electronics

MDPI

Article

Quality Assessment of View Synthesis Based on Visual Saliency and Texture Naturalness

Lijuan Tang [1], Kezheng Sun [1], Shuaifeng Huang [1], Guangcheng Wang [2,*] and Kui Jiang [3]

[1] School of Electronics and Information, Jiangsu Vocational College of Business, Nantong 226000, China; 2014114@jsbc.edu.cn (L.T.); 2014111@jsbc.edu.cn (K.S.); 2019026@jsbc.edu.cn (S.H.)
[2] School of Transportation and Civil Engineering, Nantong University, Nantong 226019, China
[3] School of Computer Science, Wuhan University, Wuhan 430072, China; kuijiang@whu.edu.cn
* Correspondence: wanggc@ntu.edu.cn

Abstract: Depth-Image-Based-Rendering (DIBR) is one of the core techniques for generating new views in 3D video applications. However, the distortion characteristics of the DIBR synthetic view are different from the 2D image. It is necessary to study the unique distortion characteristics of DIBR views and design effective and efficient algorithms to evaluate the DIBR-synthesized image and guide DIBR algorithms. In this work, the visual saliency and texture natrualness features are extracted to evaluate the quality of the DIBR views. After extracting the feature, we adopt machine learning method for mapping the extracted feature to the quality score of the DIBR views. Experiments constructed on two synthetic view databases IETR and IRCCyN/IVC, and the results show that our proposed algorithm performs better than the compared synthetic view quality evaluation methods.

Keywords: Depth-Image-Based-Rendering; view synthesis; quality assessment; visual saliency; texture measurement

Citation: Tang, L.; Sun, K.; Huang, S.; Wang, G.; Jiang, K. Quality Assessment of View Synthesis Based on Visual Saliency and Texture Naturalness. *Electronics* **2022**, *11*, 1384. https://doi.org/10.3390/electronics11091384

Academic Editor: Gwanggil Jeon

Received: 4 April 2022
Accepted: 24 April 2022
Published: 26 April 2022

1. Introduction

3D applications have become more and more popular in recent years, because they can provide users with a fully immersive experience, such as Augmented Reality (AR), Virtual Reality (VR), Free Viewpoint Videos (FVV), Mixed Reality (MR), and Multi-View Videos (MVV) [1–3]. Through these, 3D applications support the ability for people to see the same scene from different perspectives, leading to information redundancy and costly storage space. Hence, researchers often only transmit and save two texture images and a depth map, while the others are synthesized by utilizing the DIBR techniques at the receiving terminal [4]. The complete view synthesis includes the collection, processing, and transmission of texture images and depth maps, as well as DIBR view synthesis [5]. The procedure of DIBR view synthesis consists of two steps. The first step is 3D image warping, in which the original viewpoint is back-projected to the 3D scene, then re-projected to the virtual view by the depth map. In 3D warping, this may produce geometric distortions, such as minor cracks and slight shifts, since the pixel position in the synthesized view may not be an integer. Figure 1a gives an example of small cracks. For a change in the viewpoint, the synthesized view may appear as black holes. The second step is disoccluded hole filling [5,6]. Researchers have used many in-painting methods to fill the black holes, such as image in-painting with Markov chains [7] and context-driven hybrid image in-painting [8]. However, these methods are not designed for view synthesis; thus, they may introduce stretching, object warping, and blurry regions in the DIBR-synthesized views. Figure 1b gives examples of blurry regions, and Figure 1c shows examples of object warping and stretching. By the above analysis, both 3D image warping and the disocclusion hole filling will introduce different types of distortion, which is different from the traditional distortions. Therefore, 2D image quality assessment (IQA) methods are not ideal for assessing DIBR views.

Figure 1. Some examples of distortions in the DIBR views.

Due to the difficulty in capturing the geometric distortion of DIBR views by traditional IQA methods [9–16], some studies on evaluating DIBR synthesized images have been proposed, albeit without considering geometric distortions [17]. These metrics can be classified into three types according to using the reference views: Full Reference (FR), Reduced Reference (RR), and No Reference (NR) [1]. A number of works are briefly reviewed here [6,18–25].

LOGS: The geometric distortions score was derived from a combination of the sizes and distortion strength of the dis-occluded region. A reblurring-based strategy generated the global sharpness score. The final score was derived from a combination of the geometric distortions score and the global sharpness score [6].

AR-plus thresholding method: The AR (autoregression)-based local image description evaluates the DIBR-synthesized image quality. After the AR prediction, the geometry distortion can be accurately captured between a DIBR-synthesized image and its AR-predicted image. Finally, the proposed method used visual saliency to improve algorithm performance [19].

MW-PSNR: Morphological Wavelet Peak Signal-to-Noise Ratio metric [20,21]. The morphological wavelet decomposition preserves geometric structures such as edges in lower resolution images. Firstly, morphological wavelet transform was used to decompose the synthesis view and the reference view at multiple scales. Then, the mean square error (MSE) of detail sub-band was calculated. The wavelet MSE was obtained by collecting the MSE values of each scale.

MP-PSNR: The design principle of this metric was based on Pyramid representations that have much in common with the eye's visual system. The morphological pyramids decomposed the reference and synthesized views. The quality score only used detailed images from the higher pyramid [22].

OUT: This method detected the geometrically distorted of 3D Synthesized images using outlier detection. The nonlinear median filtering was used to capture geometric and structural distortion levels and remove outliers [23].

NIQSV: This algorithm assumed that the high-quality image includes sharp edges and flat areas. Morphological operations were not sensitive to images, but local thin deformation can be easily detected. The NIQSV measure first detected thin distortions using an opening operation and then filled the black hole with a closing operation of the larger structural unit [24].

MNSS: Multiscale Natural Scene Statistical (MNSS) analysis measurement [18]. Two Natural Scene Statistics (NSS) models are utilized to evaluate the DBIR views. One of the NSS models was used to capture the geometric distortion introduced by DIBR, which destroys the local self-similarity of the images. Another NSS model was based on statistical regularity, which was destroyed in DIBR synthetic views at different scales.

GDSIC: This method utilized the edge similarity in the Discrete Wavelet Transform (DWT) domain to capture the geometric distortion. The sharpness was estimated by the energies of sub-bands. Then two filters were used to calculate the image complexity. The three parts were combined to produce the final DIBR quality score [2].

CLGM: This method considers both geometric distortion and sharpness distortion, which combining Local and Global Measures. Through the analysis of local similarity, the distortion of the disoccluded area is obtained. Moreover, the typical geometrical deformation stretching is discovered and tested by computing its similarity to the adjacent areas of equal size. Considering the scale invariance, the distance between the distorted DIBR image and the down-sampled image was taken to measure global sharpness. The final score was generated by linearly combining the two geometric distortion and sharpness scores [25].

The DIBR quality metrics mentioned above usually measure the specific distortion during view synthesis. This means that prior knowledge about distortion is needed. Moreover, the computational complexity of the algorithm is also very important in practical applications. Therefore, an effective and efficient quality evaluation method for DIBR visual synthesis is needed.

This paper develops a NR quality metric for DIBR views based on visual saliency and texture naturalness. The design principle for the proposed metric is based on the following facts. First, Human Visual System (HVS) usually searches for and locates critical areas when facing images. This mechanism of visual saliency is of great significance for visual information processing in daily life. Visual saliency mechanism has been successfully applied to target recognition, compression coding, image quality assessment, facial expression recognition, and other visual tasks [26–32]. Second, the HVS has multi-scale characteristics and can extract multi-scale information from images. Moreover, it is sensitive to texture information. Inspired by this, we extract a set of related quality features and adopt a machine learning model for mapping the obtained features to the quality score of DIBR views. Extensive experiments constructed on two public DIBR views databases, IRCCyN-IVC [33], and IETR [34]. The results show the advantages of the proposed method over the relevant state-of-the-art DIBR views quality assessment algorithms.

The rest of this article is arranged as follows. Section 2 explains our designed metric in detail. Section 3 shows experimental studies to verify the effectiveness of our proposed DIBR view quality assessment method. Section 4 gives the conclusion.

2. Proposed Method

This part will describe our designed DIBR views quality assessment algorithm in detail. The distortions in the DIBR synthesized views induce some critical areas and texture degradation. Visual saliency is first used to simulate eye movement "fixation" and "saccade" [35]. Second, Local Binary Pattern (LBP) [36] is utilized to extract texture features of the DIBR images, and the histogram is used to compare texture naturalness. Finally, the extracted features are input into the regression model to train the quality model to predict the quality of DIBR images. The flowchart of our proposed metric is given in Figure 2.

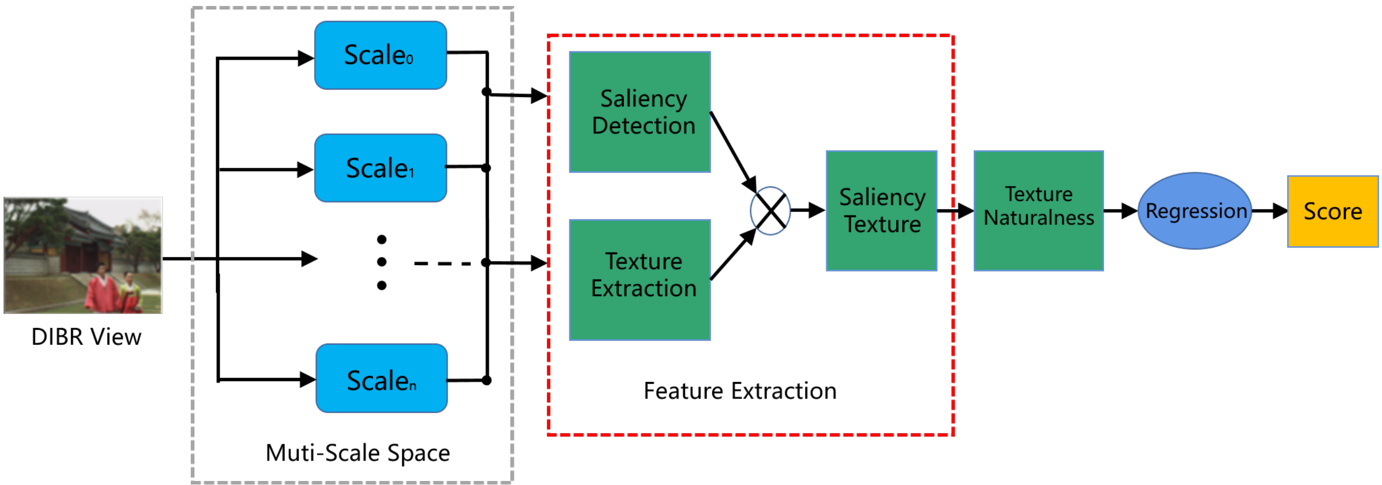

Figure 2. The flowchart of our proposed measurement algorithm.

2.1. Visual Saliency Detection

Most visual saliency models acquire the saliency map by computing the center-surround differences. This work considers the human eye sensitivity alter due to foveation. Hence, the saliency detection model used in this work considers both global and local center-surround differences [37]. The Gaussian low-pass filter simulates the "fixation" and multi-space representation of the visual attention. For a DIBR view $V(m,n)$, the smoothed version of the ith scale $V_i(m,n)$ is defined as:

$$V_G(m,n) = V(m,n) * G(m,n,\sigma_i), \tag{1}$$

where $*$ and σ_i denote the convolution operator and the standard deviation of Gaussian model at the ith scale, respectively. The kernel function is defined as:

$$G(x,\ y,\sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp(-\frac{x^2+y^2}{2\sigma_i^2}), \tag{2}$$

Then we measure local similarity between the smoothed versions and the DIBR view at each scale:

$$S_g(V) = \frac{2VV_G + C}{V^2 + V_G^2 + C}, \tag{3}$$

where V and V_G denotes the DIBR view and its smoothed version, respectively. $S(\cdot)$ calculates the similarity.

This work adopts the method proposed in [35] to approximate the saccade-inspired visual saliency. The similarity is defined as:

$$S_m(V) = \frac{2VV_M + C}{V^2 + V_M^2 + C}, \tag{4}$$

where V_M is generated by convolving V with a MB kernel which is calculated as:

$$M(x,\ y) = \begin{cases} \frac{1}{n} \ if(x \cdot \sin\theta + y \cdot \cos\theta) = 0, x^2 + y^2 \leq \frac{n^2}{4} \\ 0 \ otherwise \end{cases}, \tag{5}$$

where n denotes the amount of motion pixels, and θ represents the motion direction. Then, we integrated the saliency model of fixation and saccade inspired, and obtained the final visual saliency model by using the sample linear weighting strategy as follows.

$$S(V) = \frac{S_g(V) + \alpha S_m(V)}{1+\alpha}, \tag{6}$$

where α is an integer that deals with the relative importance of two components.

2.2. Texture Naturalness Detection

Our proposed method evaluates texture naturalness by the gradient-weighted histogram of the LBP calculated on the gradient map [36]. First, the Scharr operator extracts the gradient magnitude of a DIBR view. It is defined by applying convolution masks to a DIBR view V:

$$GM(V) = \sqrt{(V * gm_x)^2 + (V * gm_y)^2)}, \tag{7}$$

where $*$ denotes the convolution operation, gm_x and gm_y are defined as:

$$gm_x = \frac{1}{16}\begin{bmatrix} +3 & 0 & -3 \\ +10 & 0 & -10 \\ +3 & 0 & -3 \end{bmatrix} \otimes I, \tag{8}$$

$$gm_y = \frac{1}{16}\begin{bmatrix} +3 & 10 & -3 \\ 0 & 0 & 0 \\ +3 & -10 & -3 \end{bmatrix} \otimes I. \tag{9}$$

Then, the LBP operator is used to describe the texture naturalness, which is defined as follows:

$$LBP_{N,R}(x_c, y_c) = \sum_{i=0}^{p-1} t(gm_n - gm_c)2^i. \tag{10}$$

The $t(\cdot)$ is computed as

$$t(gm_n - gm_c) = \begin{cases} 1, & gm_n - gm_c \geqslant 0 \\ 0, & gm_n - gm_c < 0 \end{cases} \tag{11}$$

where N is the total number of neighbors and R is the radius. gm_c and gm_n are the gradient magnitudes at the center location and its neighbor. The setting of values N and R will be introduced in the Section 3, the experimental part.

After extracting the texture by LBP, the texture map is combined with the saliency as follows:

$$W = LBP(V) \cdot S(V), \tag{12}$$

where W is the saliency texture map of the DIBR view, and V represents the DIBR view.

2.3. Quality Assessment Model

In the proposed method, the SVR regression model is adopted to train the quality prediction model for DIBR images from the extracted features to the quality score [38]. Let parameters $\theta > 0$ and $\lambda > 0$, the SVR is defined as:

$$\begin{aligned} \min_{w,\delta,v,v'} \quad & \frac{1}{2}w^T w + \theta \sum_{i=1}^{n}(v_i + rv'_i) \\ \text{subject to} \quad & s_i - w^T\phi(d_i) - \delta \leq \lambda + v'_i, \\ & w^T\phi(d_i) + \delta - s_i \leq \lambda + v_i, \\ & v_i, v'_i \geq 0,\, i = 1, \ldots, 11. \end{aligned}$$

where $K(d_i, d_j) = \phi(d_i)^T\phi(d_j)$ is the kernel function. λ, θ and k are determined by training samples. Subsequently, the trained model is used to predict the quality of the input DIBR views.

3. Experimental Results

3.1. Experimental Databases

The performance of our proposed metric is tested on two public DIBR view synthesis databases, namely the IRCCyN/IVC [33] and IETR [34] databases.

The IRCCyN/IVC database [33] consists of 12 reference images with three sequences (BookArrival, Lovebird, Newspaper) and associated 84 synthesized views that are generated by seven different DIBR algorithms. Figure 3 shows some examples from the IRCCyN/IVC database. The Absolute Category Rating (ACR) algorithm is adopted to test the subjective assessment of the IRCCyN/IVC database. The resolution of each of the synthesized views is 1024×768.

Figure 3. Examples of the DIBR-sythesized views of IRCCyN/IVC database.

The IETR database [34] consists of 10 reference images and 140 associated synthesized views that are generated by seven different DIBR algorithms. Figure 4 shows some examples from the IETR database. The sequences of the database include BookArrival, Lovebird, Newspaper, Balloons, Kendo, Dancer, Shark, Poznan_Street, PoznanHall, and GT_fly. The subjective assessment of the IETR database is based on the ACR algorithm and the Subjective Assessment Methodology for Video Quality (SAMVIQ) algorithm. The resolution of each of the synthesized views is 1920 × 1088.

Figure 4. Examples of the DIBR sythesized views of IETR database.

3.2. Performance Evaluation Criteria

In this work, we adopt a five-parameter nonlinear fitting function to compute evaluation criteria:

$$F(v) = \theta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\theta_2 (v - \theta_3)}} \right) + \theta_4 v + \theta_5, \tag{13}$$

where $F(v)$ denotes subjective score; v represents the corresponding objective score; θ_i is the fitting parameters. Next, three popularly used criteria are employed for performance evaluation. PLCC (Pearson's Linear Correlation Coefficient) is employed to assess the prediction accuracy. The definition of the PLCC is given as follows:

$$\mathrm{PLCC} = \frac{\sum_i (b_i - \bar{b})(l_i - \bar{l})}{\sqrt{\sum_i (b_i - \bar{b})^2 \sum_i l_i - \bar{l})^2}}, \tag{14}$$

where b_i is the estimated value of the i-th DIBR view. $\bar{b}$ is the mean value of all b_i.

The monotony of algorithm prediction is measured by the SRCC (Spearman's Rank ordered Correlation Coefficient). The SRCC is calculated by

$$\text{SRCC} = 1 - \frac{6}{M(M^2-1)} \sum_{i=1}^{M} d_i^2, \tag{15}$$

where M represents the total number of images in the test database; d_i is the rank difference between the objective and subjective evaluation of the i-th image.

The final evaluation index, RMSE (Root Mean Square Error), is used to evaluate the accuracy of the algorithm. The RMSE is computed as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_i (s_i - o_i)^2}, \tag{16}$$

where s_i represents subject assessment values and o_i represents the predicted value. A good DIBR quality evaluation algorithm should obtain high SRCC and PLCC values and low RMSE values.

3.3. LBP Parameter Settings and Computational Complexity Analysis

In the actual IQA system, the efficiency and accuracy of the algorithm are very important. Therefore, we conduct experiments to test the influence of LBP radius and the number of neighbors on algorithm performance, and the corresponding computational complexity. The experimental results are shown in Table 1. From the Table 1, it can be seen that the radius R is set to 1 and the number of neighbors N is set to 8, the PLCC reaches 0.8877 on the IRCCyN/IVC database and 0.8586 on the IETR database. Moreover, the feature extraction times of the two databases are 53 s and 132 s, respectively. Therefore, in combination with efficiency and accuracy, N is set to 8, and R is set to 1 in the proposed metric.

Table 1. Performance comparison of our designed algorithm with different LBP parameter settings.

Database	Radius	Neighbors	PLCC	SRCC	RMSE	Time (S)
IRCCyN/IVC	1	8	0.8877	0.8511	0.2264	53
IRCCyN/IVC	1.5	12	0.9593	0.9301	0.1703	87
IRCCyN/IVC	2	16	0.9780	0.9510	0.1184	125
IRCCyN/IVC	3	24	0.9760	0.9441	0.1235	6087
IETR	1	8	0.8349	0.8029	0.0925	132
IETR	1.5	12	0.7642	0.7477	0.088	195
IETR	2	16	0.7770	0.7277	0.0873	338
IETR	3	24	0.8586	0.8158	0.0824	8392

3.4. Compared with Existing DIBR View Synthesis Metrics

To verify the performance superiority of the proposed algorithm, we compare the proposed metric with nine existing DIBR view synthesis IQA methods, including MW-PSNR [20,21], MP-PSNR [22], LOGS [6], APT [19], OUT [23], NIQSV [24], GDSIC [2], MNSS [18], and CLGM [25]. First, we conducted the experiments on the IRCCyN/IVC database. In the experiments, 80% of the images were randomly chosen as training models, and the other 20% of the images were utilized for testing. The training tests were conducted 1000 times, and the median performance values are reported. Table 2 gives the experimental results, and the best results are shown in bold.

Second, we conduct comparison experiments on the IETR database. The experimental Settings are similar to the IRCCyN/IVC database. Table 3 gives the experimental results, and the best results are shown in bold.

Table 2. Performance comparison of our designed algorithm with the current mainstream DIBR IQA models on the IRCCyN/IVC database.

Model	Type	PLCC	SRCC	RMSE
MP-PSNR	FR	0.6174	0.6227	0.5238
MW-PSNR	FR	0.5622	0.5757	0.5506
LOGS	FR	0.8256	0.7812	0.3601
APT	NR	0.7307	0.7157	0.4546
OUT	NR	0.7678	0.7036	0.4266
NIQSV	NR	0.7114	0.6668	0.4679
GDSIC	NR	0.7867	0.7995	0.4000
MNSS	NR	0.7704	0.7854	0.4122
CLGM	NR	0.6750	0.6528	0.4620
Proposed	NR	**0.8877**	**0.8511**	**0.2264**

Table 3. Performance comparison of our designed algorithm with the current mainstream DIBR IQA models on the IETR database.

Model	Type	PLCC	SRCC	RMSE
MP-PSNR	FR	0.6190	0.5809	0.1947
MW-PSNR	FR	0.5389	0.4875	0.2088
LOGS	FR	0.6638	0.6679	0.1854
APT	NR	0.4225	0.4141	0.2252
OUT	NR	0.2409	0.2378	0.2406
NIQSV	NR	0.2095	0.2190	0.2429
GDSIC	NR	0.4338	0.4254	0.2244
MNSS	NR	0.2285	0.3387	0.2333
CLGM	NR	0.1146	0.0860	0.2463
Proposed	NR	**0.8349**	**0.8029**	**0.0925**

It can be found in Tables 2 and 3 that our designed metric obtains the highest PLCC and SRCC values and the lowest RMSE values on both databases. In contrast, none of the DIBR IQA metrics performed better than the proposed metrics; that is, the proposed approach is highly correlated with human visual perception of DIBR view distortion.

3.5. Generalization Ability Study

Generalization ability is vital for learning-based methods. This part tested the generalization ability of the proposed method using cross-validation. First, we trained the model in the IRCCyN/IVC database, and then the trained model was used for testing the IETR database. Second, the IETR database was used to train the model, and the trained model was used to test the IRCCyN/IVC database. The cross-validation simulation results are given in Tables 4 and 5.

Table 4. The proposed method is trained on the IRCCyN/IVC database, and the performance of our proposed method is tested on the IETR database.

Model	PLCC	SRCC	RMSE
Proposed	0.4901	0.4408	0.2031

Table 5. The Proposed Method is trained on the IETR database, and the performance of our proposed method is tested on the IRCCyN/IVC database.

Model	PLCC	SRCC	RMSE
Proposed	0.7489	0.7390	0.4139

Table 4 shows the cross-validation performance of our proposed metric on the IETR database. Because the distortions in the synthesized views of the IETR database do not include geometric data, the performance of the proposed method obtained PLCC, SRCC, and RMSE values on IETR of 0.4901, 0.4408, and 0.2031, respectively. Table 5 shows the cross-validation performance of the proposed algorithm on the IRCCyN/IVC database. The IETR database was adopted as the trained model, and the PLCC, SRCC, and RMSE values obtained by the proposed algorithm on IRCCyN/IVC are 0.7489, 0.7390, and 0.4139, respectively. The experimental results are very encouraging, which are still higher than the performance of other comparable algorithms. From the results, the proposed method has an excellent generalization ability.

4. Conclusions

A blind quality index for DIBR views with visual saliency and textural naturalness is put forward in this work. Our proposed algorithm is compared with the current mainstream DIBR view quality assessment methods on the IRCCyN/IVC and IETR databases. The experimental results show that our proposed blind quality measure has good performance on the two public DIBR synthesis view databases. It is worth mentioning that our proposed algorithm has good generalization ability, but the cross-database test results are not satisfactory. We will consider using the ability of dual network to extract contour and texture features for DIBR view synthesis to further improve the performance of this algorithm in future work [39].

Author Contributions: Conceptualization, L.T. and G.W.; methodology, L.T. and K.S.; software, G.W.; validation, K.S. and S.H.; formal analysis, S.H. and K.J.; investigation, G.W. and K.J.; resources, G.W., L.T. and K.S.; data collection, L.T. and K.S.; writing—original draft preparation, L.T. and G.W.; writing—review and editing, K.S. and K.J.; visualization, G.W. and K.S.; supervision, L.T.; project administration, L.T.; funding acquisition, L.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported in part by the Science and Technology Planning Project of Nantong, grant number JCZ20143, and the Jiangsu Qinglan Engineering Project.

Data Availability Statement: The experiment uses two public DIBR databases, including the IRCCyN/IVC and IETR databases.

Acknowledgments: The authors would like to thank the reviewers and the editor for their careful reviews and constructive suggestions to help us improve the quality of this paper.

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

References

1. Tian, S.; Zhang, L.; Zou, W.; Li, X.; Su, T.; Morin, L.; Deforges, O. Quality assessment of DIBR-synthesized views: An overview. *Neurocomputing* **2021**, *423*, 158–178. [CrossRef]
2. Wang, G.; Wang, Z.; Gu, K.; Li, L.; Xia, Z.; Wu, L. Blind Quality Metric of DIBR-Synthesized Images in the Discrete Wavelet Transform Domain. *IEEE Trans. Image Process.* **2020**, *29*, 1802–1814. [CrossRef] [PubMed]
3. PhiCong, H.; Perry, S.; Cheng, E.; HoangVan, X. Objective Quality Assessment Metrics for Light Field Image Based on Textural Features. *Electronics* **2022**, *11*, 759. [CrossRef]
4. Huang, H.Y.; Huang, S.Y. Fast Hole Filling for View Synthesis in Free Viewpoint Video. *Electronics* **2020**, *9*, 906. [CrossRef]
5. Zhou, Y.; Li, L.; Wang, S.; Wu, J.; Fang, Y.; Gao, X. No-Reference Quality Assessment for View Synthesis Using DoG-Based Edge Statistics and Texture Naturalness. *IEEE Trans. Image Process.* **2019**, *28*, 4566–4579. [CrossRef] [PubMed]
6. Li, L.; Zhou, Y.; Gu, K.; Lin, W.; Wang, S. Quality Assessment of DIBR-Synthesized Images by Measuring Local Geometric Distortions and Global Sharpness. *IEEE Trans. Multimed.* **2018**, *20*, 914–926. [CrossRef]
7. Gellert, A.; Brad, R. Image inpainting with Markov chains. *Sinal Image Video Process.* **2020**, *14*, 1335–1343. [CrossRef]
8. Cai, L.; Kim, T. Context-driven hybrid image inpainting. *IET Image Process.* **2015**, *9*, 866–873. [CrossRef]
9. Sun, K.; Tang, L.; Qian, J.; Wang, G.; Lou, C. A deep learning-based PM2.5 concentration estimator. *Displays* **2021**, *69*, 102072. [CrossRef]
10. Wang, G.; Shi, Q.; Wang, H.; Sun, K.; Lu, Y.; Di, K. Multi-modal image feature fusion-based PM2.5 concentration estimation. *Atmos. Pollut. Res.* **2022**, *13*, 101345. [CrossRef]

11. Sun, K.; Tang, L.; Huang, S.; Qian, J. A photo-based quality assessment model for the estimation of PM2.5 concentrations. *IET Image Process.* **2022**, *16*, 1008–1016. [CrossRef]
12. Gu, K.; Lin, W.; Zhai, G.; Yang, X.; Zhang, W.; Chen, C.W. No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization. *IEEE Trans. Cybern.* **2017**, *47*, 4559–4565. [CrossRef]
13. Gu, K.; Tao, D.; Qiao, J.F.; Lin, W. Learning a No-Reference Quality Assessment Model of Enhanced Images with Big Data. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 1301–1313. [CrossRef] [PubMed]
14. Gu, K.; Zhai, G.; Lin, W.; Yang, X.; Zhang, W. No-Reference Image Sharpness Assessment in Autoregressive Parameter Space. *IEEE Trans. Image Process.* **2015**, *24*, 3218–3231.
15. Li, L.; Lin, W.; Wang, X.; Yang, G.; Bahrami, K.; Kot, A.C. No-Reference Image Blur Assessment Based on Discrete Orthogonal Moments. *IEEE Trans. Cybern.* **2016**, *46*, 39–50. [CrossRef] [PubMed]
16. Okarma, K.; Lech, P.; Lukin, V.V. Combined Full-Reference Image Quality Metrics for Objective Assessment of Multiply Distorted Images. *Electronics* **2021**, *10*, 2256. [CrossRef]
17. Wang, G.; Wang, Z.; Gu, K.; Jiang, K.; He, Z. Reference-Free DIBR-Synthesized Video Quality Metric in Spatial and Temporal Domains. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 1119–1132. [CrossRef]
18. Gu, K.; Qiao, J.; Lee, S.; Liu, H.; Lin, W.; Le Callet, P. Multiscale Natural Scene Statistical Analysis for No-Reference Quality Evaluation of DIBR-Synthesized Views. *IEEE Trans. Broadcast.* **2020**, *66*, 127–139. [CrossRef]
19. Gu, K.; Jakhetiya, V.; Qiao, J.F.; Li, X.; Lin, W.; Thalmann, D. Model-Based Referenceless Quality Metric of 3D Synthesized Images Using Local Image Description. *IEEE Trans. Image Process.* **2018**, *27*, 394–405. [CrossRef]
20. Sandic Stankovic, D.; Kukolj, D.; Le Callet, P. DIBR synthesized image quality assessment based on morphological wavelets. In Proceedings of the 2015 Seventh International Workshop on Quality of Multimedia Experience (QOMEX), Messinia, Greece, 26–29 May 2015.
21. Sandic-Stankovic, D.; Kukolj, D.; Le Callet, P. DIBR-synthesized image quality assessment based on morphological multi-scale approach. *EURASIP J. Image Video Process.* **2016**, 4. [CrossRef]
22. Sandic-Stankovic, D.; Kukolj, D.; Le Callet, P. Multi-scale Synthesized View Assessment based on Moprhological Pyramids. *J. Electr.-Eng.-Elektrotechnicky Cas.* **2016**, *67*, 3–11.
23. Jakhetiya, V.; Gu, K.; Singhal, T.; Guntuku, S.C.; Xia, Z.; Lin, W. A Highly Efficient Blind Image Quality Assessment Metric of 3-D Synthesized Images Using Outlier Detection. *IEEE Trans. Ind. Inform.* **2019**, *15*, 4120–4128. [CrossRef]
24. Tian, S.; Zhang, L.; Morin, L.; Deforges, O. NIQSV: A No Reference Image Quality Assessment Metric for 3D Synthesized Views. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, New Orleans, LA, USA, 5–9 March 2017; pp. 1248–1252.
25. Yue, G.; Hou, C.; Gu, K.; Zhou, T.; Zhai, G. Combining Local and Global Measures for DIBR-Synthesized Image Quality Evaluation. *IEEE Trans. Image Process.* **2019**, *28*, 2075–2088. [CrossRef] [PubMed]
26. Zheng, H.; Zhong, X.; Huang, W.; Jiang, K.; Liu, W.; Wang, Z. Visible-Infrared Person Re-Identification: A Comprehensive Survey and a New Setting. *Electronics* **2022**, *11*, 454. [CrossRef]
27. Jiang, K.; Wang, Z.; Yi, P.; Chen, C.; Wang, G.; Han, Z.; Jiang, J.; Xiong, Z. Multi-Scale Hybrid Fusion Network for Single Image Deraining. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**. [CrossRef] [PubMed]
28. Jiang, K.; Wang, Z.; Yi, P.; Chen, C.; Wang, Z.; Wang, X.; Jiang, J.; Lin, C.W. Rain-Free and Residue Hand-in-Hand: A Progressive Coupled Network for Real-Time Image Deraining. *IEEE Trans. Image Process.* **2021**, *30*, 7404–7418. [CrossRef]
29. Wang, Z.; Jiang, J.; Wu, Y.; Ye, M.; Bai, X.; Satoh, S. Learning Sparse and Identity-Preserved Hidden Attributes for Person Re-Identification. *IEEE Trans. Image Process.* **2020**, *29*, 2013–2025. [CrossRef]
30. Wang, Z.; Jiang, J.; Yu, Y.; Satoh, S. Incremental Re-Identification by Cross-Direction and Cross-Ranking Adaption. *IEEE Trans. Multimed.* **2019**, *21*, 2376–2386. [CrossRef]
31. Varga, D. Full-Reference Image Quality Assessment Based on Grünwald–Letnikov Derivative, Image Gradients, and Visual Saliency. *Electronics* **2022**, *11*, 559. [CrossRef]
32. Jiang, K.; Wang, Z.; Yi, P.; Wang, G.; Gu, K.; Jiang, J. ATMFN: Adaptive-Threshold-Based Multi-Model Fusion Network for Compressed Face Hallucination. *IEEE Trans. Multimed.* **2020**, *22*, 2734–2747. [CrossRef]
33. Bosc, E.; Pepion, R.; Le Callet, P.; Koeppel, M.; Ndjiki-Nya, P.; Pressigout, M.; Morin, L. Towards a New Quality Metric for 3-D Synthesized View Assessment. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 1332–1343. [CrossRef]
34. Tian, S.; Zhang, L.; Morin, L.; Deforges, O. A Benchmark of DIBR Synthesized View Quality Assessment Metrics on a New Database for Immersive Media Applications. *IEEE Trans. Multimed.* **2019**, *21*, 1235–1247. [CrossRef]
35. Gu, K.; Wang, S.; Yang, H.; Lin, W.; Zhai, G.; Yang, X.; Zhang, W. Saliency-Guided Quality Assessment of Screen Content Images. *IEEE Trans. Multimed.* **2016**, *18*, 1098–1110. [CrossRef]
36. Li, Q.; Lin, W.; Fang, Y. No-Reference Quality Assessment for Multiply-Distorted Images in Gradient Domain. *IEEE Signal Process. Lett.* **2016**, *23*, 541–545. [CrossRef]
37. Fang, Y.; Lin, W.; Lee, B.S.; Lau, C.T.; Chen, Z.; Lin, C.W. Bottom-Up Saliency Detection Model Based on Human Visual Sensitivity and Amplitude Spectrum. *IEEE Trans. Multimed.* **2012**, *14*, 187–198. [CrossRef]
38. Scholkopf, B.; Smola, A.; Williamson, R.; Bartlett, P. New support vector algorithms. *Neural Comput.* **2000**, *12*, 1207–1245. [CrossRef]
39. Gu, K.; Xia, Z.; Qiao, J.; Lin, W. Deep Dual-Channel Neural Network for Image-Based Smoke Detection. *IEEE Trans. Multimed.* **2020**, *22*, 311–323. [CrossRef]