

Article

Quality Assessment of 3D Synthesized Images Based on Textural and Structural Distortion Estimation

Hafiz Muhammad Usama Hassan Alvi ^{1,†}, Muhammad Shahid Farid ^{1,*,†} , Muhammad Hassan Khan ^{1,†}  and Marcin Grzegorzec ^{2,†} 

¹ Punjab University College of Information Technology, University of the Punjab, Lahore 54000, Pakistan; mscsf17m027@pucit.edu.pk (H.M.U.H.A.); hassankhan@pucit.edu.pk (M.H.K.)

² Institute of Medical Informatics, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany; grzegorzec@imi.uni-luebeck.de

* Correspondence: shahid@pucit.edu.pk

† All authors contributed equally to this work.

Abstract: Emerging 3D-related technologies such as augmented reality, virtual reality, mixed reality, and stereoscopy have gained remarkable growth due to their numerous applications in the entertainment, gaming, and electromedical industries. In particular, the 3D television (3DTV) and free-viewpoint television (FTV) enhance viewers' television experience by providing immersion. They need an infinite number of views to provide a full parallax to the viewer, which is not practical due to various financial and technological constraints. Therefore, novel 3D views are generated from a set of available views and their depth maps using depth-image-based rendering (DIBR) techniques. The quality of a DIBR-synthesized image may be compromised for several reasons, e.g., inaccurate depth estimation. Since depth is important in this application, inaccuracies in depth maps lead to different textural and structural distortions that degrade the quality of the generated image and result in a poor quality of experience (QoE). Therefore, quality assessment DIBR-generated images are essential to guarantee an appreciative QoE. This paper aims at estimating the quality of DIBR-synthesized images and proposes a novel 3D objective image quality metric. The proposed algorithm aims to measure both textural and structural distortions in the DIBR image by exploiting the contrast sensitivity and the Hausdorff distance, respectively. The two measures are combined to estimate an overall quality score. The experimental evaluations performed on the benchmark MCL-3D dataset show that the proposed metric is reliable and accurate, and performs better than existing 2D and 3D quality assessment metrics.



Citation: Alvi, H.M.U.H.; Farid, M.S.; Khan, M.H.; Grzegorzec, M.; Quality Assessment of 3D Synthesized Images Based on Textural and Structural Distortion Estimation. *Appl. Sci.* **2021**, *11*, 2666. <https://doi.org/10.3390/app11062666>

Academic Editor: Marius Pedersen

Received: 1 February 2021

Accepted: 15 March 2021

Published: 17 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: 3D image quality assessment; objective quality metric; view synthesis; 3D television

1. Introduction

Three-dimensional (3D) technologies, e.g., augmented reality, virtual reality, mixed reality, and stereoscopy, have lately enjoyed remarkable growth due to their numerous applications in the entertainment industry, gaming industry, for electro-medical equipment, etc. 3D television (3DTV) and the recent free-viewpoint television (FTV) [1] have enhanced users' television experience by providing immersion. 3DTV projects two views of the same scene from slightly different viewpoints to provide the depth sensation. The FTV, in addition to the immersive experience, enables the viewer to enjoy the scene from different viewpoints by changing his/her position in front of the television. To provide a full parallax, FTV needs dozens of views, ideally an infinite number of views. Capturing, coding, and transmitting such a large number of views is not practical due to various financial and technological constraints, such as limited available bandwidth. Therefore, novel 3D video (3DV) formats and representations have been explored to design compression-friendly and cost-efficient solutions. The multiview video plus depth (MVD) format is considered to be

the most suitable for 3D televisions. In addition to color images, MVD also provides the corresponding depth maps, which represent the geometry of the 3D scene.

The additional dimension of the depth in MVD provides the ability to generate novel views from a set of available views using the depth-image-based rendering (DIBR) technique [2], thus enabling the stereoscopy. The quality of the synthesized views is important for a pleasant user experience. Since the depth maps are usually generated using stereo-matching algorithms [3], they are not accurate. The inaccuracies in depth maps, when used in DIBR, might introduce various distortions in the synthesized images degrading their quality and resulting in a poor quality of experience (QoE). Thus, assessing the quality of the DIBR-synthesized views is necessary to ensure a satisfactory user experience.

Inaccuracies in depth maps cause textural and structural distortions such as ghost artifacts and inconsistent object shifts in the synthesized views [4–8]. Texture and depth compression also introduce artifacts in the virtual images [9,10]. Another factor that causes degradation in virtual image quality is occluded areas in the original view that become visible in the virtual view, which are called holes. These holes are usually estimated using image inpainting techniques that do not always produce a pleasant reconstruction. Figure 1a shows the artifacts introduced in a synthesized view due to visible occluded regions. Note the distorted face of a spectator in Figure 1b because of erroneous depth in DIBR.



Figure 1. Effect of inaccurate depth maps on the depth-image-based rendering (DIBR)-synthesized images. (a) Occluded regions becoming visible in the synthesized image and generating holes highlighted in red rectangles, (b) structural distortions in the synthesized view distorting the face of a spectator.

The various structural and textural distortions introduced in DIBR images may affect the picture quality, the depth sensation, and the visual comfort, which are considered three main factors of user quality-of-experience (QoE) [6]. Besides viewing experience, studies show that the distortion in 3D images can affect the performance of various applications designed for the 3D environment, such as image saliency detection, video target tracking, face detection, and event detection [11–13]. This means that the image quality is very important not only for viewer satisfaction in a stereoscopic environment but also for various 3D applications built for this environment. Therefore, 3D image quality assessment (3D-IQA) is an essential part of the 3D video processing chain.

In this paper, we propose a 3D-IQA metric to estimate the quality of DIBR-synthesized images. The proposed metric aims to measure the structural and textural distortions introduced in the synthesized image due to depth-image-based rendering and combines them to predict the overall quality of the image. The structural details in an image are considered important for their quality as the human visual system (HVS) is more sensitive to them [14,15]. It is the difference between luminance or color that makes the representation of an object or the main features of an image distinguishable. The distortion in these features, referred to as textural distortion, is also important for a true image quality estimation. The textural and structural metric scores are combined to obtain an overall quality score.

The rest of the paper is organized in the following way. Section 2 reviews the related literature, Section 3 presents the proposed 3D-IQA technique. The experimental evaluation of the proposed metric is carried out in Section 4 and we conclude the research in Section 5.

2. Related Work

The quality of an image can be either assessed through subjective tests or by using an automated objective metric [16]. As human eyes are the ultimate receiver of the image, a subjective test is certainly the best and the most reliable way to assess the visual image quality. In such tests, a set of human observers assigns quality scores to the image, which are averaged to get one score. This method, however, is a time-consuming and expensive approach. Therefore, it was felt necessary to introduce an automatic and fast way to assess the quality of an image. This provides the opportunity for researchers to introduce objective metrics for quantitative image quality evaluation, which proves to be a significant improvement in the field of image quality assessment.

Objective image and video quality metrics can be grouped into three classes based on the availability of the original reference images: full-reference (FR), no-reference (NR), and reduced-reference (RR) [17]. The IQA metric that requires the original reference image to evaluate the quality of its distorted version is referred to as a full-reference metric. The IQA approach that assesses the quality of an image in the absence of a corresponding reference image is classified as the no-reference metric. The reduced-reference metrics lie between the two categories, they do not require the reference images but some of their features must be available for comparison.

In the literature, several 2D and 3D objective quality assessment metrics have been proposed to assess visual image quality. Initially, 2D metrics were used to assess the quality of 3D content, however, the use of conventional 2D metrics was found inappropriate to assess the true quality of 3D images due to several additional factors of 3D videos that were not considered by 2D-IQA algorithms [18–20]. Therefore, novel IQA algorithms were needed to evaluate the quality of 3D videos. Such algorithms, in addition to 2D-related artifacts, must also consider artifacts introduced due to the additional dimension of depth in the videos.

In recent years, several algorithms have been proposed to evaluate the quality of 3D images. Many of them utilize the existing 2D quality metrics for this purpose, e.g., [21–24]. Since these algorithms rely on metrics especially designed for 2D images, they do not consider the most important factor of 3D images, i.e., depth, and therefore they are not accurate and reliable.

Many 3D-IQA techniques consider depth/disparity information while assessing the quality of 3D images, e.g., [25–28]. You et al. [19] adopted a belief-propagation-based method to estimate the disparity and combined the quality maps of distorted image and distorted disparity computed using conventional 2D metrics. The method proposed in [25] exploits the disparity as well as binocular rivalry to determine the quality. It uses the Multi-scale Structural Similarity Index Measure (MSSIM) [29] metric to evaluate the quality of disparity of stereo images. Zhan et al. [26] presented a machine-learning-based method that works by learning the features from 2D-IQA metrics and specially designed 3D features using the Scale Invariant Feature Transform (SIFT) flow algorithm [30], and was used to obtain the depth information. The different features of disparity and three types of distortions (blur, noise, and compression) were used by [28] in evaluating the quality of 3D images. These features were used to train a quality prediction model by using the random forest regression algorithm. The method proposed in [18] addressed the issue of structural distortion in a synthesized view due to DIBR, but the method is limited to structural distortions so it cannot be used to evaluate the overall quality of the image.

The 3D-IQA method presented in [31] identifies the disocclusion edges in the synthesized image and inversely maps them to the original image, and the corresponding regions are then compared to assess the quality. The algorithm in [32] uses feature matching points in the synthesized and reference images to compute the quality degradation. The Just

Notice Difference (JND) model is exploited in [33] to compute the global sharpness and distortion in holes in the DIBR image to assess its quality. The quality metric proposed in [34] identifies the critical blocks in the DIBR synthesized image and the reference image. The texture and color contrast similarities between these blocks are compared to estimate the quality of the synthesized image. The method in [35] works by extracting the features of energy-weighted spatial and temporal information and entropy. Then, support vector regression uses these features for depth estimation. Gorley et al. proposed a stereo-band-limited contrast method in [36] that considers contrast sensitivity and luminance changes as important factors for the assessment of image quality. The method presented in [37] extracts the natural scene features from a discrete cosine transform (DCT) domain, and a deep belief network (DBN) model was trained to get the deep features. These generated deep features and DMOS values were used to train a support vector regression (SVR) model to predict the image quality. The learning framework proposed in [38] also uses a regression model to learn the features and besides assessing the quality, it also improves the quality of stereo images. The method proposed in [39] considers the global visual characteristics by using structural similarities and the local quality was evaluated by computing the local magnitude and local phase. The global and local quality scores were combined to get the final score.

Binocular perception or binocular rivalry is an important factor in 3D image quality assessment [40,41]. Humans perceive images with both eyes and it is obvious that there is a difference between the perceptions of the left and the right eye in relation to an image. Indeed, binocular rivalry is the visual perception phenomenon in which there exists a difference in the perception of an image when it is seen from the left eye and the right eye. This difference is called the binocular parallax or binocular disparity. The binocular disparity can be divided into horizontal and vertical parallax. The horizontal parallax affects depth perception and the vertical parallax affects visual comfort [37]. This binocular perception was taken into account in [42] and a binocular fusion process was proposed for quality assessment of stereoscopic images. The 3D-IQA metric proposed in [41] is also based on binocular visual characteristics. A learning-based metric [43] uses binocular receptive field properties for assessing the quality of stereo images. Shao et al. [44] proposed a metric that simplifies the process of binocular quality prediction by dividing the problem into monocular feature encoding and binocular feature combination.

Lin et al. combine binocular integration behaviors such as binocular combination and binocular frequency integration with conventional 2D metrics in [45] to evaluate the quality of stereo images. Binocular spatial sensitivity influenced by binocular fusion and binocular rivalry properties was taken into consideration in [46]. The method proposed in [47] uses binocular responses, e.g., binocular energy response (BER), binocular rivalry response (BRR), and local structure distribution, for 3D-IQA. Quality assessment of asymmetrically distorted stereoscopic images was targeted in [48]. The method is inspired by binocular rivalry and it uses estimated disparity and Gabor filter responses to create an intermediate synthesized view whose quality is estimated using 2D-IQA algorithms. A multi-scale model using binocular rivalry is presented in [49] for quality assessment of 3D images. Numerous other 3D-IQA algorithms use binocular cues for evaluating the quality of 3D images, e.g., [50–52].

3. The Proposed Technique

In multiview video-plus-depth (MVD) format, depth-image-based rendering (DIBR) is used to generate virtual views at novel viewpoints to support 3D vision in stereoscopic and autostereoscopic displays. The DIBR obtains the virtual view by warping the original left and right views to a virtual viewpoint with the help of the corresponding depth maps. As discussed earlier, when the virtual view is generated its quality may degrade due to several structural or textural distortions introduced during synthesis. The major cause of these distortions is the inaccurate depth. This inaccuracy in the depth estimates and other compression-related artifacts can cause several distortions in the synthesized image,

such as ghost artifacts, holes, and blurry regions, as shown in Figure 1. These distortions degrade the image quality and eventually result in poor overall user quality of experience (QoE). Estimating the quality of the synthesized image is therefore important to ensure better QoE. We propose a 3D-IQA metric that attempts to estimate the distortions introduced in synthesized images. Specifically, the proposed metric is a combination of two measures: one estimates the variations in the texture and the other calculates the deterioration in the structures in the image.

3.1. Estimating the Textural Distortion

Textures are complex visual patterns, composed of spatially organized entities that have characteristic brightness, color, shape, and size. The texture is an important discriminant characteristic of an image region [53] and can be used for various purposes such as segmentation, classification, and synthesis [54]. Image texture gives us information about the spatial arrangement of color or intensities in an image or a selected region of an image. During the process of DIBR, the texture of the synthesized image can be adversely affected due to object shifting, incorrect rendering of textured areas, and blurry regions [55]. Object shifting may cause translation or changes in the size of the region in the synthesized view. Due to the translation of objects, the occluded areas in the original view may become visible in the synthesized view, and these are known as holes. These holes are usually estimated using image inpainting techniques that do not always produce accurate reconstruction and result in the incorrect rendering of texture areas and blurry regions in the synthesized view. Given a DIBR-synthesized image and its corresponding reference, the proposed metric estimates the texture distortion by computing the local variations in their contrasts.

Image contrast is an important feature of texture, a basic perceptual attribute and also an important characteristic of the human visual system (HVS) [56,57]. Contrast sensitivity is one of the dominating factors in the research of visual perception [58]. It can be defined as the difference between luminance or color that makes the representation of an object distinguishable. The most famous contrast computation methods are the Michelson and Weber contrast formulas [58]. There are a few methods that use some form of contrast to assess the quality of images [36,59–61].

The proposed metric captures the local variation in contrast of the synthesized image and its reference image. The two images are low-pass filtered to smooth their high spatial frequencies. This is achieved with a small Gaussian filter w of size 3×3 .

$$w_{ij} = \frac{1}{\alpha_g} \exp\left(-\frac{i^2 + j^2}{2\sigma_g^2}\right) \quad (1)$$

where α_g is a normalization term that ensures $\sum w_{i,j} = 1$ and σ_g is Gaussian variance, which controls the weight distribution and the filter size. Let I and R be the filtered synthesized image and its reference image of size $M \times N$. Let x_{ij} represent a block of size $m \times n$ in image I centered at pixel (i, j) , and y_{ij} be its corresponding block in reference image R centered at pixel location (i, j) . Let x_i and y_i represent the i -th corresponding blocks of I and R . The mean μ , variance σ^2 , and standard deviations σ of a block x_{ij} are computed.

$$\mu_{ij} = \frac{1}{mn} \sum_{a=-\frac{m}{2}}^{\frac{m}{2}} \sum_{b=-\frac{n}{2}}^{\frac{n}{2}} x(i+a, j+b) \quad (2)$$

$$\sigma_{ij}^2 = \frac{1}{mn-1} \sum_{a=-\frac{m}{2}}^{\frac{m}{2}} \sum_{b=-\frac{n}{2}}^{\frac{n}{2}} (x(i+a, j+b) - \mu_{x_{ij}})^2 \quad (3)$$

These statistics for y_{ij} are computed analogously. The variation in contrast ψ_{ij} between the blocks x_{ij} and y_{ij} is then computed.

$$\psi_{ij} = \frac{2\sigma_{ij}\sigma_{ij} + c}{\sigma_{ij}^2 + \sigma_{ij}^2 + c} \quad (4)$$

where c is a small constant used to stabilize the equation. The ψ scores of all pixels in I are computed and averaged to obtain the texture distortion score T of the synthesized image.

$$T = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \psi_{ij} \quad (5)$$

3.2. Estimating the Structural Distortion

The study presented in [14] shows that the human visual system (HVS) is highly adapted for extracting structural information from the image. The inaccuracies and compression artifacts in the depth map adversely affect the structural details of the image during the process of DIBR, generally distorting the edges and gradients in the images [55,60]. The depth compression may cause the pixels to be lost or wrongly projected in the synthesized view. Similarly, the estimation inaccuracies in the depth cause ghost artifacts, inconsistent object shift, and distortion of edges in the synthesized view. These distortions in the image affect both the texture and the structure of the image. Therefore it is equally important to compute the structural dissimilarities in the image to assess its quality. Several methods are proposed to compute the structural similarity in 2D images, e.g., [29,60,62–65].

We used the Hausdorff distance [66] to compute the structural similarity score. The Hausdorff distance measures the degree of mismatch between two sets [66,67]. Similar to a texture distortion metric, this mismatch is also computed locally. The Hausdorff distance can be computed for grayscale images, e.g., [68], and for binary images [66,67]. In the proposed metric, since we want to estimate the distortion in the structural details in the warped image compared to the reference image, the edges in the two images are detected and these edge images are used to estimate the degree of mismatch. Any edge detector can be used for this purpose, however, similar to [66], in our study we used the Canny edge detector [69] to compute the edge maps. The Hausdorff distance between two image blocks x_{ij} and y_{ij} of size $m \times n$ centered at location (i, j) in image I and R , respectively, as defined in the preceding section, is computed as follows:

$$HD(x_{ij}, y_{ij}) = \max(hd(x_{ij}, y_{ij}), hd(y_{ij}, x_{ij})) \quad (6)$$

The function $hd(x_{ij}, y_{ij})$ is called the directed Hausdorff distance from x_{ij} to y_{ij} and it can be defined as

$$hd(x_{ij}, y_{ij}) = \max_{a \in x_{ij}} \min_{b \in y_{ij}} \|a - b\|. \quad (7)$$

Equation (7) identifies the point a in x_{ij} that is the farthest from any point in y_{ij} and measures its distance from the nearest neighboring point in y_{ij} . The function $hd(x_{ij}, y_{ij})$ then ranks each point of x_{ij} according to its distance from the nearest neighbor in y_{ij} and picks the largest distant point from these ranked distances because it is the most mismatched point between the reference and distorted image blocks. Similarly, the directed Hausdorff distance from y_{ij} to x_{ij} is computed. In $hd(x_{ij}, y_{ij})$ and $hd(y_{ij}, x_{ij})$, the former represents the degree of mismatch between the synthesized and original image block and the latter represents the degree of mismatch between the original and the synthesized image blocks. Then the largest of the two is chosen as the mismatch score (Equation (6)). The obtained value is normalized.

$$HD_{ij} = \frac{HD(x_{ij}, y_{ij})}{255 \times \sqrt{mn}} \quad (8)$$

The value of HD_{ij} falls in the interval $[0, 1]$. Recall that HD_{ij} is the degree of mismatch, and the structural similarity score is computed by subtracting this normalized Hausdorff score (HD_{ij}) from 1.

$$\eta_{ij} = 1 - HD_{ij} \quad (9)$$

The structural scores for all k blocks are computed and averaged to obtain a single score S .

$$S = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \eta_{ij} \quad (10)$$

3.3. Final Quality Score

The textural and structural scores of the synthesized image computed using Equations (5) and (10), respectively, are combined to compute the overall quality score Q .

$$Q = \alpha T + (1 - \alpha)S. \quad (11)$$

The parameter α is used to adjust the relative importance of textural and structural scores. Its value is empirically set to 0.7. Figure 2 shows the results of the proposed metric on a few sample images from the testing dataset. In a stereoscopic environment, the quality scores obtained by the proposed metric for both views are averaged to get a single quality score. Figure 2b–e are images obtained by DIBR synthesis from the two source color and depth images, the source view images were artificially degraded by introducing additive white noise (AWN) at four different levels, with noise control parameters 5, 17, 33, and 53, respectively. Figure 2a shows the corresponding ground truth image. Below each image, the scores estimated by the proposed metric and the respective subjective scores are reported. It can be noted that the visual quality of the synthesized images degrades as the noise in the source color and depth images increases and our metric effectively captures this quality degradation.

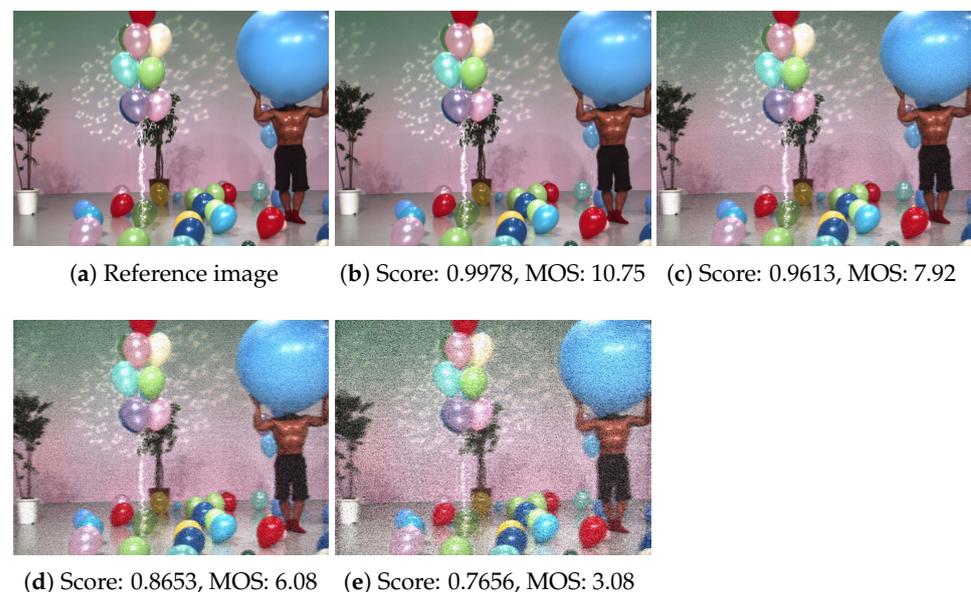


Figure 2. Results of the proposed quality assessment metric on sample images from the test dataset. (a) Original reference image, (b–e) are the images generated using DIBR from color and depth images polluted with additive white noise (AWN) at four different levels, with noise control parameters 5, 17, 33, and 53, respectively. The objective quality scores estimated by the proposed method and the subjective scores of these images are shown below them.

4. Experiments and Results

The performance of the proposed 3D video quality assessment metric was evaluated on the benchmark stereoscopic Media Communications Lab – MCL-3D dataset [70] and compared with other 2D and 3D-IQA metrics. We conducted multiple experiments of different types to evaluate the performance and statistical significance of our proposed method. The results were also compared with existing 2D- and 3D-IQA algorithms.

4.1. Dataset

The MCL-3D dataset was used to evaluate the performance of the proposed quality metric. The dataset analyzes the impact of different distortions on the quality of depth-image-based rendering (DIBR) synthesized images. MCL-3D dataset was created by Media Communication Lab, University of Southern California, and is publicly available [70]. The dataset was created from 9 multiview-video-plus-depth (MVD) sequences. The resolution of 3 MVD sequences is 1024×768 whereas the remaining 6 sequences have a resolution of 1920×1080 . The dataset reports 648 mean opinion scores (MOSs) of stereo image pairs generated using DIBR from distorted texture and/or depth images. Six different types of distortions (Gaussian blur (Gauss), additive white noise (AWN), down-sampling blur (Sample), JPEG compression (JPEG), JPEG 2000 compression (JP2K), and transmission loss (transloss)) with four different levels are applied either on texture images, depth images, or both. The distorted texture images and depth maps are used to generate the intermediate middle virtual images by using view synthesis reference software (VSRS) [71], a benchmark DIBR technique for generating synthesized views. Sample reference and distorted DIBR images are shown in Figure 3.

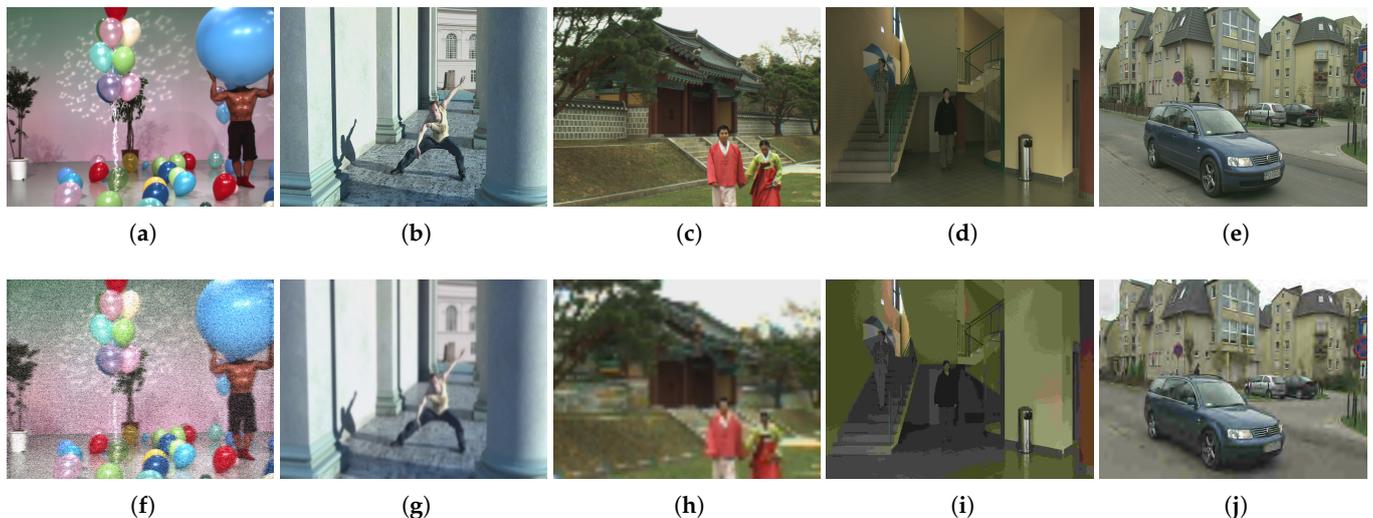


Figure 3. Top row: sample reference images from Balloon (a), Dancer (b), Love_bird1 (c), Poznan_Hall2 (d), and Poznan Street (e) 3D sequences, from the MCL-3D dataset. Bottom row: top image with AWN (f), Gauss (g), Sample (h), JPEG (i), and JP2K (j) distortions, respectively.

4.2. Performance Evaluation Parameters

To evaluate the performance of the proposed method we used different statistical tools, including the Pearson linear correlation coefficient (PLCC), Spearman rank-order correlation coefficient (SROCC), Kendall rank-order correlation coefficient (KROCC), root-mean-square error (RMSE), and mean absolute error (MAE). Before computing these parameters, the scores obtained by objective quality metrics were mapped to subjective differential mean opinion score (DMOS) values using the nonlinear logistic regression described in [72]:

$$DMOS_p = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(o - \beta_3))} \right) + \beta_4 o + \beta_5 \quad (12)$$

where o is the score obtained by the objective quality metric, $DMOS_p$ is the mapped score, and β_1, \dots, β_5 are the regression parameters.

The Pearson linear correlation coefficient (PLCC) is used to determine the linear correlation between two continuous variables. Since this method is based on covariance computation, it is considered the best method for measuring statistical relationships. The method was used in the prediction accuracy test. Let x represent the MOS values, y represent the mapped scores, and \bar{x} and \bar{y} represent the mean values of x and y , respectively. PLCC is computed as

$$PLCC = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (13)$$

The Pearson correlation coefficient describes how strong the relationship between subjective MOS and evaluated objective scores is. The value lies between -1 and 1 . Values closer to 1 represent a strong relationship.

The Spearman rank-order correlation coefficient (SROCC) is a nonparametric measure of rank correlation. It assesses how well the relationship between two variables can be described using a monotonic function. The difference between PLCC and SROCC is that the former only assesses linear relationships whereas the latter assesses monotonic relationships that may or may not be linear. For n observations, the SROCC can be computed as

$$SROCC = 1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)} \quad (14)$$

The Kendall rank correlation coefficient (KROCC) is another nonparametric measure to determine the relationship between two continuous variables. Like SROCC, it assesses associations based on the ranks of data. It is used to test the similarities in data when they are ranked by quantities.

$$KROCC = \frac{n_c - n_d}{\frac{1}{2}n(n-1)} \quad (15)$$

where n is the sample size, n_c is the number of concordant pairs and n_d is the number of discordant pairs.

Root-mean-square error (RMSE) is the most widely used performance evaluation measure and it computes the prediction error [73]. Since the method takes the square of the error before computing the average, it gives a relatively high weight to a large error and that is why it is considered an important method for performance evaluation.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (16)$$

Mean absolute error (MAE) is another method to compute the difference between two continuous variables. MAE is a linear score that means that all the individual differences are weighted equally in average.

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (17)$$

Since PLCC, SROCC, and KROCC are the correlations and MAE and RMSE are the errors, large values of correlations and small values of errors indicate a better performance of the quality metric.

4.3. Performance Comparison with 2D and 3D-IQA Metrics

To evaluate the effectiveness of the proposed method, we compared its performance with various existing 2D and 3D-IQA metrics. We compared the performance of the proposed metric with widely used 2D quality assessment metrics: PSNR, SSIM [60], VSNR [74],

IFC [75], MSSIM [29], VIF [76], and UQI [77]. Before computing the performance parameters, the objective scores computed by these metrics were also mapped to MOS values by using the same logistic function mentioned in Equation (12). In all experiments, the implementation of the method provided by the authors or other parties was used. The comparison in terms of all five performance parameters is presented in Table 1. The best results are highlighted in bold for convenience. The results reveal that the proposed algorithm outperforms all the compared 2D-IQA algorithms in all performance parameters. Specifically, the proposed method achieves PLCC 0.8909, SROCC 0.8979, and KROCC 0.7095 with minimum RMSE 1.1816.

Table 1. Overall performance comparison of the proposed and the compared 2D image quality assessment metrics on MCL-3D dataset. The best results are highlighted in bold.

Metric	PLCC	SROCC	KROCC	RMSE	MAE
PSNR	0.8226	0.8400	0.6393	1.4794	1.1830
SSIM	0.7656	0.7777	0.5796	1.6737	1.2810
VSNR	0.8287	0.8368	0.6357	1.4560	1.1520
IFC	0.5104	0.7423	0.5495	2.2376	1.8331
MSSIM	0.8666	0.8762	0.6852	1.2983	1.0188
VIF	0.7983	0.7933	0.5999	1.5668	1.2240
UQI	0.7537	0.7564	0.5595	1.7099	1.3066
Ours	0.8909	0.8979	0.7095	1.1816	0.9245

We also evaluated the performance the proposed metric with thirteen well-known and recent 3D image quality assessment metrics. They include 3DSwIM [55], StSD [78], Chen [25], Benoit [79], Campisi [21], Ryu [50], PQM [80], Gorley [36], You_l, You_g [19], SIQM [73], NIQSV [81], and ST-SIAQ [82]. The evaluation results presented in Table 2 show that the proposed method outperforms all the compared methods in each performance parameter and achieves the best PLCC (0.8909) with the minimum RMSE (1.1816). The other measures, SROCC, KROCC, and MAE, also reveal that the proposed method performs better than other 3D-IQA metrics.

Table 2. Overall performance comparison of the proposed and the compared 3D image quality assessment metrics on MCL-3D dataset. The best results are highlighted in bold.

Metric	PLCC	SROCC	KROCC	RMSE	MAE
3DSwIM	0.6497	0.5683	0.4088	1.9777	1.4902
StSD	0.6995	0.7008	0.5084	1.8593	1.4563
Chen	0.8615	0.8708	0.6769	1.3212	1.0362
Benoit	0.7425	0.7518	0.5589	1.7429	1.3480
Campisi	0.7656	0.7777	0.5796	1.6737	1.2810
PQM	0.8612	0.8646	0.6729	1.3225	1.0029
Gorley	0.7099	0.7196	0.5270	1.8323	1.4749
You _l	0.7504	0.7567	0.5600	1.7196	1.3325
SIQM	0.7744	0.7756	0.5648	1.6461	1.2757
You _g	0.3650	0.6609	0.4938	2.4222	2.0123
Ryu	0.8752	0.8824	0.6923	1.2584	0.9859
NIQSV	0.6783	0.6208	0.4384	1.9118	1.6032
ST-SIAQ	0.7133	0.7034	0.5118	1.8233	1.4028
Ours	0.8909	0.8979	0.7095	1.1816	0.9245

To further investigate the effectiveness of the proposed method, its performance with the compared 2D and 3D-IQA metrics was also evaluated for individual distortion types, i.e., AWN, Gauss, Sample, Transloss, JPEG, and JP2K. Recall that the stereopair images in the dataset were generated through DIBR from the polluted depth and/or the color images with these types of noise. The results of the comparison with 2D and 3D quality metrics in terms of PLCC are reported in Tables 3 and 4. These results show that the proposed metric performs better than the compared method for most individual distortion types. Similar observations were made when evaluated using other performance parameters, i.e., SROCC, KROCC, RMSE, and MAE, which are not shown here to save space.

Table 3. Comparison of the proposed and the compared 2D image quality assessment (2D-IQA) metrics on individual distortion types in terms of Pearson linear correlation coefficient (PLCC). The best results are highlighted in bold.

Metric	AWN	Gauss	Sample	Tansloss	JPEG	JP2K
PSNR	0.9125	0.9047	0.9051	0.8602	0.7491	0.8719
SSIM	0.8652	0.8594	0.8535	0.8542	0.7849	0.8052
VSNR	0.8943	0.8533	0.8449	0.7571	0.7352	0.6150
IFC	0.9114	0.7116	0.7733	0.6004	0.8394	0.5985
MSSIM	0.9103	0.9077	0.9106	0.9011	0.9097	0.8963
VIF	0.9555	0.9285	0.9128	0.8168	0.9417	0.9155
UQI	0.8501	0.8669	0.8608	0.8494	0.8896	0.7537
Ours	0.9081	0.9332	0.9310	0.8195	0.9502	0.9369

Table 4. Comparison of the proposed and the compared 3D image quality assessment (3D-IQA) metrics on individual distortion types in terms of PLCC. The best results are highlighted in bold.

Metric	AWN	Gauss	Sample	Tansloss	JPEG	JP2K
3DSwIM	0.4640	0.8218	0.8127	0.7566	0.6431	0.6478
StSD	0.7472	0.8429	0.8392	0.6527	0.7372	0.7836
Chen	0.9058	0.8978	0.9088	0.8842	0.9078	0.8928
Benoit	0.9102	0.8600	0.8544	0.6796	0.8044	0.8064
Campisi	0.8652	0.8594	0.8535	0.8542	0.7849	0.8052
Ryu	0.9343	0.9409	0.9163	0.8685	0.8935	0.9284
PQM	0.8653	0.9346	0.9337	0.8052	0.8732	0.9027
Gorley	0.7735	0.8550	0.8544	0.6043	0.8326	0.9051
You _l	0.9278	0.8560	0.8564	0.7930	0.7934	0.8071
SIQM	0.7798	0.8738	0.8673	0.6851	0.8858	0.7057
NIQSV	0.8184	0.9156	0.8862	0.719	0.2798	0.7070
ST_SIAQ	0.7884	0.8215	0.8070	0.8264	0.742	0.7884
You _g	0.8856	0.8685	0.7138	0.5373	0.8599	0.9121
Ours	0.9081	0.9332	0.9370	0.8195	0.9502	0.9369

4.4. Variance of the Residual Analysis

Variance is the squared deviation of a measure from its mean. It is generally used in evaluating the efficiency of an image quality assessment metric by finding how much the scores computed by an objective IQA metric are closer to the subjective scores. This is achieved by computing the difference between the predicted scores and actual scores. A small difference indicates that the results computed by the metric are reliable and close

to the actual scores. To compute this variance σ^2 , first the residual difference between the DMOS and predicted scores after non-linear mapping ($DMOS_p$) is computed:

$$\widehat{R}(i) = DMOS_p(i) - DMOS(i) \quad (18)$$

The variance σ^2 of each compared 2D and 3D quality metric was computed from its residuals \widehat{R} and the statistics are presented in Table 5. The results show that our method achieves the smallest variance among all compared methods, which means the scores estimated by the proposed method are highly correlated with the subjective ratings.

Table 5. Variance of the residuals of subjective ratings and the mapped objective scores of the proposed and the compared 2D and 3D quality metrics. The best results are highlighted in bold.

2D-IQA Metrics		3D-IQA Metrics	
Metric	Variance	Metric	Variance
3DSwiM	3.9174	PSNR	2.1919
StSD	3.4625	SSIM	2.8056
Chen	1.7483	VSNR	2.1232
Benoit	3.0423	IFC	5.0145
Campisi	2.8056	MSSIM	1.6881
Ryu	1.5860	VIF	2.4588
PQM	1.7518	UQI	2.9281
Gorley	3.3626		
You _l	2.9615		
You _g	5.8762		
SIQM	2.7138		
ST_SIAQ	3.3297		
NIQSV	3.6606		
Ours	1.3983		

4.5. Statistical Significance Test

The statistical significance test [16,72] helps to determine whether a quality metric is statistically better than the other metric. We conducted this test to statistically verify the performance of the proposed metric. In this experiment, we considered only the 3D-IQA metrics as the previous evaluations have shown that the 2D-IQA algorithms perform rather poorly compared to 3D-IQA approaches in assessing the quality of DIBR-synthesized images. The F-test procedure was used to test the significance of the difference between two quality assessment metrics. In the F-test, we compared the variances of residuals (Equation (18)) of two metrics i and j with the F-ratio threshold to find the statistical significance. The F-ratio threshold was obtained from the F-distribution look-up table with $\alpha = 0.05$.

$$\frac{\sigma_j^2}{\sigma_i^2} > F - ratio \quad (19)$$

If the $\frac{\sigma_j^2}{\sigma_i^2}$ ratio is greater than the F-ratio then the metric i is said to be significantly superior to the metric j . Similarly, the metric i is said to be significantly inferior to the metric j if this ratio is less than the p-value. The two metrics are said to be statistically indistinguishable if this ratio lies between the p-value and the F-ratio threshold. The F-ratio is called the right-tail critical value and the p-value is called the left-tail critical value. Both of these values were obtained from the F-distribution look-up table at a 95% significance level. The results of the test are presented in Table 6. Each entry in the table is a codeword of 6 characters corresponding to symbols A, G, J, j, S, and T, which represent the distortions AWN, Gauss, JP2K, JPEG, Sample, and Transloss, respectively. In the codeword, the value '1' means that the performance of metric in the row is significantly superior to the metric in the column, the value '0' means that the metric in the row is significantly inferior to

the metric in the column, and ‘-’ means that the performance of metric in the row and column is equivalent or statistically indistinguishable. The results demonstrate that except AWN and Transloss, the performance of the proposed metric is significantly superior or equivalent to all the compared 3D-IQA metrics in all distortion types.

The experimental evaluations performed on the benchmark DIBR synthesized image dataset showed that the performance of the proposed 3D-IQA metric is appreciably better than the compared 2D and 3D-IQA algorithms. Moreover, the variance and the statistical significance tests also showed that our metric is significantly superior or equivalent to most of the compared 3D-IQA metrics. All these performance analyses reveal that the proposed metric is reliable and accurate in estimating the quality of the DIBR-synthesized views.

Table 6. Statistical significance tests of proposed and other 3D-IQA metrics on MCL-3D dataset. Value ‘1’ means the metric in the row is significantly superior to that of the column. Value ‘0’ means the metric in the row is significantly inferior to that of the column and ‘-’ means both the metrics are significantly equivalent. The symbols A, G, J, j, S, and T represent the distortions for AWN, Gauss, JP2K, JPEG, Sample, and Transloss, respectively.

Metric	3DSwIM	Benoit	Campisi	Gorley	PQM	Ryu	SIQM	StSD	You _l	Chen	NIQSV	ST_SIAQ	You _g	Ours
	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST	AGJjST
3DSwIM	-----	0-00--	0-00-0	0-00-1	00000-	000000	0--0--	0-0---	0-00--	000000	00-10-	0-0---	0-0011	00000-
Benoit	1-11--	-----	1----0	1-0---	100000	-00000	1-10--	1-----	-----0	--0000	1011--	1----0	--0-1-	-00000
Campisi	1-11-1	0----1	-----	1-0--1	-0000-	00000-	1-10-1	1-----	0-----	0-000-	-011-1	1-----	--0011	00000-
Gorley	1-11-0	0-1---	0-1--0	-----	00--00	00-000	--10--	--11--	0-1--0	00-000	-011--	--11-0	0--1-	000000
PQM	11111-	011111	-1111-	11--11	-----	0-0--0	111-11	111111	01111-	01-0-0	--111-	11111-	-1--11	0-00--
Ryu	111111	-11111	11111-	11-111	1-1--1	-----	111-11	111111	-11111	111--	1111-1	11111-	11--11	1--0--
SIQM	1--1--	0-01--	0-01-0	--01--	000-00	000-00	-----	---1--	0-01-0	0-0-00	-0-1--	---110	0-0-1-	000000
StSD	1-1---	0-----	0-----	--00--	000000	000000	---0--	-----	0-----	000000	-0-10-	-----	0-001-	000000
You _l	1-11--	-----1	1-----	1-0--1	10000-	-00000	1-10-1	1-----	1-----	-00000	1011--	1-----	1-0011	-0000-
Chen	111111	--1111	1-111-	11-111	10-1-1	000---	1-1-11	111111	-11111	-----	1-11-1	111111	---111	-000-1
NIQSV	11-01-	0100--	-100-0	-100--	--000-	0000-0	-1-0--	-1-01-	0100--	0-00-0	-----	-1-010	010011	0-0000
ST_SIAQ	1-1---	0-----1	0-----	--00-1	00000-	00000-	---001	-----1	0-----	000000	-0-101	-----	0-0011	00000-
You _g	1-1100	--1-0-	--1100	1---0-	-0--00	00--00	1-1-0-	1-110-	0-1100	---000	101100	1-1100	-----	-0-000
Ours	11111-	-11111	11111-	111111	1-11--	0--1--	111111	111111	-1111-	-111-0	1-1111	11111-	-1-111	-----

5. Conclusions

In this paper, a novel 3D-IQA metric was proposed to assess the quality of DIBR-synthesized images. The proposed method merges two metrics, one computing the deterioration in the texture of the synthesized image and the other computing the structural distortions introduced in the synthesized image due to DIBR and various other types of noise. The two measures are weighted-averaged to obtain the overall quality indicator. Experimental evaluations were performed on the MCL-3D dataset, which contains DIBR-synthesized images generated from color and depth images that were subject to different types of noise. The experimental results and comparisons with existing 2D and 3D-IQA metrics demonstrate that the proposed metric is accurate and reliable in assessing the quality of DIBR-synthesized 3D images.

Author Contributions: Conceptualization, H.M.U.H.A. and M.S.F.; methodology, H.M.U.H.A., M.S.F., M.H.K., M.G.; software, H.M.U.H.A., M.S.F.; validation, H.M.U.H.A., M.S.F. and M.H.K.; investigation, H.M.U.H.A., M.S.F., M.H.K., M.G.; writing—original draft preparation, H.M.U.H.A., M.S.F., M.H.K.; writing—review and editing, M.S.F., M.H.K.; supervision, M.S.F., M.H.K., M.G.; project administration, M.S.F., M.H.K., M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the Higher Education Commission, Pakistan, under project '3DViCoQa' grant number NRPU-7389.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tanimoto, M.; Tehrani, M.P.; Fujii, T.; Yendo, T. Free-viewpoint TV. *IEEE Signal Process. Mag.* **2010**, *28*, 67–76. [\[CrossRef\]](#)
2. Fehn, C.; De La Barré, R.; Pastoor, S. Interactive 3-DTV-concepts and key technologies. *Proc. IEEE* **2006**, *94*, 524–538. [\[CrossRef\]](#)
3. Scharstein, D.; Szeliski, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vis.* **2002**, *47*, 7–42. [\[CrossRef\]](#)
4. Merkle, P.; Morvan, Y.; Smolic, A.; Farin, D.; Mueller, K.; De With, P.; Wiegand, T. The effects of multiview depth video compression on multiview rendering. *Signal Process.-Image Commun.* **2009**, *24*, 73–88. [\[CrossRef\]](#)
5. Farid, M.S.; Lucenteforte, M.; Grangetto, M. Edges shape enforcement for visual enhancement of depth image based rendering. In Proceedings of the 2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSp), Pula, Italy, 30 September–2 October 2013; pp. 406–411.
6. Bosc, E.; Le Callet, P.; Morin, L.; Pressigout, M. Visual Quality Assessment of Synthesized Views in the Context of 3D-TV. In *3D-TV System with Depth-Image-Based Rendering: Architectures, Techniques and Challenges*; Zhu, C., Zhao, Y., Yu, L., Tanimoto, M., Eds.; Springer: New York, NY, USA, 2013; pp. 439–473.
7. Farid, M.S.; Lucenteforte, M.; Grangetto, M. Edge enhancement of depth based rendered images. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 5452–5456.
8. Huang, H.Y.; Huang, S.Y. Fast Hole Filling for View Synthesis in Free Viewpoint Video. *Electronics* **2020**, *9*, 906. [\[CrossRef\]](#)
9. Boev, A.; Hollosi, D.; Gotchev, A.; Egiazarian, K. Classification and simulation of stereoscopic artifacts in mobile 3DTV content. In *Stereoscopic Displays and Applications XX*; Woods, A.J., Holliman, N.S., Merritt, J.O., Eds.; International Society for Optics and Photonics, SPIE: San Jose, CA, USA, 2009; Volume 7237, pp. 462–473.
10. Hewage, C.T.; Martini, M.G. Quality of experience for 3D video streaming. *IEEE Commun. Mag.* **2013**, *51*, 101–107. [\[CrossRef\]](#)
11. Niu, Y.; Lin, L.; Chen, Y.; Ke, L. Machine learning-based framework for saliency detection in distorted images. *Multimed. Tools Appl.* **2017**, *76*, 26329–26353. [\[CrossRef\]](#)
12. Fang, Y.; Yuan, Y.; Li, L.; Wu, J.; Lin, W.; Li, Z. Performance evaluation of visual tracking algorithms on video sequences with quality degradation. *IEEE Access* **2017**, *5*, 2430–2441. [\[CrossRef\]](#)
13. Korshunov, P.; Ooi, W.T. Video quality for face detection, recognition, and tracking. *ACM Trans. Multimed. Comput. Commun. Appl.* **2011**, *7*, 1–21. [\[CrossRef\]](#)
14. Wang, Z.; Bovik, A.C.; Lu, L. Why is image quality assessment so difficult? In Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, USA, 13–17 May 2002; Volume 4, p. IV-3313.
15. Wang, Z.; Lu, L.; Bovik, A.C. Video quality assessment based on structural distortion measurement. *Signal Process.-Image Commun.* **2004**, *19*, 121–132. [\[CrossRef\]](#)
16. Liu, X.; Zhang, Y.; Hu, S.; Kwong, S.; Kuo, C.C.J.; Peng, Q. Subjective and objective video quality assessment of 3D synthesized views with texture/depth compression distortion. *IEEE Trans. Image Process.* **2015**, *24*, 4847–4861. [\[CrossRef\]](#)
17. Varga, D. Multi-Pooled Inception Features for No-Reference Image Quality Assessment. *Appl. Sci.* **2020**, *10*, 2186. [\[CrossRef\]](#)
18. Bosc, E.; Le Callet, P.; Morin, L.; Pressigout, M. An edge-based structural distortion indicator for the quality assessment of 3D synthesized views. In Proceedings of the Picture Coding Symposium, Krakow, Poland, 7–9 May 2012; pp. 249–252.
19. You, J.; Xing, L.; Perkis, A.; Wang, X. Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis. In Proceedings of the International Workshop Video Process. Quality Metrics Consum. Electron, Scottsdale, AZ, USA, 13–15 January 2010; Volume 9, pp. 1–6.
20. Seuntjens, P. Visual Experience of 3D TV. Ph.D. Thesis, Eindhoven University of Technology, Eindhoven, Austria, 2006.
21. Campisi, P.; Le Callet, P.; Marini, E. Stereoscopic images quality assessment. In Proceedings of the 15th European Signal Processing Conference, Poznan, Poland, 3–7 September 2007; pp. 2110–2114.
22. Banitalebi-Dehkordi, A.; Nasiopoulos, P. Saliency inspired quality assessment of stereoscopic 3D video. *Multimed. Tools Appl.* **2018**, *77*, 26055–26082. [\[CrossRef\]](#)
23. Sandić-Stanković, D.; Kukolj, D.; Le Callet, P. DIBR-synthesized image quality assessment based on morphological multi-scale approach. *EURASIP J. Image Video Process.* **2016**, *2017*, 4. [\[CrossRef\]](#)
24. Voo, K.H.; Bong, D.B. Quality assessment of stereoscopic image by 3D structural similarity. *Multimed. Tools Appl.* **2018**, *77*, 2313–2332. [\[CrossRef\]](#)
25. Chen, M.J.; Su, C.C.; Kwon, D.K.; Cormack, L.K.; Bovik, A.C. Full-reference quality assessment of stereopairs accounting for rivalry. *Signal Process.-Image Commun.* **2013**, *28*, 1143–1155. [\[CrossRef\]](#)
26. Zhan, J.; Niu, Y.; Huang, Y. Learning from multi metrics for stereoscopic 3D image quality assessment. In Proceedings of the International Conference on 3D Imaging (IC3D), Liege, Belgium, 13–14 December 2016; pp. 1–8.

27. Jung, Y.J.; Kim, H.G.; Ro, Y.M. Critical Binocular Asymmetry Measure for the Perceptual Quality Assessment of Synthesized Stereo 3D Images in View Synthesis. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *26*, 1201–1214. [[CrossRef](#)]
28. Niu, Y.; Zhong, Y.; Ke, X.; Shi, Y. Stereoscopic Image Quality Assessment Based on both Distortion and Disparity. In Proceedings of the 2018 IEEE Visual Communications and Image Processing (VCIP), Taichung, Taiwan, 9–12 December 2018; pp. 1–4.
29. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale structural similarity for image quality assessment. In Proceedings of the Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, 9–12 November 2003; Volume 2, pp. 1398–1402.
30. Liu, C.; Yuen, J.; Torralba, A.; Sivic, J.; Freeman, W.T. SIFT Flow: Dense Correspondence across Different Scenes. In *ECCV*; Forsyth, D., Torr, P., Zisserman, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 28–42.
31. Liang, H.; Chen, X.; Xu, H.; Ren, S.; Cai, H.; Wang, Y. Local Foreground Removal Disocclusion Filling Method for View Synthesis. *IEEE Access* **2020**, *8*, 201286–201299. [[CrossRef](#)]
32. Mahmoudpour, S.; Schelkens, P. Synthesized View Quality Assessment Using Feature Matching and Superpixel Difference. *IEEE Signal Process. Lett.* **2020**, *27*, 1650–1654. [[CrossRef](#)]
33. Wang, L.; Zhao, Y.; Ma, X.; Qi, S.; Yan, W.; Chen, H. Quality Assessment for DIBR-Synthesized Images With Local and Global Distortions. *IEEE Access* **2020**, *8*, 27938–27948. [[CrossRef](#)]
34. Wang, X.; Shao, F.; Jiang, Q.; Fu, R.; Ho, Y. Quality Assessment of 3D Synthesized Images via Measuring Local Feature Similarity and Global Sharpness. *IEEE Access* **2019**, *7*, 10242–10253. [[CrossRef](#)]
35. Chen, Z.; Zhou, W.; Li, W. Blind stereoscopic video quality assessment: From depth perception to overall experience. *IEEE Trans. Image Process.* **2017**, *27*, 721–734. [[CrossRef](#)]
36. Gorley, P.; Holliman, N. Stereoscopic image quality metrics and compression. In *Stereoscopic Displays and Applications XIX*; International Society for Optics and Photonics: San Jose, CA, USA, 29 February 2008; Volume 6803, p. 680305.
37. Yang, J.; Jiang, B.; Song, H.; Yang, X.; Lu, W.; Liu, H. No-reference stereoimage quality assessment for multimedia analysis towards Internet-of-Things. *IEEE Access* **2018**, *6*, 7631–7640. [[CrossRef](#)]
38. Park, M.; Luo, J.; Gallagher, A.C. Toward assessing and improving the quality of stereo images. *IEEE J. Sel. Top. Signal Process.* **2012**, *6*, 460–470. [[CrossRef](#)]
39. Chen, L.; Zhao, J. Quality assessment of stereoscopic 3D images based on local and global visual characteristics. In Proceedings of the 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Hong Kong, China, 10–14 July 2017; pp. 61–66.
40. Huang, S.; Zhou, W. Learning to Measure Stereoscopic 3D Image Perceptual Quality on the Basis of Binocular Rivalry Response. *Appl. Sci.* **2019**, *9*, 906. [[CrossRef](#)]
41. Shao, F.; Lin, W.; Gu, S.; Jiang, G.; Srikanthan, T. Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics. *IEEE Trans. Image Process.* **2013**, *22*, 1940–1953. [[CrossRef](#)]
42. Bensalma, R.; Larabi, M.C. A perceptual metric for stereoscopic image quality assessment based on the binocular energy. *Multidimens. Syst. Signal Process.* **2013**, *24*, 281–316. [[CrossRef](#)]
43. Shao, F.; Li, K.; Lin, W.; Jiang, G.; Yu, M.; Dai, Q. Full-reference quality assessment of stereoscopic images by learning binocular receptive field properties. *IEEE Trans. Image Process.* **2015**, *24*, 2971–2983. [[CrossRef](#)]
44. Shao, F.; Li, K.; Lin, W.; Jiang, G.; Yu, M. Using binocular feature combination for blind quality assessment of stereoscopic images. *IEEE Signal Process. Lett.* **2015**, *22*, 1548–1551. [[CrossRef](#)]
45. Lin, Y.H.; Wu, J.L. Quality assessment of stereoscopic 3D image compression by binocular integration behaviors. *IEEE Trans. Image Process.* **2014**, *23*, 1527–1542. [[CrossRef](#)]
46. Wang, X.; Kwong, S.; Zhang, Y. Considering binocular spatial sensitivity in stereoscopic image quality assessment. In Proceedings of the 2011 Visual Communications and Image Processing (VCIP), Tainan, Taiwan, 6–9 November 2011; pp. 1–4.
47. Zhou, W.; Yu, L. Binocular responses for no-reference 3D image quality assessment. *IEEE Trans. Multimed.* **2016**, *18*, 1077–1084. [[CrossRef](#)]
48. Chen, M.J.; Su, C.C.; Kwon, D.K.; Cormack, L.K.; Bovik, A.C. Full-reference quality assessment of stereoscopic images by modeling binocular rivalry. In Proceedings of the 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, CA, USA, 4–7 November 2012; pp. 721–725.
49. Wang, J.; Rehman, A.; Zeng, K.; Wang, S.; Wang, Z. Quality prediction of asymmetrically distorted stereoscopic 3D images. *IEEE Trans. Image Process.* **2015**, *24*, 3400–3414. [[CrossRef](#)]
50. Ryu, S.; Kim, D.H.; Sohn, K. Stereoscopic image quality metric based on binocular perception model. In Proceedings of the 2012 19th IEEE International Conference on Image Processing, Orlando, FL, USA, 30 September–3 October 2012; pp. 609–612.
51. Shao, F.; Li, K.; Lin, W.; Jiang, G.; Dai, Q. Learning blind quality evaluator for stereoscopic images using joint sparse representation. *IEEE Trans. Multimed.* **2016**, *18*, 2104–2114. [[CrossRef](#)]
52. Liu, Y.; Yang, J.; Meng, Q.; Lv, Z.; Song, Z.; Gao, Z. Stereoscopic image quality assessment method based on binocular combination saliency model. *Signal Process.* **2016**, *125*, 237–248. [[CrossRef](#)]
53. Peddle, D.R.; Franklin, S.E. Image Texture Processing and Data Integration. *Photogramm. Eng. Remote Sens.* **1991**, *57*, 413–420.
54. Livens, S. Image Analysis for Material Characterisation. Ph.D. Thesis, Universitaire Instelling Antwerpen, Antwerpen, Belgium, 1998.

55. Battisti, F.; Bosc, E.; Carli, M.; Le Callet, P.; Perugia, S. Objective image quality assessment of 3D synthesized views. *Signal Process.-Image Commun.* **2015**, *30*, 78–88. [[CrossRef](#)]
56. Webster, M.A.; Miyahara, E. Contrast adaptation and the spatial structure of natural images. *JOSA A* **1997**, *14*, 2355–2366. [[CrossRef](#)]
57. Westen, S.J.; Lagendijk, R.L.; Biemond, J. Perceptual image quality based on a multiple channel HVS model. In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Detroit, MI, USA, 9–12 May 1995; Volume 4, pp. 2351–2354.
58. Peli, E. Contrast in complex images. *JOSA A* **1990**, *7*, 2032–2040. [[CrossRef](#)] [[PubMed](#)]
59. Damera-Venkata, N.; Kite, T.D.; Geisler, W.S.; Evans, B.L.; Bovik, A.C. Image quality assessment based on a degradation model. *IEEE Trans. Image Process.* **2000**, *9*, 636–650. [[CrossRef](#)] [[PubMed](#)]
60. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
61. Kang, K.; Liu, X.; Lu, K. 3D image quality assessment based on texture information. In Proceedings of the 2014 IEEE 17th International Conference on Computational Science and Engineering, Chengdu, China, 19–21 December 2014; pp. 1785–1788.
62. Chen, G.H.; Yang, C.L.; Xie, S.L. Gradient-based structural similarity for image quality assessment. In Proceedings of the 2006 International Conference on Image Processing (ICIP), Atlanta, GA, USA, 8–11 October 2006; pp. 2929–2932.
63. Chen, G.H.; Yang, C.L.; Po, L.M.; Xie, S.L. Edge-based structural similarity for image quality assessment. In Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), Toulouse, France, 14–19 May 2006; Volume 2, p. II.
64. Yang, C.L.; Gao, W.R.; Po, L.M. Discrete wavelet transform-based structural similarity for image quality assessment. In Proceedings of the 2008 15th IEEE International Conference on Image Processing (ICIP), San Diego, CA, USA, 12–15 October 2008; pp. 377–380.
65. Wang, Z.; Bovik, A.C.; Sheikh, H.R. Structural similarity based image quality assessment. In *Digital Video Image Quality and Perceptual Coding*; Marcel Dekker Series in Signal Processing and Communications; CRC Press: Boca Raton, FL, USA, 18 November 2005; pp. 225–241.
66. Huttenlocher, D.P.; Klanderman, G.A.; Rucklidge, W.J. Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **1993**, *15*, 850–863. [[CrossRef](#)]
67. Tsai, C.T.; Hang, H.M. Quality assessment of 3D synthesized views with depth map distortion. In Proceedings of the 2013 Visual Communications and Image Processing (VCIP), Kuching, Malaysia, 17–20 November 2013; pp. 1–6.
68. Vivek, E.P.; Sudha, N. Gray Hausdorff distance measure for comparing face images. *IEEE Trans. Inf. Forensics Secur.* **2006**, *1*, 342–349. [[CrossRef](#)]
69. Canny, J. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *PAMI-8*, 679–698. [[CrossRef](#)]
70. Song, R.; Ko, H.; Kuo, C. MCL-3D: A database for stereoscopic image quality assessment using 2D-image-plus-depth source. *arXiv* **2014**, arXiv:1405.1403.
71. Tian, D.; Lai, P.L.; Lopez, P.; Gomila, C. View synthesis techniques for 3D video. In Proceedings of the Applications of Digital Image Processing XXXII, San Diego, CA, USA, 3–5 August 2009; International Society for Optics and Photonics: San Diego, CA, USA, 2009; Volume 7443, p. 74430T.
72. Sheikh, H.R.; Sabir, M.F.; Bovik, A.C. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Trans. Image Process.* **2006**, *15*, 3440–3451. [[CrossRef](#)]
73. Farid, M.S.; Lucenteforte, M.; Grangetto, M. Evaluating virtual image quality using the side-views information fusion and depth maps. *Inf. Fusion* **2018**, *43*, 47–56. [[CrossRef](#)]
74. Chandler, D.M.; Hemami, S.S. VSNR: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Trans. Image Process.* **2007**, *16*, 2284–2298. [[CrossRef](#)]
75. Sheikh, H.R.; Bovik, A.C.; De Veciana, G. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Process.* **2005**, *14*, 2117–2128. [[CrossRef](#)] [[PubMed](#)]
76. Sheikh, H.R.; Bovik, A.C. Image information and visual quality. *IEEE Trans. Image Process.* **2006**, *15*, 430–444. [[CrossRef](#)]
77. Wang, Z.; Bovik, A.C. A universal image quality index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84. [[CrossRef](#)]
78. De Silva, V.; Arachchi, H.K.; Ekmekcioglu, E.; Kondoz, A. Toward an impairment metric for stereoscopic video: A full-reference video quality metric to assess compressed stereoscopic video. *IEEE Trans. Image Process.* **2013**, *22*, 3392–3404. [[CrossRef](#)] [[PubMed](#)]
79. Benoit, A.; Le Callet, P.; Campisi, P.; Cousseau, R. Quality assessment of stereoscopic images. *EURASIP J. Image Video Process* **2008**, 659024, 2008.
80. Joveluro, P.; Malekmohamadi, H.; Fernando, W.C.; Kondoz, A. Perceptual video quality metric for 3D video quality assessment. In Proceedings of the 2010 3DTV-Conference: The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON), Tampere, Finland, 7–9 June 2010; pp. 1–4.

81. Tian, S.; Zhang, L.; Morin, L.; Deforges, O. NIQSV: A no reference image quality assessment metric for 3D synthesized views. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 1248–1252.
82. Ling, S.; Le Callet, P. Image quality assessment for free viewpoint video based on mid-level contours feature. In Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME), Hong Kong, China, 10–14 July 2017; pp. 79–84.