contains large peaks around the local maximum. Using non-maximum suppression to refine the smoothed image Gradient, the edges in the image in possible x and y coordinates are determined.

In the fourth phase, a threshold is employed to eliminate erroneous points and to detect and connect the discrete points of the image's edges. This process can be executed with a single threshold level by making the values below the threshold level zero. Incorrect threshold level selection may result in erroneous or partial findings [9-12].

In recent years, some scholars have suggested and applied various methods based on the Canny algorithm to practical engineering. The researchers experimented with several threshold selection strategies in order to improve the Canny algorithm's successrate. Er-Sen Li et al. gain some improvement in edge detection outcomes by employing the Otsu approach in threshold selection [11][12].

This study's threshold level runs from 0 to 1. The Otsu operator and adaptive threshold determination approach presented in [12] are employed in this investigation to select the threshold level. As a consequence of the graphic work completed, it is considered that the threshold level compatible with the HVS should be 0.08.

Finally, the Canny algorithm and threshold level combination method suppress the pixels that are not regarded edges. In other words, these pixels have a pixel value of 0. The pixels judged to be edges are given the value 1. The total number of ones is then normalized with the spatial resolution in this study, which is known as Normalized Numbers of Ones (NNO).

3. Results and Discussions

In order to derive results, the color sequences of 10 different color+depth map based 3D video test sequences (i.e, Windmill, Interview, Breakdance, Chess, Advertisement, Ballet, Butterfly, Farm, Football and Newspaper) are coded using the Joint Scalable Video Model (JSVM) codec version 9.13.1 with 5 different Quantization Parameters (QPs) (i.e., 25, 30, 35, 40, 45) at 25 frame rate.

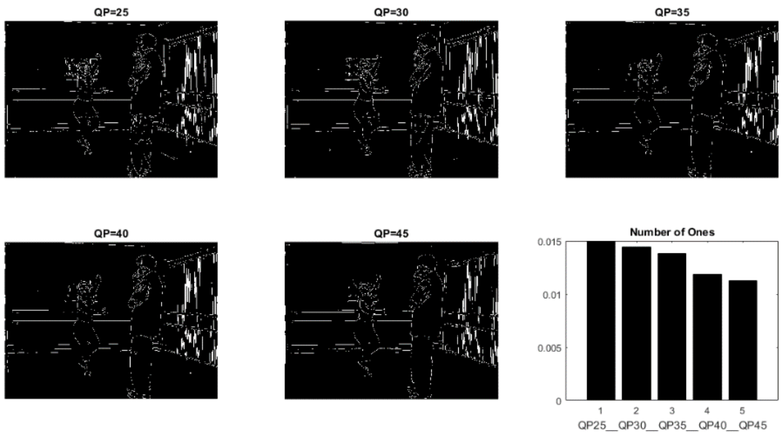

Fig. 1. The edge information of a snapshot of the "Ballet" video sequence and its associated NNOs for the 5 QPs

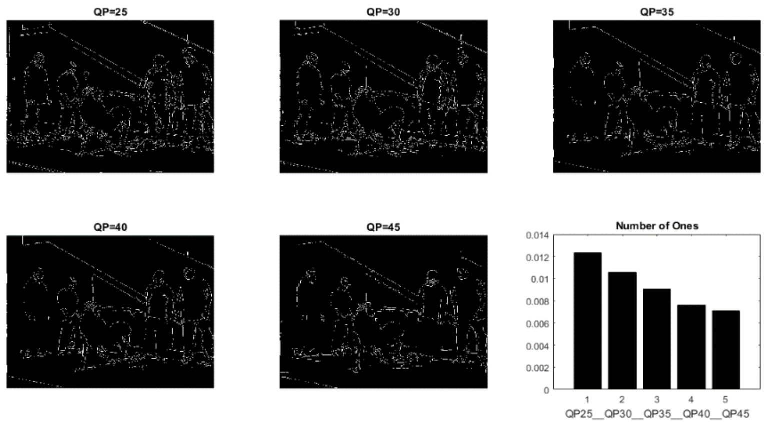

Fig. 2. The edge information of a snapshot of the "Breakdance" video sequence and its associated NNOs for the 5 QPs

Then, the edge information is obtained using the Canny algorithm and the NNOs are also calculated for the same encoded 7 video sequences and 5 different QPs. Figs. 1 and 2 show how the edge information and the NNOs differ for the Ballet and Breakdance video sequences encoded with 5 different QPs, respectively. These results are used as representative results for presenting how the edge information and the NNOs differ for the video sequences. As observed from the figures, the NNOs decrease as the video quality decreases. Similar results are obtained for the remaining 5 color video sequences.

4. Conclusions

In this study, the edge detection features envisaged as significant for the HVS, have been utilized to derive results. These results can be considered as preliminary results for developing a NR 3D VQA metric. These preliminary results present that the edge information is quite promising to be used in the development of a 3D VQA metric. By this study, it will be easier and more practical to measure 3D video quality on the receiving side, which is a critical demand for multimedia technologies. In future studies, it is aimed to evaluate other edge detection methods and develop an efficient NR VQA metric using these preliminary results.

Acknowledgement

This study has been supported by the TUBITAK 1001 project no: 122E015, namely: “An Ambient Light, Spatial Resolution, and Depth Cue Based 3 Dimensional Adaptation Model Enabling Enhanced Depth Perception Satisfaction and Efficient Communication Network Usage”.

Author Contributions

Author 1: Methodology, Results and Discussions, Software. Author 2: Review and Editing, Conclusions.

References

[1] G. Nur, H. Kodikara Arachchi, S. Dogan, and A.M. Kondoz, “Seamless Video Access for Mobile Devices by Content-Aware Utility-Based Adaptation,” Springer Multimedia Tools and Applications, vol. 70, pp. 689-719, 2014.

[2] D.V.S.X. De Silva, G. Nur, E. Ekmekcioglu, and A. Kondoz, “QoE of 3D Media Delivery Systems,” Media Networks: Architectures, Applications, and Standards, CRC Press Taylor and Francis Group, 14 May 2012.

[3] G. Nur, V. De Silva, H. Kodikara Arachchi, A. Kondoz, W. Fernando, M.O. Martínez-Rach and S. Dogan, “Sensitivity of the HVS for Binocular Disparity Cue in 3D Displays under Different Ambient Illumination Conditions,” IEEE International Conference on Consumer Electronics, Las Vegas, Nevada, USA, 13-16 January 2012.

[4] E. Cemiloglu and G. Nur Yilmaz, “Blind Video Quality Assessment via Spatiotemporal Statistical Analysis of Adaptive Cube Size 3D-DCT Coefficients,” IET Image Processing, 2020.

[5] Kai Zhao, Kun Yuan, Ming Sun, Xing Wen; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 1302-1310, 2023.

[6] G. Nur Yilmaz, “Context and Content Based Scalable Video Adaptation for Ubiquitous Multimedia Communication Services,” Springer Multimedia Tools and Applications, 2021.

[7] Pinson M.H. and Wolf S., “A New Standardized Method for Objectively Measuring Video Quality,” IEEE Trans. Broadcasting, vol. 50, no. 3, pp. 312-322, Sep. 2004.

[8] Wang Z., Lu L., and Bovik A. C., “Video Quality Assessment Based on Structural Distortion Measurement,” Proc. of Signal Processing: Image Com., vol. 19, no. 2, pp. 121-132, Feb. 2004.

[9] E S Li, etc., “An Adaptive Edge-detection Method Based on the Canny Operator,” International Conference on Environmental Science and Information Application Technology, pp.465–469, 2009.

[10] Yu Chai, Sujing Wei, Xinchun Li , "The Multi-Scale Hough Transform Lane Detection Method Based on the Algorithm of Otsu and Canny," Advanced Materials Research, vol. 1042, pp. 126-130. 001: I .4028 /www.scientific.netlAMR.I042.126, 2014.

[11] Kaur, Er Amandeep, and Er Vikas Wasson. "A novel approach to no-reference image quality assessment using canny magnitude based upon neural network." Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2017 2nd IEEE International Conference on. IEEE, 2017.

[12] Gong, Y., Li, X., Zhang, H., Liu, Q., & Sun, Y., “An improved Canny algorithm based on adaptive 2D-Otsu and Newton iterative”, IEEE 2nd International Conference on Image, Vision and Computing (ICIVC), 2017.