Do We Need a New Large-Scale Quality Assessment Database for Generative Inpainting Based 3D View Synthesis? (Student Abstract)

Sadbhawna, Vinit Jakhetiya, Badri N. Subudhi, Harshit Shakya, Deebha Mumtaz

Indian Institute of Technology Jammu, India

{2018rcs0013, vinit.jakhetiya, subudhi.badri, 2020pcs2064, 2018rcs0011}@iitjammu.ac.in

Abstract

The advancement in Image-to-Image translation techniques using generative Deep Learning-based approaches has shown promising results for the challenging task of inpainting-based 3D view synthesis. At the same time, even the current 3D view synthesis methods often create distorted structures or blurry textures inconsistent with surrounding areas. We analyzed the recently proposed algorithms for inpainting-based 3D view synthesis and observed that these algorithms no longer produce stretching and black holes. However, the existing databases such as IETR, IVC, and IVY have 3D-generated views with these artifacts. This observation suggests that the existing 3D view synthesis quality assessment algorithms can not judge the quality of most recent 3D synthesized views. With this view, through this abstract, we analyze the need for a new large-scale database and a new perceptual quality metric oriented for 3D views using a test dataset.

Introduction

A good 3D-synthesized images/videos can provide consumers with a more engaging and better immersive experience. Free Viewpoint Video (FVV), 3D-Television, 360°video, Virtual Reality (VR) are some of the applications of 3D-synthesis, famous because of their realistic and interactive experience (Shih et al. 2020; Niklaus et al. 2019). Unfortunately, the rendered 3D views, even using the contemporary methods, cannot generate the perfect novel 3D view (Shih et al. 2020). These methods cannot perform efficiently on complex surfaces and produce some artifacts, as shown in Fig. 1. The artifacts in the 3D synthesized views are different from the conventional artifacts in regular natural images. With the advancement of efficient algorithms for generating 3D synthesized views, it is required to have an image quality assessment (IQA) algorithm which can automatically judge the perceptual quality of generated 3D synthesized views that match with the human visual system. The 3D IQA algorithms can judge the perceptual quality and are also helpful in the fast development of Image Restoration (IR) (IR includes tasks such as super-resolution (SR), denoising, enhancement, etc.) algorithms. With this view, through this abstract, we will address the research problem statement that is there a need for creating a large-scale 3D synthesized IQA database which is generated using the recently proposed 3D synthesized view generation algorithms?

Figure 1: (a). A synthesized view. (b). The failure (green arrows) of a random patch (red window) in a 3D synthesized view. Synthesized Using: (Shih et al. 2020)

Related Work

Novel View Synthesis(NVS): From proxy geometry based traditional methods to Generative Adversarial Networks (GANs) based deep learning methods, NVS has evolved with the advancement in computer vision techniques. (Shih et al. 2020; Niklaus et al. 2019) are some of the new view synthesis techniques. Synthesizing novel views requires stipulations such as comprehensive scene understandings, preserving structures observed in the input semantics, lowered baseline requirements, etc. Even the newest view synthesis methods in the literature lack in one or the other of these stipulations. Hence, much research is being done to make the view synthesis perfect. Figure 1 shows an example of the failure of such a contemporary 3D synthesized method. From this Figure, it is clear that view synthesis does not render a clear view in certain circumstances, such as complex surfaces.

Image Quality Assessment(IQA): SSIM(Wang et al. 2004) is the most widely used IQA method as it introduced the structural similarity in comparing images as compared to the Mean Square Error (MSE) value, which are FR (Full-Reference) IQAs. There are various other popular NR (No-Reference) IQAs, such as BRISQUE, NIQE, etc., in the literature. There are various new IQAs based on deep features

Dataset	IVC	IVY	IETR
Dataset proposed in year	2011	2016	2019
Number of syn. Views	84	84	140
3D syn. algo. used	4	7	7
Year of most recent 3D algo.	2010	2014	2016
Obsolete distotions?	Yes	Yes	Yes
Type of obsolete distortion	BH	Ghos.	Stre.
Gen. methods included?	No	No	No

syn: synthesized, algo: algorithm, BH: black-holes, Ghos: Ghosting, Stre: Stretching, Gen: Generative

Table 1: Comparison of existing IQA datasets.

Metric	Oriented for	PLCC	SRCC	RMSE
LOGS	3D Views	0.6350	0.6021	0.8400
PSNR	Natural Images	0.2869	0.1772	1.0417
SSIM	Natural Images	0.1610	0.1231	1.1735
LPIPS	Natural Images	0.1921	0.0132	1.3874
APT	3D Views	0.1717	0.0013	1.3849

PLCC: Pearson Linear Correlation Coefficient, SRCC: Spearman Rank Correlation Coefficient, RMSE: Root Mean Square Error

Table 2: Performance of state-of-the-art IQA algorithms for the proposed test dataset.

such as LPIPS Metric (Zhang et al. 2018). BAPPS (Zhang et al. 2018) and PIPAL(Gu et al. 2020) are two examples of large-scale contemporary IQA datasets. All these IQA are fundamentally designed and oriented for Natural Images.

In 2012, Bosc et al. (Bosc et al. 2011) first analyzed that the IQAs developed for natural images work poorly for 3D synthesized images and proposed a dataset named IRC-CyN/IVC dataset for this analysis of 84 images. After that, there are many different 3D-IQAs and datasets proposed for this task in the literature, such as IVY, IETR (Tian et al. 2021). A summarized comparison of these three datasets is given in Table 1 with their drawbacks.

Motivation

Out of various existing 3D IQA and synthesis algorithms, we have encountered the following drawbacks:

1. To the best of our knowledge, there is no large-scale dataset for quality evaluation of 3D synthesized views, and subsequently, no generic IQA algorithms are proposed to judge the quality of 3D synthesized views.
2. The quality evaluation datasets and metrics used by contemporary 3D synthesis methods for various purposes are designed for natural images (for example, to determine the threshold in Section 3.1 in the paper (Shih et al. 2020), authors used the LPIPS metric (Zhang et al. 2018) which is initially designed for naturally degraded images and not for 3D images).

The process of the creation of the IQA dataset and its subjective testing is hectic. To this context, in order to validate that the proposed problem is worth pursuing, we created a small test dataset of 60 3D views generated using two recent 3D algorithms (i.e., (Shih et al. 2020; Niklaus et al. 2019)), tested using five expert subjects. This subjective testing is also validated using Cohen's Kappa coefficient. The performance of five popular IQA metrics (LOGS (Tian et al. 2021), Peak Signal to Noise Ratio(PSNR), SSIM (Wang et al. 2004), LPIPS (Zhang et al. 2018), APT (Tian et al. 2021)) oriented for natural as well as 3D images are given in Table 2. The comparison in Table 1, the value of correlation coefficients, and the error in Table 2 suggest that the literature has no proper algorithm for this purpose.

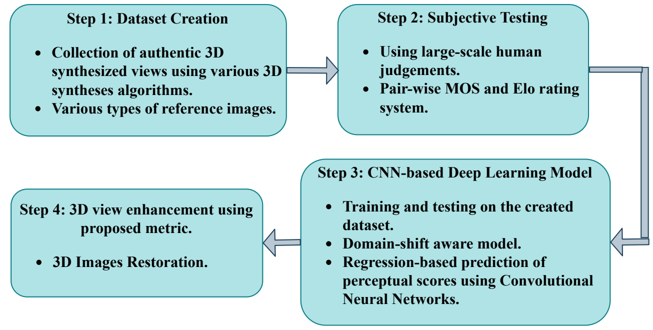

Figure 2: Step-wise flow of the proposed future work.

Conclusions and Future Work

Our preliminary analysis suggests a need for a new perceptual metric designed explicitly for 3D views for 3D image restoration and enhancement. For this purpose, the future steps involved in the proposed process is summarized in Figure 2.

References

Bosc, E.; Pepion, R.; Le Callet, P.; Koppel, M.; Ndjiki-Nya, P.; Pressigout, M.; and Morin, L. 2011. Towards a New Quality Metric for 3-D Synthesized View Assessment. *IEEE Journal of Selected Topics in Signal Processing*, 5(7): 1332–1343.

Gu, J.; Cai, H.; Chen, H.; Ye, X.; Ren, J.; and Dong, C. 2020. PIPAL: a Large-Scale Image Quality Assessment Dataset for Perceptual image Restoration. In *ECCV*.

Niklaus, S.; Mai, L.; Yang, J.; and Liu, F. 2019. 3D Ken Burns Effect from a Single Image. *ACM Transactions on Graphics*.

Shih, M.-L.; Su, S.-Y.; Kopf, J.; and Huang, J.-B. 2020. 3D Photography using Context-aware Layered Depth Inpainting. In *CVPR*.

Tian, S.; Zhang, L.; Zou, W.; Li, X.; Su, T.; Morin, L.; and Deforges, O. 2021. Quality Assessment of DIBR-synthesized views: An Overview. *Neurocomputing*.

Wang, Z.; Bovik, A.; Sheikh, H.; and Simoncelli, E. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612.

Zhang, R.; Isola, P.; Efros, A. A.; Shechtman, E.; and Wang, O. 2018. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In *CVPR*.