OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

SURVEY ON PANORAMIC VIRTUAL REALITY VIDEO QUALITY ASSESSMENT BASED ON 3D CONVOLUTIONAL NEURAL NETWORKS

Miss. Trupti Farande[1], Prof. Mangal Kotkar[2]
PG Students[1], Assistant Professor[2], Department of Computer Engineering, Dhole Patil College of Engineering, Wagholi, Near Eon IT park., Pune-412207, Maharashtra, India [1 2]
farandetrupti@gmail.com[1]

Abstract: Virtual reality (VR), a new type of simulation and interaction technology, has aroused widespread attention and research interest. It is necessary to evaluate the virtual reality quality and provide a standard for the rapidly developing technology. To the best of our knowledge, few researchers have built benchmark databases and designed related algorithms, which has hindered the further development of VR technology. In this work, a free available dataset (VRQ-TJU) for virtual reality quality assessment is presented with subjective scores for each sample data. The validity for the designed database has been proved based on the traditional multimedia quality assessment metrics. In addition, an end-to-end 3D convolutional neural networks (CNN) is introduced to predict the VR video quality without a referenced VR video. This method can extract spatio-temporal features and does not require using hand-crafted features. At the same time, a new score fusion strategy is designed, based on the characteristics of the VR video projection process. Taking the pre-processed VR video patches as input, the network captures local spatiotemporal features and gets the score of every patch. Then the new quality score fusion strategy is applied to get the final score. Such approach shows advanced performance on this database.

Keywords: Benchmark Database, Quality Score Fusion Strategy, Spatio-temporal Features, Virtual Reality Quality Assessment, 3D Convolutional Neural Networks.

----------∴∴∴----------

I INTRODUCTION

Now a days, the increase of the variety of multimedia, human beings have more access to receive visual information. As a new simulation and interaction technology, virtual reality (VR) technology is used in many fields such as architecture, military affairs and game. It can create a virtual environment which is consistent with the real world rules, or build a complete hypothetical environment which is contrary to reality. At present, the implementation of VR technology is very challenging. On the one hand, VR requires more complex implementation conditions. People must be equipped with specific devices to feel the immersion of VR. The related equipment and application scenarios restrict the further development of VR. On the other hand, VR requires a variety of perceptual information to match each other to achieve a good quality of experience, and the content of VR is different from the traditional media.

Therefore, it is necessary to evaluate the quality of all aspects of VR to promote the more standardized development of the industry. VR video, also known as panoramic stereoscopic video, is a video work played by virtual reality output device. Its purpose is to bring immersive experience with on-the-spot interaction for users to watch videos. Good visual information can bring immersion in the virtual scene, while low quality visual information not only brings bad experience, but also can lead to physical disease. As the carrier of VR visual information, VR video requires people to design the appropriate method for virtual reality video quality assessment (VRVQA).

II LITERATURE SURVEY

Discrete Event Simulation and Virtual Reality Use in Industry: New Opportunities and Future Trends

This paper reviews the area of combined discrete event simulation (DES) and virtual reality (VR) use within industry. While establishing a state of the art for progress in this area, this paper makes the case for VR DES as the vehicle of choice for complex data analysis through interactive simulation models, highlighting both its advantages and current limitations. This paper reviews active research topics such as VR and DES real-time integration, communication protocols, system design considerations, model validation, and applications of VR and DES. While summarizing future research directions for this technology combination, the case is made for smart factory adoption of VR DES as a new platform for scenario testing and decision

making. It is put that in order for VR DES to fully meet the visualization requirements of both Industry 4.0 and Industrial Internet visions of digital manufacturing, further research is required in the areas of lower latency image processing, DES delivery as a service, gesture recognition for VR DES interaction, and linkage of DES to real-time data streams and Big Data sets.

IVQAD 2017: An Immersive Video Quality Assessment Database

This paper presents a new database, Immersive Video Quality Assessment Database 2017 (IVQAD 2017), intended for immersive video quality assessment in virtual reality environment. Video quality assessment (VQA) plays an important role in video research fields. Nowadays virtual reality technology have been widely used and playing videos in virtual reality visual system is becoming more and more popular. However, existing research in VQA fields mainly focus on traditional videos. In this paper, authors build the IVQAD which contains 10 raw videos and 150 distorted videos. Bit rate, frame rate and resolution were considered as quality degradation factors. All the videos were encoded with MPEG-4. Subjects were asked to assess the video under virtual reality environment and mean opinion score (MOS) was derived by computing. Using IVQAD 2017, researchers can explore the influence of resolution, video compression and video packet loss on immersive videos' quality.

Adaptive 360 VR video streaming: Divide and conquer!

In this paper, we propose an adaptive bandwidth-efficient 360 VR video streaming system using a divide and conquer approach. Authors propose a dynamic view-aware adaptation technique to tackle the huge bandwidth demands of 360 VR video streaming. They spatially divide the videos into multiple tiles while encoding and packaging, use MPEG-DASH SRD to describe the spatial Relationship of tiles in the 360-degree space, and prioritize the tiles in the Field of View (FoV). In order to describe such tiled representations, authors extend MPEG-DASH SRD to the 3D space of 360 VR videos. They spatially partition the underlying 3D mesh, and construct an efficient 3D geometry mesh called hexaface sphere to optimally represent a tiled 360 VR video in the 3D space. The initial evaluation results report up to 72% bandwidth savings on 360 VR video streaming with minor negative quality impacts compared to the baseline scenario when no adaptations is applied.

Estimation of optimal encoding ladders for tiled 360 vr video in adaptive streaming systems

In this context, this paper targets both the provider's and client's perspectives and introduces a new content-aware encoding ladder estimation method for tiled 360◦ VR video in adaptive streaming systems. The proposed method of this paper first categories a given 360◦ video using its features of encoding complexity and estimates the visual distortion and resource cost of each bitrate level based on the proposed distortion and resource cost models. An optimal encoding ladder is then formed using the proposed integer linear programming (ILP) algorithm by considering practical constraints. Experimental results of the proposed method in this paper are compared with the recommended encoding ladders of professional streaming service providers. Evaluations show that the proposed encoding ladders deliver better results compared to the recommended encoding ladders in terms of objective quality for 360◦ video, providing optimal encoding ladders using a set of service provider's constraint parameters.

Stereoscopic images quality assessment based on deep learning

In this paper, authors proposed a S3D image quality assessment (S3D IQA) method based on deep learning. In this method, the Convolutional Restricted Boltzmann Machines (CRBM) combined with Factored Third-Order RBM (FTO-RBM) is considered as learning model to extract feature maps from pre-processed left and right images automatically. Then an improved traversal algorithm based on two pooling strategies is put forward to optimize extracted feature maps, which improves the final quality assessment performance significantly. Experimental results show that this S3D IQA method achieves good performance on 3D databases tested.

III SYSTEM ARCHITECTURE

In this work, we try to find a VRVQA method that takes full account of VR video Characteristics. In recent years, the method of deep learning has been widely used in the field of multimedia quality assessment. In addition to the most common convolutional neural networks (CNN) models, many other models have implications for their use, such as DNN-based methods, methods based on Convolutional Restricted Boltzmann Machines, methods based on generating confrontation networks. Considering the factors that affect the quality of VR video are more complex, we decide to use the deep learning model for quality assessment rather than manually extracting features. In fact, VRVQA needs to consider 2D video quality, depth perception, visual comfort, illusion of immersion and other factors. In order to fully consider the VR video's information on the time domain, we decide to design a 3D CNN architecture to capture the spatiotemporal features. We present a freely available dataset (VRQ-TJU) for VRVQA. A large number of experiments show that our method has achieved good results.

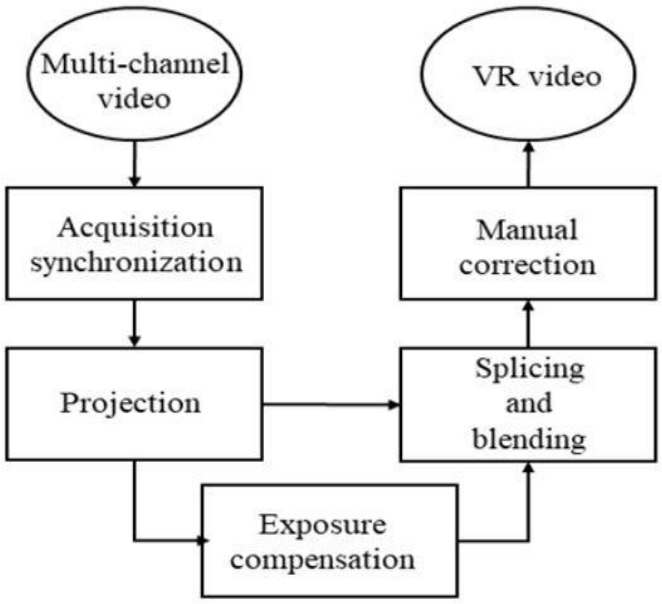

Figure 1: System Architecture

Working of System:

1) We present an end-to-end 3D CNN based framework for VRVQA, which takes VR difference video patches as input and considers the information among different frames. This is a NR VRVQA method. We can utilize the 3D CNN architecture for quality assessment without the sophisticated preprocessing. To the best of our knowledge, we are the pioneers to exploit the 3D CNN to evaluate the quality of VR video.

2) We present a dataset, VRQ-TJU, for the virtual reality quality assessment. If there is no reasonable database as a support, any evaluation algorithm will be meaningless. In essence, the establishment of the database will promote the development of virtual reality evaluation.

3) We present a quality score fusion strategy for VR plane videos. Different from the spherical model, the spatial distribution of VR videos are uneven in the plane model. This characteristic is due to the specific shooting process and projection process of VR video. Experiments show that the presented score fusion strategy can effectively improve the performance of quality assessment.

VI CONCLUSION

In this paper, a NR quality assessment method for VR video was presented. This work will help VR technology develop more mature. Before this, no one has assessed the quality of VR video and used 3D CNN to the field of quality assessment. We present a method of using 3D CNN to assess the quality of VR video. Unlike traditional NR-VQA methods, this method does not require complex preprocessing or hand-crafted features. The objective prediction score of the VR video is obtained by the combination of the local spatiotemporal features and the quality score fusion strategy. Experiments show that the results of the algorithm are consistent with the subjective quality assessment.

REFERENCES

[1] C. J. Turner, W. Hutabarat, J. Oyekan, And A. Tiwari, “Discrete Event Simulation And Virtual Reality Use In Industry: New Opportunities And Future Trends”, IEEE Transactions On Human-Machine Systems, Vol. Pp, No. 99, Pp. 113, 2016.

[2] H. Duan, G. Zhai, X. Yang, D. Li, And W. Zhu, “Ivqad 2017: An Immersive Video Quality Assessment Database”, In International Conference On Systems, Signals And Image Processing, 2017, Pp. 15.

[3] M. Hosseini And V. Swaminathan, “Adaptive 360 Vr Video Streaming: Divide And Conquer!”, Pp. 107110, 2016.

[4] C. Ozcinar, A. D. Abreu, S. Knoor, A. Smolic, C. Ozcinar, A. D. Abreu, S. Knoor, A. Smolic, C. Ozcinar, And A. D. Abreu, “Estimation Of Optimal Encoding Ladders For Tiled 360 Vr Video In Adaptive Streaming Systems”, In The Ieee International Symposium On Multimedia, 2017.

[5] K.Wang, J. Zhou, N. Liu, And X. Gu, “Stereoscopic Images Quality Assessment Based On Deep Learning”, In Visual Communications And Image Processing, 2016, Pp. 14.

[6] F. Gao, J. Yu, S. Zhu, Q. Huang, And Q. Tian, “Blind Image Quality Prediction By Exploiting Multi-Level Deep Representations”, Pattern Recognition, 2018.

[7] F. Gao And J. Yu, “Biologically Inspired Image Quality Assessment”, Elsevier North-Holland, Inc., 2016.