This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2888861, IEEE Transactions on Industrial Informatics

A Highly Efficient Blind Image Quality Assessment Metric of 3D-Synthesized Images using Outlier Detection

Vinit Jakhetiya, Ke Gu, Trisha Singhal, Sharath C. Guntuku, Zhifang Xia, and Weisi Lin, *Fellow, IEEE*

***Abstract*—With multitudes of image processing applications, Image Quality Assessment (IQA) has become a prerequisite for obtaining maximally distinctive statistics from images. Despite the widespread research in this domain over several years, existing IQA algorithms have a number of key limitations concerning different image distortion types and algorithms' computational efficiency. Images which are synthesized using depth image-based rendering have applications in various disciplines such as free viewpoint videos, which enable synthesis of novel realistic images in the referenceless environment. In the literature, very few no-reference quality assessment metrics of 3D synthesized images are proposed, and most of them are computationally expensive which makes it difficult for them to be deployed in real-time applications. In this work, we attribute the geometrically distorted pixels as outliers in 3D synthesized images. This assumption is validated using the 3 $sigma$ rule-based robust outlyingness ratio. We propose a novel fast and accurate blind IQA metric of 3D-synthesized images using non-linear median filtering since the median filtering has the capability of identifying and removing outliers. The advantages of the proposed algorithm are two-fold. First, it uses a simple technique, i.e. median filtering, to capture the level of geometric and structural distortions (up to some extend). Second, the proposed algorithm has higher computational efficiency. Experiments show the superiority of the proposed no-reference IQA algorithm over existing state-of-the-art full-, reduced-, and no-reference IQA methods, in terms of both predicting accuracy and computational complexity.**

***Index Terms*—Image Quality Assessment (IQA), Median Filtering, Robust Outlyingness Ratio (ROR), Outliers**

This work was supported in part by the National Science Foundation of China under Grant 61527804, 61703009, the Beijing Advanced Innovation Center for Future Internet Technology under Grant 110000546619001, the Young Elite Scientist Sponsorship Program by China Association for Science and Technology under Grant 2017QNRC001, and the Nova Programme Interdisciplinary Cooperation Project under Grant Z161100004916041. (Corresponding author: Ke Gu.)

V. Jakhetiya is with Department of Computer Science and Engineering, Indian Institute of Technology, Jammu, India (email: vinit.jakhetiya@iitjammu.ac.in).

K. Gu is with Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing Advanced Innovation Center for Future Internet Technology, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China (email: guke.doctor@gmail.com).

T. Singhal is with Department of Computer Science and Engineering, Hindustan College of Science and Technology, India (email: trishasinghal.hcst.cs15@sgei.org).

S. C. Guntuku is with the Centre for Digital Health, University of Pennsylvania, USA (email:sharathg@sas.upenn.edu).

Z. Xia is with the State Information Center of P. R. China, Beijing 100045, China (email: xiazhifang@cegn.gov.cn)

W. Lin is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore, 639798 (email: wslin@ntu.edu.sg).

I. Introduction

As one of the rapidly growing technologies, image processing has received substantial attention owing to the fact that it has extensive applications and exerts influences on several technical fields. Image processing can be more precisely interpreted as the means of translating a digital image to the human visual system (HVS). Display devices and transmission of images impose noise in the image which tends to degrade the quality of the image. The objective of image quality assessment metrics is to predict the quality score of images which matches to that of the HVS. Even after years of persistent efforts, no IQA algorithm succeeds in the effective and efficient management of certain constraints, predominantly, different types of distortions which include structural distortions, and geometric distortions in images and their computational complexity.

With the exponential development of 3D-related technologies [2] viz. Augmented Reality (AR), Mixed Reality (MR) and Virtual Reality (VR), Free Viewpoint Television (FTV) came into limelight because of its numerous applications in areas like remote education, remote surveillance, entertainment, medical applications, etc. These technologies enable the users to navigate through the unhindered spatio-temporal space within the realistic virtual environment. In Free Viewpoint Videos (FVVs), users experience multiple views which are generated via the Depth-Image-Based-Rendering (DIBR) techniques [1]. Operational rendering technique generally generates warp and geometric distortion in views which may create deplorable artifacts in the entire FVV. The IQA metrics are not only useful for predicting the quality of DIBR images also in the image/video compression [3], [4], [29]. To resolve these obstructions, it is required to have an assessment metric to evaluate the quality score of DIBR-synthesized images with maximum computational efficiency possible.

II. Related Work

In previous years, numerous IQA metrics of natural, screen-content, re-targeted, and 3D-synthesized images have been developed. These IQA algorithms are broadly categorized into three divisions, such as Full-Reference (FR), Reduced-Reference (RR) and No-Reference (NR) class. These three classes of IQA models incorporate complete information, partial information and no information about the source image respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2888861, IEEE Transactions on Industrial Informatics

Perception of natural images by the HVS is extremely structured and can be used for extracting structural information from the images. Based on the dependencies demonstrated by the pixels of images, researchers proposed a more straightforward method to compute the differences between the reference and warped images. Various IQA methods were proposed based on structural similarity and its variants.

The major limitation of structural similarity based IQA metrics is that these metrics are not applicable for geometrically distorted images [7]. Hence, these algorithms cannot be deployed to judge the quality of 3D synthesized images.

On the same line, a lot of NR IQA algorithms for natural images were proposed in the literature keeping Natural Scene Statistics (NSS) as their fundamental foundation. For example, Gu et al. proposed an IQA algorithm for the images under multiple distortions [8] and contrast degradations [9]. Authors of [10] proposed a completely blind assessment approach known as Natural Image Quality Evaluator (NIQE), which did not involve any learning. The new model followed the feature-based approach which evaluates the model statistics, for an instance, analysis of Mean Subtracted Contrast Normalized (MSCN) coefficients to derive the quality score. Yue *et al.* proposed a blind IQA algorithm for tone-mapped images based on the responses of double-opponent (DO) and single-opponent (SO) cells [11], [12]. However, Min *et al.* [13] suggested a combined quality assessment measure, unified content-type adaptive (UCA) for compressed natural scene images (NSIs), computer graphic images (CGIs) and screen content images (SCIs) incorporating the subjective evaluation. Furthermore, Min *et al.* [14] presented distortion determined NR metric based on the pseudo reference image (PRI).

The 3D synthesized images are contaminated with both structural and geometrical distortions. The IQA algorithms proposed for natural images are seen to perform well when structural distortions are present in the distorted image. Further, these algorithms do not have the capability of capturing geometric distortions. Consequently, these algorithms have a poor performance when applied to 3D synthesized images.

The 3D synthesized images produced by rendering techniques comprise of geometric distortions affecting the visual quality. With this concern, Bosc *et al.* [15] created a database for 3D synthesized images and a few IQA algorithms were proposed for this category of images. Sandic *et al*. proposed Morphological Wavelet Decomposition (MW-PSNR) method [16] and Morphological Pyramid Decomposition (MP-PSNR) method [17] using morphological filters. The approach computed the SNR (Signal-to-noise ratio) for evaluating the perceptual quality of 3D-synthesized images. The Reduced-Reference (RR) IQA model, MP-PSNR-reduced [18], was deployed as a replacement of MW-PSNR [16]. The revised model possessed lower computational complexity when compared to MP-PSNR [17]. Jang *et al.* proposed an IQA metric for depth images which considers the occlusion effects [5]. Also, Wang *et al.* proposed a quality assessment algorithm for both multi-view and depth images [6] and applied it on the MCL-3D database [30]. Yue *et al.* proposed a new IQA algorithm for 3D synthesized images which combines both local and global measures [28]. Very recently, Gu *et al.* [7] proposed the NR IQA metric of DIBR-synthesized images by employing local image description based on autoregressive (AR) modeling. The performance of this metric is satisfactory but the computational complexity associated with this algorithm is quite high. Jakhetiya *et al.* [19] proposed a reduced-reference IQA method based on distortion categorization and local prediction. In the literature, we found very few blind IQA algorithms for 3D synthesized images and at the same time, the proposed algorithms are quite computationally expensive.

In order to overcome the above-described limitation, we propose a new IQA metric for 3D-synthesized images which uses a very simple technique, median filtering. The main contributions of the proposed algorithm are three-fold:

1) We proposed to model the behavior of geometric and structural distortions as outliers and validated this assumption using the 3 $sigma$ rule-based Robust Outlying Ratio (ROR).
2) We proposed that median filtering has the capability of highlighting both the geometric and structural (up to some extent) distortions in 3D synthesized images and subsequently utilized them to determine the image quality of these images.
3) The proposed metric exhibits much better performance than the existing IQA algorithms yet with less computational complexity.

In what follows, Section III has the comprehensive explanation of the presented approach of the novel IQA metric for 3D synthesized images including the justification for all the techniques applied and enlightenment on the drawbacks related to the existing IQA algorithms. Section IV deals with the experimental outcomes, comprising the values of correlation coefficients, error estimation, and computational efficiency. Section V states the culminating comments and future works.

III. Proposed Algorithm

In the given work, a highly efficient quality assessment approach for 3D synthesized images in the blind environment with high implementation efficiency is proposed. The foremost concern related with IQA of DIBR-synthesized images is the capturing of both geometric and structural distortions which degrade the quality of images. The previously proposed blind IQA algorithms [7] are able to highlight geometric regions accurately, and simultaneously, are able to effectively predict the quality score when geometric distortions are predominant in the 3D synthesized images. But these IQA metrics perform poorly when both geometric and structural distortions are available. Also, the algorithms are quite computationally expensive and cannot judge the quality of 3D synthesized images in real-time scenarios.

A. Pixel outliers in geometric regions

One of the main motivations of the proposed algorithm is to identify the behavior of pixels in geometric distorted regions and subsequently, devise an efficient algorithm employing an uncomplicated image processing technique. This will be helpful in two ways: reduction in the complexity

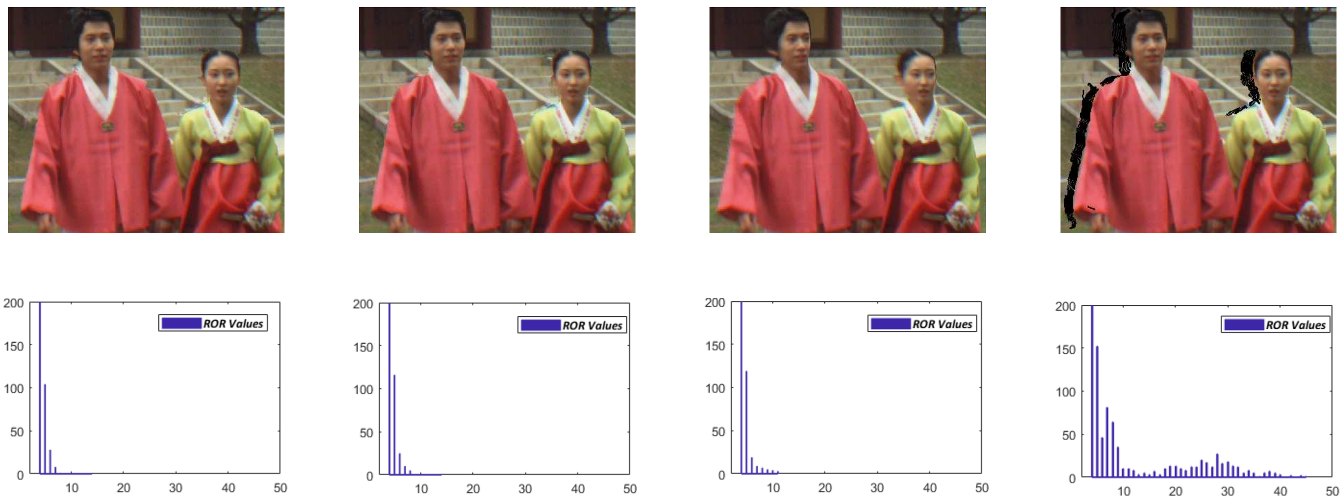


Fig. 1: Relationship between four 3D synthesized images from IRCCyN database [15] (first row) and histograms (second row) of corresponding ROR values. From left to right, images are sorted in order of decreasing perceptual quality (MOS values). These four images have MOS values 3.72, 3.70, 3.302, and 2.88 respectively and corresponding ROR values greater than four (outliers) are 142, 159, 166, and 684, respectively.

of the algorithm immeasurably and identifying the behavior of geometrically distorted pixels to give directions to the future researchers. The correlation between the pixels in the geometric distortion region is quite low and in this work, we assume these pixels to behave like outliers [19], [7]. At the same time, our visual system is quite sensitive to the outliers and it can sense even a small amount of outlier [20]. Subsequently, these outliers severely degrade the perceptual quality of 3D synthesized images. In the literature, 3 $sigma$ rule-based ROR [21] is efficiently able to identify the outliers. Keeping this in view, in order to check the behavior of the pixels of the 3D synthesized image in the geometric regions, we propose to estimate impulsiveness (outlierness) of pixels based upon the 3 $sigma$ rule. The 3 $sigma$ statistics for the i^{th} pixel [21] in the image y is defined as:

$$t_i = \frac{(y_i - \bar{y_i}) + \epsilon}{\sigma_i + \epsilon} \tag{1}$$

where $\bar{y_i}$, σ_i and ϵ are the samples mean, sample standard deviation of the pixel of interest and a very small constant number to avoid division-by-zero respectively. While t_i is the measurement of outlierness [21] of the pixel of interest. The 3 $sigma$ rule suggests that a pixel is most probably an outlier with probability (P$\geq$ 0.97) if the value of t_i is greater than 3. In the situation of a large number of outliers in an image, these outliers might interact with each other and some outliers remain unnoticed via the 3 $sigma$ rule. With this view, we propose to use 3 $sigma$ rule-based Robust Outlyingness Ratio (ROR) statistics [21] to check whether pixels in geometrical distortion regions behave like outliers or not. The ROR statistics is defined as:

$$ROR(y_{i,j}) = \frac{|y_{i,j} - Med(y_{i,j})|}{MADN(y_{i,j})} \tag{2}$$

where MADN and Med are the normalized median absolute deviations and median of the pixel of interest respectively. The MADN is defined as:

$$MADN(y_{i,j}) = \frac{|Med(y_{i,j} - Med(y_{i,j}))|}{0.6457} \tag{3}$$

The ROR value of a pixel suggests it's level of outlierness and pixels which have ROR value greater than 4 are considered to be outliers. More details of ROR statistics can be found in [21]. In order to support our assumption that pixels with geometric distortion in 3D synthesized images are behaving like the outliers, we propose to estimate ROR statistics for such pixels. In Fig.1, four cropped 3D synthesized images from the IRCCyN/IVC [15] database and corresponding distributions (histograms) of ROR values are shown. These four images are sorted in order of decreasing MOS values (perceptual quality). From this figure, the following arguments can be made:

1) Only very few pixels have ROR values greater than 4 where geometrical distortions are not predominant in an image (Figs. 1(a), (b) and (c)).
2) As for an input image contaminated with the geometrical distortion, ROR values are much greater than the 4 and subsequently, these pixels behave like outliers. So these ROR statistics can be used to estimate the level of geometric distortions.
3) With the decrement of MOS values, both geometric distortions and ROR values increase, which clearly shows a positive correlation between the ROR values and the geometric distortions.
4) It is also clear that even without the presence of strong geometrical distortions (Fig.1 (a), (b) and (c)), the number of outliers and MOS values show a correlation which suggests that the strong structural distortions are also behaving like the outliers to some extent.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2888861, IEEE Transactions on Industrial Informatics

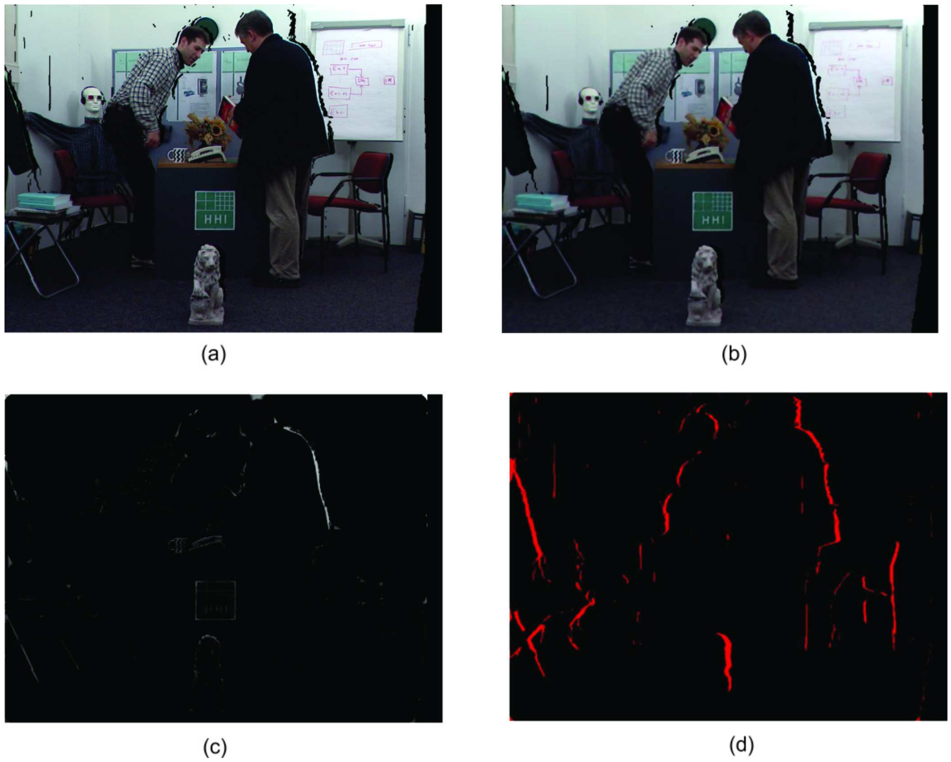


Fig. 2: Capability of median filtering in highlighting the quality degrading distortions. Pre-processing: (a) Geometrically distorted image; (b) Filtered image using Median Filtering; (c) Residual Image (between original and median filtered image) (d) Modified residual filtered image.

From the above arguments and empirical study, it is clear that pixels with geometrical and strong structural distortions behave like outliers. An efficient estimation of outliers and counting of outliers can help in finding the quality score of 3D synthesized images.

B. IQA of 3D synthesized images using median filtering

From the above empirical study, we can observe that efficient identification and removal of outliers can lead to an efficient estimation of the perceptual quality of 3D synthesized images. It is well known that the median filtering is performing much better than the non-linear filters (such as bilateral and non-local means), and autoregressive modeling for automatic identification of outliers and also for removing them. With this view, in the work, computationally efficient median filtering is used for the automatic identification and removal of outliers. With the increment of distortions (both geometric and structural), the number of outliers will increase and median filtering has the capability of removing them up to some extent.

The statistically based non-linear signal processing technology, median filtering, is basically used for removing perceptual quality deteriorating noise from the images while preserving the edges [22]. The noisy pixel is replaced with the median value of the neighboring pixels in a window. In order to show that the median filtering is capable of automatically identifying and removing the outliers to some extent, we proposed to apply median filtering on the 3D synthesized image as shown in Fig. 2. Fig. 2(c) represents the residual image between the 3D synthesized image and the corresponding median filtered image (Fig. 2(b)). In Fig. 2(d) the modified residual image is shown, where the pixels greater than a certain threshold are represented using the red color for the purpose of better visibility. In this work, our goal is to identify the geometric distortions and not to remove them. From these figures, it is clear that median filtering has the remarkable capability of highlighting both the geometric and structural distortions. With the increment of distortions, the magnitude of the residual image (R) will increase and result in the quality degradation of 3D synthesized images. The predicted value of a pixel using median filtering and residual image are as follows:

$$\hat{I}(m,n) = Median\{I(i,j)(i,j) \in w\}, \tag{4}$$

$$R(m,n) = |I(m,n) - \hat{I}(m,n)|. \tag{5}$$

In general, the geometrical distortion more affects the perceptual quality as compared to the structural distortions [19], [7]. With this view, we proposed to classify the residual image (between original and median filtered image) $R(m,n)$ into three categories, which are respectively

1) regions with overall (both structural and geometric) distortions (M_{SG}),
2) regions with strong geometric distortions (M_G), and
3) regions which do not have perceptual quality degrading geometric and structural distortions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2888861, IEEE Transactions on Industrial Informatics

$$M_{SG}(m,n) = \begin{cases} R(m,n), & \text{if } R(m,n) > \Upsilon_1 \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

$$M_{G}(m,n) = \begin{cases} M_{SG}(m,n), & \text{if } M_{SG}(m,n) > \Upsilon_2 \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where Υ_1, and Υ_2 are the two thresholds used to distinguish between overall distortion and strong geometric distortions and Υ_2 should be greater than Υ_1. The idea behind Equations (6) and (7) is for estimating the over-all distortion in the 3D synthesized image and further identifying the severe geometric distortions which predominantly affects the perceptual quality of these images. If the variation in the map M_G is high, it clearly indicates the corresponding 3D synthesized image has geometric distortions and the perceptual quality is really poor. At the same time, if both maps M_G and M_{SG} are smooth, it clearly indicates that the 3D synthesized image is free from the geometric and structural distortions, and this image has better perceptual quality.

The previous NR IQA algorithm, [7] designed for 3D synthesized images is classifying these images into only two categories such as geometric and non-geometric distortion category. With this view, this algorithm is able to perform well when only geometric distortion is available in the 3D synthesized image and when both structural and geometric distortions are available, the performance of this algorithm is not satisfactory. Also, this algorithm uses binary maps to estimate the perceptual quality and subsequently, ignoring the impact of level of distortions. As standard deviation has the better capability of highlighting variation in signals as compared to mean values. With this view, We also propose to use the standard deviation of both M_G and M_{SG} maps to estimate the level of geometrical and structural distortions in 3D synthesized images.

$$B_G = Std(M_G); \quad B_{SG} = Std(M_{SG}). \tag{8}$$

Based upon B_G and B_{SG}, it is possible to have two conditions

1) When both B_G and B_{SG} have high values and are very close to each other, which clearly suggests both geometric and structural distortions are available in the 3D synthesized image and perceptual quality is quite poor.
2) When B_{SG} has quite high value as compared to B_G, it suggests that the structural distortion is dominating over the geometrical distortion and perceptual quality of 3D synthesized image is moderately better.

We proposed to compare the maps of overall distortion and the geometric distortion for judging the perceptual quality of 3D synthesized images. Finally, we proposed to pool variations (B_G and B_{SG}) of the maps (M_G and M_{SG}) as follows to estimate the final quality score estimated as:

$$Q_s = \frac{2 * B_G * B_{SG} + \epsilon}{B_G^2 + B_{SG}^2 + \epsilon} \tag{9}$$

where ϵ is a constant number having very small value for eluding divide-by-zero complication. The quality score for 3D-synthesized images can be promptly estimated by using two levels of geometric distortions as the influence of input image contents and eliminating them from comprehensive geometric distortions. Median filtering and dual thresholding were the principal strategies applied. From the equation (9), it can be readily discerned that our proposed metric has no knowledge regarding the reference image, hence the proposed metric is totally blind (referenceless). Also, the proposed algorithm requires only one median filtering and two thresholds to estimate the quality score, which clearly suggests that the proposed algorithm is computationally simple.

IV. Experimental Results

In the work, a very potent and computationally efficient IQA algorithm for the 3D-synthesized images is proposed. Thus, to examine the performance of the proposed metric, the algorithm is applied to the IRCCyN/IVC database [15]. The IRCCyN database contains a total of 96 images in which 84 images are synthetically created (virtual) views and 12 images are pristine images in nature. These 84 images are synthetically created using several 3D synthesis algorithms and more information about these synthesis algorithms can be found in [15]. We have compared it with the existing several IQA algorithms to compare the results developed for both, natural and 3D-synthesized images.

We compared the proposed algorithm with the several previously given algorithms in respect of Pearson linear correlation coefficient (PLCC), Spearman's rank correlation coefficient (SRCC), and root mean square error (RMSE). PLCC and SRCC retained higher values while a lower value is computed for RMSE which represents a better IQA algorithm. The objectively predicted outcomes are plotted to the subjective human ratings using a five factored nonlinear logistic function, such as:

$$f(Q_s) = \beta_1 \left(0.5 - \frac{1}{1 + e^{\beta_2 (Q_s - \beta_3)}}\right) + \beta_4 Q_s + \beta_5 \tag{10}$$

where $f(Q_s)$ and Q_s represent the mapped scores and forecasted scores using the IQA algorithms, respectively. $\beta_i(i{=}1,2,3,4,5)$ are the five free arguments to be found using the regression technique. For evaluating the outcomes, Table I presents the comparison results of the proposed algorithm with prevailing IQA metrics. Observations depict that the given work acquired preferable values of PLCC, SRCC, and RMSE than the contending metrics as 0.7678, 0.7036 and 0.4266. The following analysis can be procured from the table:

1) Existing assessment metrics of natural images are not successful in a convincing quality evaluation. In the first category, Most Apparent Distortion (MAD) [23] emerged as the prime metric with 0.6077, 0.5994 and 0.5288 of PLCC, SRCC and RMSE respectively. Other metrics in the first category are PSNR, and SSIM. Among the RR metrics (Fourier Transform based Quality Measure (FTQM) [24] and Reduced-Reference Entropic Differencing (RRED) [25]) and NR metrics (NIQE [10],

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2888861, IEEE Transactions on Industrial Informatics

TABLE I: Performance comparison of the proposed algorithm with previous FR, RR, and NR quality assessment metrics in terms of PLCC, SRCC and RMSE.

Metric	PLCC	SRCC	RMSE	Designed for	Category
PSNR	0.3976	0.3095	0.6109	Natural Images	Full-Reference
SSIM	0.485	0.4368	0.5823	Natural Images	Full-Reference
MAD [23]	**0.6077**	**0.5994**	**0.5288**	**Natural Images**	**Full-Reference**
FTQM [24]	**0.5628**	**0.5543**	**0.5504**	**Natural Images**	**Reduced-Reference**
RRED [25]	0.4072	0.3090	0.608d1	Natural Images	Reduced-Reference
NIQE [10]	0.4374	0.3739	0.5987	Natural Images	No-Reference
SISBLIM [8]	**0.5225**	**0.3832**	**0.5677**	**Natural Images**	**No-Reference**
MW-PSNR [16]	0.5622	0.5757	0.5506	DIBR-Synthesized Images	Full-Reference
MP-PSNR [17]	0.6174	0.6227	0.5238	DIBR-Synthesized Images	Full-Reference
MP-PSNR-reduce [18]	**0.6772**	**0.6634**	**0.4899**	**DIBR-Synthesized Images**	**Reduced-Reference**
3D-SWIM [26]	0.6584	0.6156	0.5011	DIBR-Synthesized Images	Full-Reference
VSQA [27]	0.5742	0.5233	0.5451	DIBR-Synthesized Images	Full-Reference
RRLP [19]	**0.7145**	**0.6293**	**0.4659**	**DIBR-Synthesized/Screen Content Images**	**Reduced-Reference**
APT [7]	**0.7307**	**0.7157**	**0.4546**	**DIBR-Synthesized Images**	**No-Reference**
NIQSV+ [31]	**0.654**	**0.490**	**0.504**	**DIBR-Synthesized Images**	**No-Reference**
CLGM [28]	**0.675**	**0.652**	**0.462**	**DIBR-Synthesized Images**	**No-Reference**
Proposed Algo.	**0.7678**	**0.7036**	**0.4266**	**DIBR-Synthesized Images**	**No-Reference**

and SIx-Step BLInd Metric (SISBLIM) [8]). FTQM [24] and SISBLIM [8] approaches exhibit the optimum value of PLCC as 0.5628 and 0.5225 respectively. These testing results validate the incapability of these metrics in seizing the geometric distortions in DIBR-synthesized views.

2) In the second category, the IQA metrics are listed for DIBR-synthesized views achieved better implementation results than those for natural images. These methods include MW-PSNR [16], MP-PSNR [17], MP-PSNR-reduc [18], 3D-SWIM [26], VSQA [27], APT [7], NIQSV+ [31], CLGM [28] and RRLP [19]. The NR metric APT [7] has the supreme performance among this category and this IQA model achieves 0.7307 and 0.7157 values of PLCC and SRCC respectively. These metrics have lower values than our NR IQA metric acquired higher values of PLCC and SRCC. In addition to this, the prevailing FR, RR and NR models also have very high computational complexity, which is one of the important issues allocated with image quality assessment.

Comprehensively, the proposed referenceless IQA metric achieved higher values of PLCC and SRCC with the lesser values of RMSE by comparison with the considered quality assessment metrics. The proposed method also has very high computational efficiency.

In Table II, the performance of the proposed algorithm for different reconstruction algorithms used in the IRCCyN database is shown to analyze the performance of the proposed algorithm for individual structural and geometric distortions. In the IRCCyN database created virtual views from algorithms A3, A4, A5, A7 are contaminated with the geometric distortions. While created virtual views from algorithms A1, A2, and A6 are predominantly affected by the structural distortions. From Table II, we noticed that the APT algorithm is able to perform well when virtual views are affected with the geometrical distortions. While for the virtual views which are predominantly contaminated with the structural distortions, the performance of the APT algorithm is poor, as shown in the 2^{nd}, 3^{rd} and 7^{th} row. Also, IQA algorithms developed for natural images are able to perform well for the A2 and A6 category. While these algorithms are performing poorly when geometric distortions are present in the virtual view. At the same time, the proposed algorithm is able to perform well for virtual views which are affected by both the geometric and structural distortions. Keeping this in view, the proposed algorithm is more suitable to compute the perceptual quality of the 3D synthesized images in contrast with the prevailing IQA metrics.

In Fig. 3, the scatter plot between the DMOS values and predicted quality scores using the SISBLIM [8] , MW-PSNR [16], MP-PSNR [17], MP-PSNR-Reduc [18], APT [7] and the proposed IQA algorithm. In Fig. 3, the synthesized images generated using seven algorithms are shown by using different colored markers. The figure shows that the proposed algorithm has the most centralized distribution with respect to other

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2888861, IEEE Transactions on Industrial Informatics

TABLE II: Performance of the proposed model and existing IQA models (in terms of PLCC) for different reconstruction methods used in IRCCyN database. A1 to A7 represent 7 methods used to generate 3D virtual views in IRCCyN/IVC database.

Algorithm	SSIM	RRED	SISBLIM	MW-PSNR	MP-PSNR	MP-PSNR-Reduc	APT	Prop.
A1	0.2961	0.5415	0.3998	0.1516	0.2033	0.1271	Nan	0.3598
A2	0.6680	0.5305	0.4697	0.8632	0.8945	0.8765	0.2307	0.5799
A3	0.7854	0.1370	0.7263	0.6732	0.7041	0.6638	0.8695	0.8089
A4	0.7449	0.7065	0.4095	0.9321	-0.9514	0.9454	0.8161	0.6931
A5	0.7398	0.7358	0.6912	0.7294	0.8599	0.8986	0.8474	0.8415
A6	0.5998	0.3421	0.009	-0.6230	0.6454	0.7681	0.5489	0.8254
A7	0.8627	0.6238	0.413	0.8064	0.8352	0.9237	0.8306	0.5692
Overall	0.485	0.4072	0.5225	0.5622	0.6174	0.6772	0.7307	**0.768**

IQA metrics. For example, the distribution of the predicted scores using the APT algorithm does not have any change after the 3.5 DMOS values. This observation clearly suggests that the performance of the APT algorithm is really poor when geometrical distortion is missing in the 3D synthesized images. In table III, the processing time of the proposed metric with respect to the existing algorithms is shown. It is noticeable from this table that the other existing NR algorithm APT takes approximately two minutes to judge the quality of 3D synthesized image while the proposed algorithm requires only 0.041 seconds. With this view, the processing time of the proposed algorithm is really small and even may be applied to judge the perceptual quality of the 3D-synthesized videos in real-time.

In Table IV, the statistical significance between the proposed algorithm and the existing 8 algorithms is done using the F-Test and shown in Table IV. Here +1, -1 and 0 represent that the proposed algorithm is statistically better, worse and comparable than the other algorithm. From this table, it can be easily observed that the proposed algorithm is statistically better as compared to the all of the existing algorithms except APT [7]. Although the proposed metric and APT algorithms are statistically comparable but the computational requirements of the proposed algorithm is much lesser than the APT algorithm.

From extensive experimental results, the following arguments can be made:

1) The proposed algorithm is able to perform far better than the previous IQA algorithms (including both algorithms devised for natural and 3D-synthesized images).
2) The proposed algorithm is computationally very simple and can calculate the perceptual standards of 3D syn-

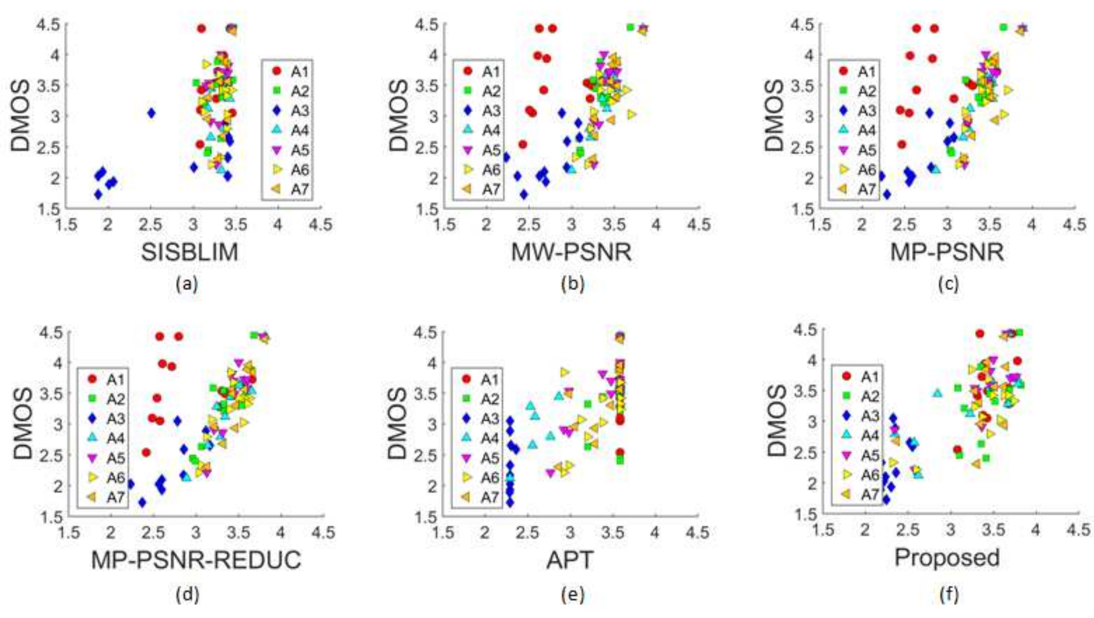


Fig. 3: Scatter plots of MOS values versus IQA models designed for natural images (NR metric: SISBLIM) and for 3D synthesized images (FR metrics: MW-PSNR and MP-PSNR, RR metric: MP-PSNR-reduc, NR metrics: APT and the proposed algorithm).

TABLE III: Computational requirements (in seconds) of existing relevant IQA models.

Metric	MW-PSNR	MP-PSNR	MP-PSNR-reduc	RRLP	APT	Prop.
Avg. Running Time	0.095	0.879	0.549	0.849	101.66	**0.041**

TABLE IV: Statistical comparison of our algorithm with existing IQA models.

Metric	SISBLIM	RRED	SSIM	MAD	MW-PSNR	MP-PSNR	MW-PSNR-reduc	APT
S	+1	+1	+1	+1	+ 1	+1	+1	0

thesized images in real-time applications.

3) Also, the proposed algorithm is able to perform well when virtual views are affected with either of geometrical or structural distortions.

From the above given experimental results and arguments, it can be validated that the proposed algorithm is very efficient and is suitable for estimating the quality score of 3D synthesized images in real-time scenarios.

V. Conclusion and Future Work

High-quality experience of virtual views in FVVs is synthesized with DIBR techniques encompassing DIBR-synthesized images. Nevertheless, highly coherent and fast blind IQA metric for DIBR-synthesized images was missing in the literature. Keeping this as the emphasis, the given work proposed a highly efficient blind metric for judging the quality of 3D-synthesized images. In this work, we assumed geometrical distortions behave like outliers and we empirically proved this assumption using the 3 $sigma$ rule-based ROR statistics. With this view, we proposed an IQA algorithm using computationally simple median filtering to highlight both geometric and structural distortions and further fused these highlighted distortions to predict the quality of 3D synthesized images. Experiments show that the proposed algorithm outperforms previous FR, RR, and NR metrics and has minimal computational complexity. The Inspection also shows the proficiency of the approach in capturing geometric distortions as well.

The performance of the proposed algorithm is satisfactory when geometrical distortions have predominantly occurred in the 3D synthesized images. As for structural distortions, the performance of the proposed algorithm is not optimal. With this view, in the future, we will work to propose an algorithm which can efficiently identify both structural and geometrical distortions and fuse them to accurately predict the perceptual quality of 3D synthesized images. One way is to identify structural distortions using state-of-art IQA algorithms (such as SISBLIM, or IL-NIQE) and estimate geometrical distortions using the ROR and fuse them based upon the number of distortions. Future work should also consider adding users to the loop in developing IQA algorithms by considering human-factors as perception of image quality is shown to vary based on several user traits such as culture and personality.

References

[1] K. Muller, P. Merkle, and T. Wiegand, "3-D video representation using depth maps," *Proc. IEEE*, vol. 99, no. 4, pp. 643-656, Apr. 2011.

[2] P. Kauff et al., "Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability," *Signal Processing: Image Communication*, vol. 22, pp.217-234, 2007.

[3] Y. Zhang, X. Yang, X. Liu, Y. Zhang, G. Jiang, and S. Kwong, "High-Efficiency 3D Depth Coding Based on Perceptual Quality of Synthesized Video," *IEEE Transactions on Image Processing*, vol. 25, no. 12, pp. 5877-5891, Dec. 2016.

[4] X. Liu, Y. Zhang, S. Hu, S. Kwong, C. C. Kuo, and Q. Peng, "Subjective and Objective Video Quality Assessment of 3D Synthesized Views With Texture/Depth Compression Distortion," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4847-4861, Dec. 2015.

[5] W. Jang, T. Chung, J. Sim, and C. Kim, "FDQM: Fast quality metric for depth maps without view synthesis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 7, pp. 1099-1112, Jul. 2015.

[6] J. Wang, S. Wang, K. Zeng, and Z. Wang, "Quality assessment of multi-view-plus-depth images," *IEEE International Conference on Multimedia and Expo*, Hong Kong, pp. 85-90, 2017.

[7] K. Gu, V. Jakhetiya, J. Qiao, X. Li, W. Lin, and D. Thalmann, "Model-based referenceless quality metric of 3D synthesized images using local image description," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 394-405, Jan. 2018.

[8] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Trans. Broadcast.*, vol. 60, no. 3, pp. 555-567, Sep. 2014.

[9] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "No-reference quality metric of contrast-distorted images based on information maximization," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4559-4565, Dec. 2017.

[10] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 22, no. 3, pp. 209-212, Mar. 2013.

[11] G. Yue, C. Hou, K. Gu, S. Mao, and W. Zhang, "Biologically Inspired Blind Quality Assessment of Tone-Mapped Images," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2525-2536, Mar. 2018.

[12] G. Yue, C. Hou, and T. Zhou, "Blind Quality Assessment of Tone-Mapped Images Considering Colorfulness, Naturalness and Structure," *IEEE Transactions on Industrial Electronics*, 2018.

[13] X. Min, K. Ma, K. Gu, G. Zhai, Z. Wang, and W. Lin, "Unified blind quality assessment of compressed natural, graphic, and screen content images," *IEEE Trans. Image Process.*, vol. 26, no. 11, pp. 5462-5474, 2017.

[14] X. Min, K. Gu, G. Zhai, J. Liu, X. Yang, and C. W. Chen, "Blind quality assessment based on pseudo reference image," *IEEE Trans. Multimedia*, vol. 20, no. 8, pp. 2049-2062, 2018.

[15] E. Bosc, R. Pepion, P. Le Callet, M. Koppel, P. Ndjiki-Nya, M. Pressigout, and L. Morin, "Towards a new quality metric for 3-D synthesized view assessment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 7, pp. 1332-1343, Nov. 2011.

[16] D. Sandic-Stankovic, D. Kukolj, and P. Le Callet, "DIBR-synthesized image quality assessment based on morphological wavelets," *in Proc. IEEE Int. Workshop on Quality of Multimedia Experience*, pp. 1-6, Jan. 2015.

[17] D. Sandic-Stankovic, D. Kukolj, and P. Le Callet, "DIBR-synthesized image quality assessment based on morphological pyramids," *The True Vision-Capture, Transmission and Display of 3D Video*, pp. 1-4, Oct. 2015.

[18] D. Sandić-Stanković, D. Kukolj, and P. Le Callet, "Multi-scale synthesized view assessment based on morphological pyramids", *Journal of Electrical Engineering*, vol. 67, no. 1, pp. 3-11, 2016.

[19] V. Jakhetiya, K. Gu, W. Lin, Q. Li, and S. Prasad Jaiswal, "A prediction backed model for quality assessment of screen content and 3-D synthesized images," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 652-660, 2018.

[20] V. Jakhetiya, W. Lin, S. Jaiswal, K. Gu, and S. C. Guntuku, "Just noticeable difference for natural images using RMS contrast and feedback mechanism," *Neurocomputing*, vol. 275, pp. 366-376, 2018.

[21] B. Xiong and Z. Yin, "A universal denoising framework with a new impulse detector and nonlocal means," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1663-1675, Apr. 2012.

[22] A. Roy, L. Manam, and R. H. Laskar, "Region adaptive fuzzy filter: An approach for removal of random valued impulse noise," *IEEE Transactions on Industrial Electronics*, vol. PP, no. 99, pp. 1-1, 2018.

[23] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, Jan. 2010.

[24] M. Narwaria, W. Lin, I. V. McLoughlin, S. Emmanuel, and L. T. Chia, "Fourier transform-based scalable image quality measure," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3364-3377, Aug. 2012.

[25] R. Soundararajan and A. C. Bovik, "RRED indices: Reduced-reference entropic differencing for image quality assessment," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 517-526, Feb. 2012.

[26] F. Battisti, E. Bosc, M. Carli, and P. Le Callet, "Objective image quality assessment of 3D synthesized views," *Signal Process. Image Commun.*, vol. 30, pp. 78-88, Jan. 2015.

[27] P. H. Conze, P. Robert, and L. Morin, "Objective view synthesis quality assessment," *Electron. Imag. Int. Society for Optics and Photonics*, pp. 8288-8256, Feb. 2012.

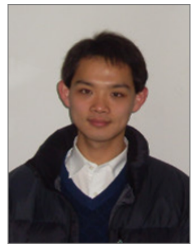
[28] G. Yue, C. Hou, K. Gu, T. Zhou, and G. Zhai, "Combining Local and Global Measures for DIBR-Synthesized Image Quality Evaluation," *IEEE Transactions on Image Processing*, 2019.

[29] B. T. Oh, J. Lee, and D. Park, "Depth map coding based on synthesized view distortion function," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 7, pp. 1344-1352, Nov. 2011.

[30] R. Song, H. Ko, and C. C. Kuo, "MCL-3D: A database for stereoscopic image quality assessment using 2D-image-plus-depth source," *arXiv preprint arXiv:1405.1403*, 2014.

[31] S. Tian, L. Zhang, L. Morin, and O. Deforges, "NIQSV+: A no-reference synthesized view quality assessment metric," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 1652-1664, 2018.

Vinit Jakhetiya received the B.Tech. degree in computer and communication engineering from LNM Institute of Information Technology, India, in 2011, and the Ph.D. degree in electronics and computer engineering from Hong Kong University of Science and Technology, Hong Kong, in 2016. From January 2015 to December 2015, he was a Visiting Student with the School of Computer Science and Engineering, Nanyang Technological University, Singapore, and later he joined as a Project Officer at the same university. Currently, he is working as an assistant professor, Indian Institute of Technology, Jammu, India. His research interests include image/video processing, image quality assessment, and visual perceptual modeling.



Ke Gu received the B.S. and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015, respectively. He is currently a Professor with the Beijing University of Technology, Beijing, China. His research interests include environmental perception, image processing, quality assessment, and machine learning. He received the Best Paper Award from the IEEE Transactions on Multimedia, the Best Student Paper Award at the IEEE International Conference on Multimedia and Expo in 2016, and the Excellent Ph.D. Thesis Award from the Chinese Institute of Electronics in 2016. He was the Leading Special Session Organizer in the VCIP 2016 and the ICIP 2017, and serves as a Guest Editor for the Digital Signal Processing Journal. He is currently an Associate Editor of the IEEE Access and the IET Image Processing. He is a Reviewer for 20 top SCI journals.

Trisha Singhal is an undergraduate student in Computer Science and Engineering stream and expected to receive her Bachelor's degree from Hindustan College of Science and Technology, India, in 2019. She has a strong interest in the field of image processing and machine learning.

Sharath Chandra Guntuku received B.E. (Hons). Computer Science from Birla Institute of Technology and Science (BITS) - Pilani, India, in 2013 and PhD Computer Science from Nanyang Technological University in 2017. He is a research scientist at the University of Pennsylvania. His research interests include multimedia analysis, image processing and applications in health and psychology.

Zhifang Xia received the B.S. degree in measuring and control instrument from Anhui University, Hefei, China in 2008 and received the Master degree in control science and engineering from Tsinghua University, Beijing, China in 2012. She is currently an engineer and a registered consultant(investment) with State Information Center, Beijing, China. Her interests include image processing, quality assessment, machine learning and e-government. She won the second prize of National excellent engineering consultation award in 2016.

Weisi Lin (M'92-SM'98-F'16) received his Ph.D. from King's College, London University, U.K. He served as the Lab Head of Visual Processing, Institute for Infocomm Research, Singapore. Currently, he is an Professor in the School of Computer Science and Engineering. His areas of expertise include image processing, video compression, and multimedia communication, in which he has published 200 journal papers and 200+ conference papers, filed 7 patents, and authored 2 books. He is an AE for IEEE Trans. on Image Processing, and IEEE Trans. Circuits and Systems for Video Tech. He is an IEEE fellow.