IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 11, 2019, accepted December 24, 2019, date of publication January 1, 2020, date of current version January 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2963498

Quality Evaluation of DIBR-Synthesized Images Based on Holes and Expanded Regions

YUE ZHAO, LAIHUA WANG, SUMIN QI, WEISHENG WANG, AND QING JIA
School of Software, Qufu Normal University, Shandong 273165, China

Corresponding author: Laihua Wang (wlh@tju.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant (No. 61601261).

ABSTRACT With the Depth-image-based rendering (DIBR) technique, virtual views can be drawn at any position between two reference viewpoints. However, DIBR algorithm may introduce various distortions in the synthesized images, such as holes, artifact, and stretching. This paper proposes a quality evaluation method for DIBR-synthesized images by analyzing the holes and expanded regions. First, the holes regions are obtained with the watershed algorithm and used to evaluate the quality of the DIBR-synthesized images. Then, the synthesized image is compared with the reference image in the sub-channel, and the number of different pixel values is used to assess the global quality of synthesized images. Finally, the two measures are pooled to calculate the overall quality score of the DIBR synthesized image. Experimental results show that our scheme can better predict the image quality and performs better in terms of implementation time compared with other schemes.

INDEX TERMS DIBR, synthesized image, quality evaluation, watershed algorithm, expanded regions.

I. INTRODUCTION

In recent years, stereoscopic multi-view video has developed rapidly and occupied an important position not only in 2D-to-3D video conversion system but also in free viewpoint TV (FTV) [1], [2]. Stereoscopic multi-view video gives the audience an immersive experience and allows viewers to experience the beauty of the world [3]. However, it needs a lot of space and bandwidth to store and transfer data respectively. In order to reduce the space and bandwidth, the depth-image-based-rendering (DIBR) technique was proposed, which used the depth map information and the original reference image to obtain the virtual viewpoint image. However, during the process of DIBR, a series of distortions are generated, such as holes, stretching, and artifact. Therefore, an effective evaluation method for DIBR-synthesized images is beneficial to improve the quality of the synthesized virtual image and it is also important for the development of 3D technology.

Generally, there are three quality evaluation methods: full-reference (FR), reduced-reference (RR) and no-reference (NR). The full reference compares the original image with the synthesized image and often uses the differential-mean-opinion-score (DMOS). However, in most cases, the investigator cannot obtain all the original image information, so the reduced-reference and no-reference quality evaluation are employed as the remedies. Wang and Simoncelli [4] proposed an image quality assessment method which used the Kullback-Leibler distance. Because wavelet transforms can't explicitly extract the image geometric information (e.g., lines and curves), Gao *et al.* [5] developed a new method to mimic the human visual system (HVS). Chandler and Hemami [6] presented an efficient metric for quantifying the visual fidelity of natural images based on near-threshold and suprathreshold properties of human vision. Liu *et al.* [7] gave a no-reference image quality assessment method using the curvelet transform. Serir *et al.* [8] used the blur effect on real images to measure the image quality. The noise estimation in the digital domain [9] and other quality assessment methods for certain distorted types had also achieved effective results. Hassen *et al.* [10] provided an image sharpness assessment based on local phase coherence. Gu *et al.* [11] used simple convolution operators to develop a new objective metric for research on perceptual quality assessment of distorted screen content images.

At present, there have been many methods for evaluating the quality of DIBR synthesized images. Li *et al.* [12] proposed a DIBR synthesized images quality evaluation based on local geometric distortion and global sharpness. In the

The associate editor coordinating the review of this manuscript and approving it for publication was Gangyi Jiang.

paper, they firstly used SIFT-flow-based warping to detect the disoccluded regions. Then, a re-blurring-based strategy was proposed to quantify the global sharpness. Finally, they pooled the two parts to get the final result. In [13], dictionary and image sharpness (SPARISH) index were used to represent sparse. Jung *et al.* [14] proposed a critical binocular asymmetry (CBA) metric designed for characterizing the binocular asymmetry property of human eyes. In the method, critical areas were first detected using the synthesized left-, right- view images and the corresponding disparity maps. The average SSIM scores of the critical areas in left- and right- view images were measured as the overall CBA score. Bokan *et al.* [15] used sparsity based features of morphologically to estimate the distorted level of the image and calculate the quality score by a general regression neural network. Farid *et al.* [16] used the original views from which the virtual image was generated to estimate the distortion induced by the DIBR process. In their method a block-based perceptual feature matching based on signal phase congruency metric was devised to estimate the synthesis distortion. Zhou *et al.* [17] gave a self-adaptive scale transform mode and detected the disoccluded regions by comparing the absolute difference between the preprocessed synthesized image and the warped image of the preprocessed reference image. Wang *et al.* [18] proposed a perceptual NR blur evaluation method using a new machine learning technique, i.e., extreme learning machine (ELM). Zhou *et al.* [19] used DoG-based edge statistics and texture naturalness to evaluate the image quality. Hou and Lin [20] trained gradient image characteristics and the grid characteristics by support vector regression (SVR) network to predict the stereoscopic image quality. Zhou *et al.* [21] used the random forest regression model to learn the quality model for the multiply distorted images.

These methods have laid the foundation for follow-up researches, but there are some disadvantages in practical applications. The combination of the left and right view images (method in [14]) needs to consider the accuracy and scientificity of mixing the left and right views. Some other methods belongs to local quality evaluation methods. However, the distortions in the DIBR-synthesized image are so complicated that the regions detection and feature extraction methods fail to extract some special local distorted regions. In some methods (such as those in [20] and [21]) because of lacking data sets, these algorithms either split the images into macro blocks to increase the data sets, or extracted features from the images and input them into a regression model to obtain a quality evaluation model. Although these methods improve the performance, the complexity of the algorithm increases because of splitting or features extracting.

Therefore, it is necessary to propose a method with a fast running speed and superior performance. Unlike the general distortions, the distortions of the DIBR-synthesized image are mostly generated in the synthesis process. The initial virtual view obtained by 3D warping has a lot of holes regions. Two representative distortions are shown in Fig. 1. Fig. 1(a) is the original reference image of Fig. 1(b) and Fig. 1(c).

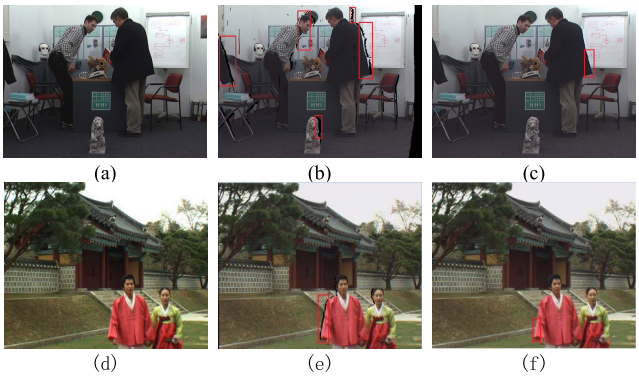


FIGURE 1. Illustrations of the synthesized distortions: (a) reference view of (b)-(d);(b) holes regions;(c) expanded regions;(d) reference view of (e)-(f);(e) holes regions;(f) expanded regions. The distorted regions are marked with red rectangles in the figure.

The distorted versions are shown in Fig. 1(b) and Fig. 1(c). There are a large number of holes regions in Fig. 1(b), and it can be seen that these regions are mainly distributed in the intersection of the foreground object and the background. In Fig. 1(c), there are many expanded regions at the junction of the men's clothes and the whiteboard on the right side. Fig. 1(d) is the original reference image of Fig. 1(e) and Fig. 1(f). There are a large number of holes regions in Fig. 1(e) and a lot of expanded regions in Fig. 1(f).

A DIBR synthetic image quality evaluation method is proposed in this paper based on the holes and expanded regions distortion mentioned above. Firstly, the holes are extracted and marked as the connected regions, and then the connected regions are used to evaluate the holes regions. Secondly, in order to evaluate the expanded regions, the synthesized image is compared with the reference image in the subchannel, and the number of different pixel values is used as the measurement standard. Finally, the two parts are pooled together to calculate the image quality.

The rest of this paper is organized as follows. In Section II, the proposed method is described in detail. In Section III, the experimental results of the IRCCyN/IVC DIBR image database are presented and compared with the current popular quality evaluation methods. In section IV, a summary is made of the future work.

II. PROPOSED METHOD

In this section, we introduce the proposed method in detail. In this paper, holes and expanded regions are used to evaluate the quality of the DIBR synthesized image. The framework of the proposed method is shown in Fig. 2. During the quality evaluation of the holes regions, in order to calculate the dispersed extent and the area proportion of the holes regions, we obtain the holes regions by the watershed algorithm [22]. The number and the area of the holes regions are obtained by extracting connected regions. For evaluating the expanded regions, we measure the difference between the synthesized image and the reference image in each channel, and we calculate the area of the expanded regions. Finally, we combine the

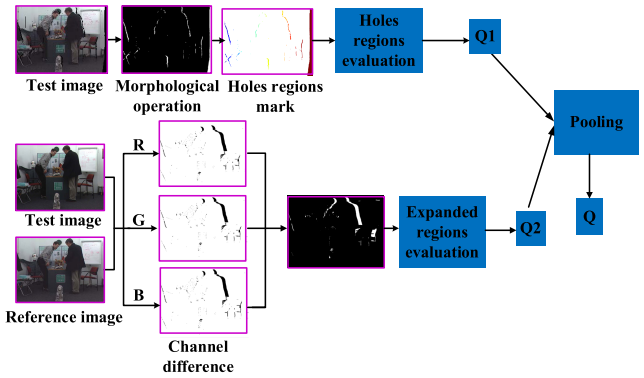


FIGURE 2. The framework of proposed method.

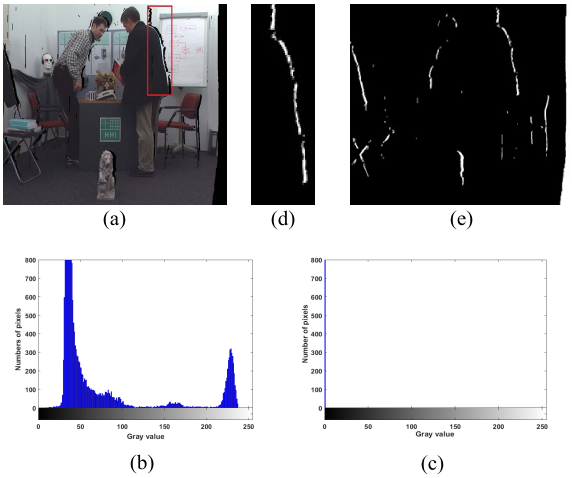


FIGURE 3. Holes regions analysis: (a) Synthesized virtual view with holes regions;(b) gray histogram of surrounding natural regions;(c) gray histogram of holes regions; (d) Extraction of holes regions marked by the red rectangle;(e) Extraction of holes regions of whole picture.

two parts and get the overall quality evaluation of the DIBR synthesized image.

A. QUALITY EVALUATION BASED ON HOLES REGIONS

The holes region can't be seen in the original reference image but is visible in the synthesized image. Fig. 3(a) shows a synthesized virtual image with holes regions. These holes regions seriously affect the quality of the synthesized virtual view. It is mentioned in [23] that all pixels of the holes regions have the same gray value. As is shown in Fig. 3(a), the natural regions in the distorted image are diversified in color, while the color change in the holes regions is not obvious, and the color is relatively simple. In Fig.3(b), the natural region in the red rectangle in Fig. 3(a) is analyzed by the gray histogram and it can be seen that the natural regions have a rich gray value. In Fig. 3(c), the holes regions in the red rectangle in Fig. 3(a) are analyzed by the gray histogram and it can be seen that the holes regions have single gray value. Based on the above analysis, the holes regions are regarded as a collection basin in geodesy, while other regions in the image are regarded as another collection basin. The steps of using watershed algorithm to extract the holes regions are as follows:

1) Extract the gradient image $g = \{g(x, y)\}$ of the distorted image $f = \{f(x, y)\}$:

$$g(x,y)=\{[f(x,y)-f(x-1,y)]^2+[f(x,y)-f(x,y-1)]^2\}^{1/2} \tag{1}$$

2) By simulating the immersion process, the regions with the same gray value in g are segmented to form a watershed at the boundary, and the calculation formula is as follows:

$$T[n] = \{(\mathrm{x}, y)|g(x, y) < n\} \tag{2}$$

where $T[n]$ is the coordinate set of the point in g below the plane $\mathrm{g}(x, y) = n$, and we empirically set $n = 1$.

3) Using $T[n]$ to operate binary g, the coordinates of the point below the plane $\mathrm{g}(x, y) = n$ are marked as the holes regions, and the remaining regions are marked as natural regions. The label map $D = \{D(x, y)\}$ of g is as follow:

$$D(x, y) = \begin{cases} 1, & g(\mathrm{x}, y) \in T[\mathrm{n}] \\ 0, & else \end{cases} \tag{3}$$

where $D(x, y) = 1$ indicates the pixel in holes regions, and $D(x, y) = 0$ indicates the pixel in natural regions.

4) However, over-segmentation is prone to occur in water-shed segmentation, so D is processed by the erosion operati-on:

$$H(x, y) = D(x, y)\Theta S \tag{4}$$

where $H = \{H(x, y)\}$ is the holes image, S is the struct-ural element, we select a circular structural element with a ra-dius of 1, and Θ is a symbol of erosion.

5) To eliminate the salt-and-pepper noise in H, the median filtering is employed:

$$h(x, y) = med\{H(x - u, y - l), (u, l \in W)\} \tag{5}$$

where $h = \{h(x, y)\}$ is the final holes regions map. $med\{\}$ represents median filtering. W is a median filter template which is composed of several pixels adjacent to it. u and l are the values in W.

The extraction of holes regions marked by the red rectangle is shown in Fig. 3(d), in which the white regions are the holes regions, and the black regions are the non-holes regions. The extraction result is shown in Fig. 3(e) in which the white regions are the holes regions, and the black regions are the non-holes regions. It can be shown in Fig 3(e) that the method proposed in this paper can effectively separate the holes regions from the natural regions.

After separating the holes regions from the natural regions, the next step is to consider how to use them to evaluate the image quality. In this paper, the area of the holes regions and the degree of dispersion of the holes regions are used as the measure of the holes quality. In order to calculate the area of the holes regions, the connected regions in h is first extracted:

$$Y_k = (Y_{k-1} \oplus \tau) \cap h, \quad \mathrm{k} = 1, 2, 3 \cdots \tag{6}$$

where τ is a rectangle structural element of 3×3. $\oplus$ is a symbol of dilation. Y_k represents the iteration result of the k^{th}. The condition for the iteration finishing is $Y_k = Y_{k-1}$. The larger the total area is, the worse the quality will be. The area proportion B of the holes regions to the entire picture is calculated:

$$B = \frac{\sum_{i=1}^{A_h} T_i}{M \cdot N} \tag{7}$$

where A_h is the total number of connection regions. M and N represent the length and width of the image, respectively. T_i represents the area of the i^{th} connected region.

In addition, the dispersion degree of the holes regions also affects the quality of the synthesized image. If the area of single holes region is too large in the synthesized image, it will directly affect the visual quality. When the holes regions with the same area are uniformly distributed in the synthesized image, the image quality perceived by human eyes will be better. Therefore, we also evaluate the quality of the image with the dispersion degree of the connected components. The discrete degree A of the holes regions are calculated with the standard deviation:

$$A = \sqrt{\frac{1}{A_h} \sum_{i=1}^{A_h} (T_i - \overline{T})^2} \tag{8}$$

where $\overline{T}$ represents the average area of all connected regions in the synthesized image:

$$\overline{T} = \frac{1}{A_h} \sum_{i=1}^{A_h} T_i \tag{9}$$

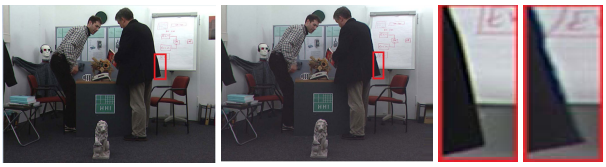
Combining the above two parts, the final holes regions score Q_1 is defined as:

$$Q_1 = \begin{cases} 1, & A_h = 0 \\ \dfrac{1}{\sqrt{\frac{1}{A_h} \sum_{i=1}^{A_h} (T_i - \overline{T})^2 + \frac{\sum_{i=1}^{A_h} T_i}{M \cdot N}}}, & others \end{cases} \tag{10}$$

If there is no hole in the image, Q_1 is 1, and the image quality is the best. When there are holes in the image, the larger the standard deviation and the image area of holes regions are, the worse the quality is.

B. QUALITY EVALUATION BASED ON EXPANDED REGIONS

In [24] the holes regions are filled with 'useful' color information during the process of the view synthesis. It will cause the expanded edge, and the outline of the object in the picture will change, as shown in Fig. 4. Compared with Fig. 4(b) the region marked by red rectangle in Fig. 4(a) is distorted. Fig. 4(c) and Fig. 4(d) are the enlarged versions of the red rectangle regions in Fig. 4(a) and Fig. 4(b) respectively. It can seen in Fig. 4(c) and Fig. 4(d) that the distance between the outline regions and the letter 'E' in the figure changes. Therefore, in this part, the step of the quality of the synthesized image with the expanded regions are as follows:



(a) (b) (c) (d)

FIGURE 4. Expanded regions analysis: (a) Synthesized virtual view of expanded regions;(b) Reference image;(c) enlarged region of (a) marked by the red rectangle;(d) enlarged image of (b) marked by the red rectangle.

1) To ensure that the images will have a higher contrast and larger dynamic range, we employ histogram specification to process the reference image $G = \{G(x, y)\}$ and the distorted image f:

$$j = (L-1) \sum_{e=0}^{U_G} \frac{d_e(G)}{\mathrm{M} \cdot N} \tag{11}$$

$$r = (L-1) \sum_{e=0}^{U_f} \frac{d_e(f)}{\mathrm{M} \cdot N} \tag{12}$$

where $j = \{j(x, y)\}$ is the processed image of G, $r = \{r(x, y)\}$ is the processed image of f. L is the gray level ($L = 256$) and e is the gray level. U_G is the total gray level of G and U_f is the total gray level of f. $d_e(G)$ is the number of e in G and $d_e(f)$ is the number of e in f.

2) After obtaining the j and the r, we can extract the difference map $p = \{p(x, y)\}$ of the expanded regions:

$$p(x, y) = |r(x, y) - j(x, y)| \tag{13}$$

3) To eliminate the noise in p, a threshold γ is used to obtain a binary image $q = \{q(x, y)\}$ so as to enhance the accuracy of searching the expanded regions. The formula is as follows:

$$q(x, y) = \begin{cases} 1, & p(x, y) > \gamma \\ 0, & else \end{cases} \tag{14}$$

4) In order to ensure the accuracy of the experimental results and avoid re-calculating the holes regions, we eliminate the holes regions in q with H by exclusive or:

$$z = q \odot H \tag{15}$$

where $\odot$ is the exclusive or, H is the extracted holes image, and $z = \{z(x, y)\}$ is the image after removing the holes. The result is shown in Fig. 5, in which the black regions are the natural regions and the white regions are the distorted regions.

After obtaining the distorted regions, we use them to evaluate the image quality. Expanded regions are compared with the corresponding regions in the original image. Therefore, based on the standard deviation formula and the PSNR algorithm, the quality assessment score of the expanded regions

FIGURE 5. Expanded regions extraction.

can be obtained by the following formula:

$$Q_2 = \log 10(\frac{M \cdot N}{\sqrt{\frac{1}{V}\sum_{i=1}^{V}(S_i - R_i)^2}}) \tag{16}$$

where V is the number of channels, S_i represents the area of the distorted region in the distorted image and R_i represents the area of the distorted region in original image. Since there is no distortion in the original image, the value of R_i is 0, so Q_2 can be further simplified as:

$$Q_2 = \log(\frac{M \cdot N}{\sqrt{\frac{1}{V}\sum_{i=1}^{V}(S_i)^2}}) \tag{17}$$

With Eq. (17), we can obtain that the larger the value of Q_2 is, the less the distorted regions in the distorted image are, and the better the quality of the image is.

C. POOLING

Based on the above evaluations of the holes and expanded regions, the two parts are combined to generate an overall score Q for the DIBR synthesized image:

$$Q = \alpha Q_1 + \beta Q_2 \tag{18}$$

where α and β are both balanced parameters, there are used to balance the contribution rate of Q_1 and Q_2. In this paper, we set $\alpha = 0.8312$, $\beta = 0.1688$.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the IRCCyN/IVC database is used to make an objective analysis of the proposed algorithm. Then some classical algorithms are selected to compare with the algorithm to verify its superiority.

A. EVALUATION PROTOCOL SELECTION

The IRCCyN/IVC DIBR synthesized image database has 96 images intotal, 84 of which are synthesized virtual views obtained through different DIBR synthesis algorithms [25]–[29], and the other 12 images are reference views. The database also gives the average subjective score for each distorted image. The seven synthesized algorithms are further elaborated in Table 1.

TABLE 1. Database algorithm description.

Algorithm	Name	Description
A1	Fehn cropped	The depth image is preprocessed to eliminate discontinuities, the border is cropped, and the image is resized to its original size [25].
A2	Fehn interpolated	The depth image is pre-processed to eliminate discontinuities, but the borders are not cropped [25].
A3	MPEG_VSRS	It uses the inpainting method to fill the disoccluded regions [26].
A4	Mueller	The disoccluded regions are filled with the help of depth information [27].
A5	ICME	It uses a patch-based texture synthesis to fill the missing part [28].
A6	ICIP TMM	It uses depth temporal information to improve synthesis [29].
A7	Holes	It contains unfilled holes.

B. METRICS

To measure the performance of an algorithm, the correlation is calculated between the score of the algorithm and the average subjective score given in the database. The higher the correlation is, the better the performance of the evaluation algorithm is. In this paper, SRCC, PLCC and RMSE are used as the evaluation parameters.

SRCC(the Spearman rank-order correlation coefficient) is defined as:

$$SRCC = 1 - \frac{6}{\lambda(\lambda^2 - 1)}\sum_{i=1}^{\lambda}(r_{xi} - r_{yi})^2 \tag{19}$$

where λ is the number of dataset data, r_{xi} and r_{yi} are the locations in data sets xi and yi.

PLCC (Pearson product-moment correlation coefficient) is indicated as:

$$PLCC = \frac{COV(X, Y)}{\delta X \delta Y} \tag{20}$$

where $COV(X, Y)$ is the covariance of the random variables X and Y, δX and δY are the product of the standard deviation of the random variables X and Y.

RMSE (Root Mean Square Error) is described as:

$$RMSE = \sqrt{\frac{1}{E}\sum_{i=1}^{E}(q_p^i - \overline{q_s})^2} \tag{21}$$

where E is the number of dataset data, q_p^i is the objecti-ve score of the i^{th} image, $\overline{q_s}$ is the mean of the subjective score.

After computing the objective scores, the nonlinear regression is used to map the scores to subjective ratings by a five-parameter logistic function:

$$F(x) = \delta_1 \cdot \left[\frac{1}{2} - \frac{1}{\exp(\delta_2 \cdot (x - \delta_3))}\right] + \delta_4 \cdot x + \delta_5 \tag{22}$$

TABLE 2. Performance comparison on the IRCCyN/IVC database.

Algorithm	Type	PLCC	SRCC	RMSE
SSIM[30]	2D	0.485	0.437	0.582
LTG[31]	2D	0.531	0.414	0.564
NIQE[32]	2D	0.437	0.374	0.599
ADD-SSIM[33]	2D	0.551	0.467	0.556
QAC[34]	2D	0.352	0.311	0.623
BIQI[35]	2D	0.565	0.582	0.523
IL-NIQE[36]	2D	0.499	0.535	0.577
DESQUE[37]	2D	0.473	0.398	0.529
MW-PSNR[38]	DIBR	0.562	0.576	0.551
SIQE[39]	DIBR	0.528	0.449	0.565
VSQA[40]	DIBR	0.574	0.523	0.545
MP-PSNR[41]	DIBR	0.617	0.623	0.524
Bosc[42]	DIBR	0.584	0.491	0.540
APT[43]	DIBR	0.731	0.716	0.455
Proposed	DIBR	**0.663**	**0.613**	**0.499**

TABLE 3. Average implementation time of a single image.

Algorithm	Time(s)
APT[43]	107.41
IL-NIQE[36]	4.018
NIQE[32]	0.496
Proposed	0.1522

where x is the objective score, δ_1, δ_2, δ_3, δ_4 and δ_5 are the coefficient of the regression function. $F(x)$ is the fitted scores. x is mapped to $F(x)$, and its magnitude range is similar to that of DMOS.

C. ANALYSIS OF RESULTS

In order to better verify the performance of the proposed algorithm in this paper, the following quality evaluation methods were selected for analysis and compare in the IRCCyN/IVC database: SSIM [30], LTG [31], NIQE [32], ADD-SSIM [33], QAC [34], BIQI [35], IL-NIQE [36], DESQUE [37], MW-PSNR [38], SIQE [39], VSQA [40], MP-PSNR [41], Bosc [42], APT [43]. These algorithms include both traditional 2D image quality evaluation methods and proprietary DIBR synthesized virtual view quality evaluation algorithms.

Table 2 shows the performance comparison on the IRCCyN/IVC Database. To be fair, all methods are implemented on the same test platform (MATLAB R2015b software executed on a 2.10 GHz processor with 96 GB RAM, Windows Server 2016 Datacenter Pro 64-bit desktop). For each method, tic and toc are used to record the total time on the entire IRCCyN/IVC image database. As is shown in Table 2, our algorithm has achieved good results compared with other methods, PLCC=0.663, SRCC=0.613, RMSE=0.499. The distorted types of DIBR synthesized images are different from traditional 2D images, so the performance of traditional 2D evaluation methods is worse than that of DIBR synthesized view quality evaluation algorithms.

TABLE 4. The implementation time of expanded and holes regions.

Metric / Algorithm	Expanded regions	Holes regions
A1	0.0970	0.0539
A2	0.0959	0.0537
A3	0.0968	0.0590
A4	0.0966	0.0544
A5	0.1021	0.0531
A6	0.0982	0.0538
A7	0.0958	0.0533

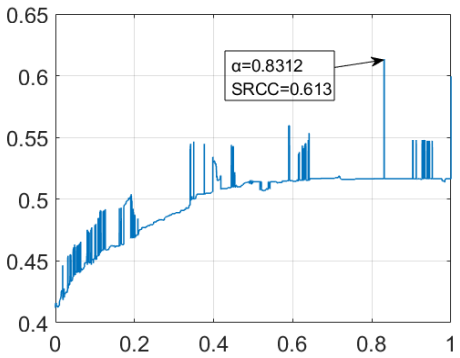


FIGURE 6. The impact of parameter (α used in Eq. (18)) on the performance of the SRCC.

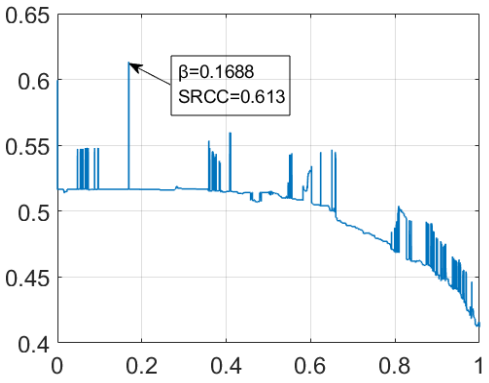


FIGURE 7. The impact of parameter (β used in Eq. (18)) on the performance of the SRCC.

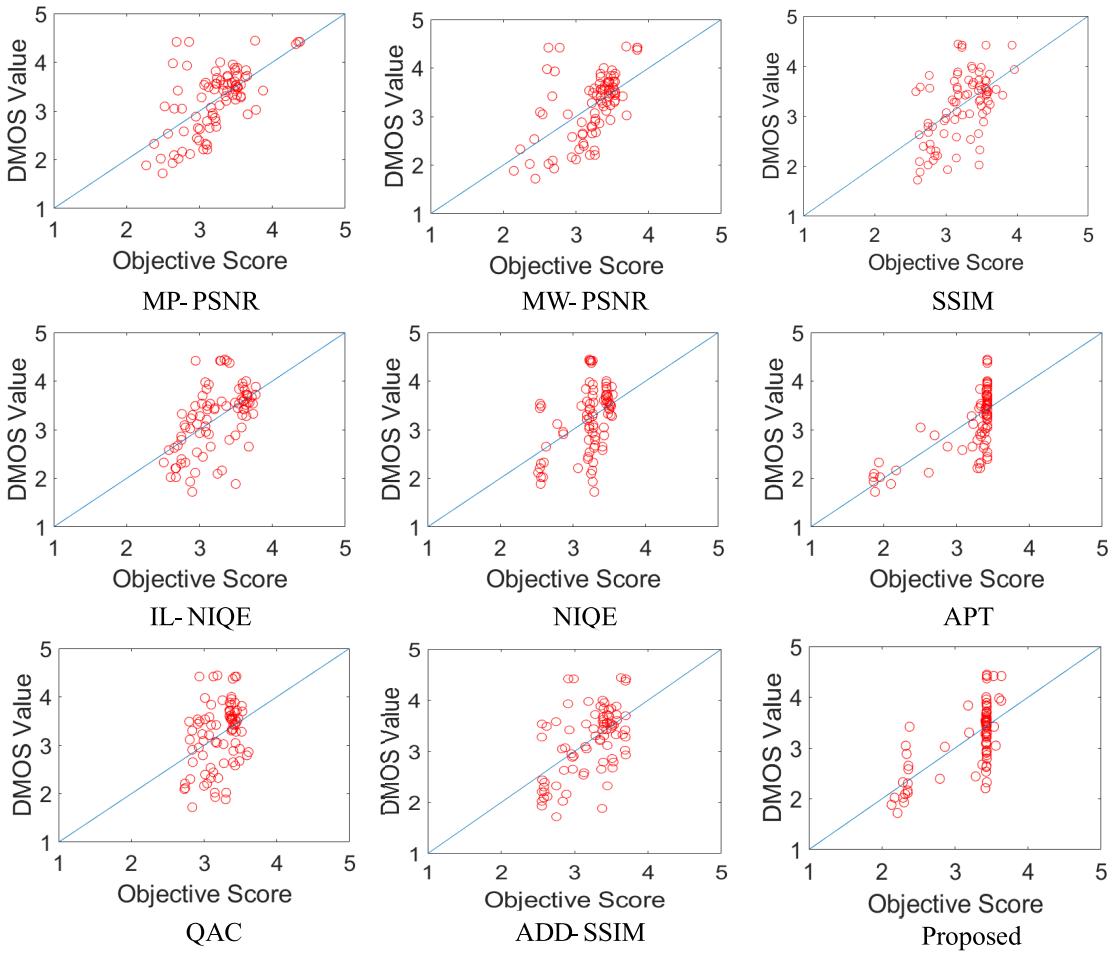


FIGURE 8. Scatter plots.

In Table 2, we compared our method with APT algorithm. It can be seen that the performance of APT algorithm is better than our method, but the algorithm is better than APT in complexity, which are proposed in Table 3.

In Table 3, we list the running time of the four algorithms. It is not difficult to find that the proposed method is better than IL-NIQE and NIQE in running time. Although the performance of the proposed method is slightly lower than that of APT in Table 2, our algorithm has a significant advantage in running time.

In Table 4, we listed the time for a single algorithm. It can be seen that most of the implementation times are less than 0.1s. Compared with the evaluation algorithms of DIBR synthesized images, our algorithm performs better in the implementation time and evaluates the expanded regions that have not been evaluated in other papers.

The holes and expanded regions are comprehensively considered and used to generate the final score. We add an ablation experiment for further study in Table 5, and present the graphs of SRCC parameter selection in Fig. 6 and Fig. 7. It is observed that the combined method in Fig. 6 and Fig. 7 achieves the best result when α and β are set as 0.8312 and 0.1688, respectively.

TABLE 5. Performance of each component considered.

Metrics	PLCC	SRCC	RMSE
Q_1	0.641	0.599	0.511
Q_2	0.354	0.412	0.623
Q	0.663	0.613	0.499

In Fig. 8, the scatter plots about the objective scores of some algorithms and DMOS values in database in MP-PSNR, MW-PSNR, SSIM, IL-NIQE, NIQE, APT, QAC and ADD-SSIM are presented. It can be seen that the proposed method has good consistency with the subjective score.

IV. CONCLUSION

In this paper, a DIBR synthesized image quality evaluation model is presented based on holes regions and expanded regions. The holes regions are the most classic problem in the DIBR synthesis process. After analyzing the visual illusion and the proportion of the distorted regions, we design a quality assessment model of holes regions. By comparing the reference image with the synthesized image, we extract the

expanded regions and use them to generate a quality evaluation score. Because of the expanded regions are difficult to observe, the evaluation score of the expanded regions isn't too high. After analysis and comparison, it is easy to find that although our algorithm has better performance than most existing methods except APT method, the following issues still need considering in future work:

1) During the quality evaluation of the holes regions, when the parameter settings are different, the stretching can be found, and the stretching is more comprehensive than the stretching extracted by the existing algorithms. So we plan to extract the stretching and use it to further improve the quality evaluation of the DIBR-synthesized images.

2) The evaluation of the expanded regions in this paper is based on the reference image. In the future, it is expected to find a non-referenced expanded regions quality evaluation algorithm.

REFERENCES

[1] G. Yue, C. Hou, K. Gu, T. Zhou, and G. Zhai, "Combining local and global measures for DIBR-synthesized image quality evaluation," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 2075–2088, Apr. 2019.

[2] L. Wang, C. Hou, S. Qi, and L. Jiang, "Analysis of maximum tolerant depth distortion in view synthesis," *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 7909–7927, Apr. 2018.

[3] G. Luo and Y. Zhu, "Hole filling for view synthesis using depth guided global optimization," *IEEE Access*, vol. 6, pp. 32874–32889, 2018.

[4] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," in *Proc. Hum. Vis. Electron. Imag. X*, Mar. 2005, pp. 149–159.

[5] X. Gao, W. Lu, D. Tao, and X. Li, "Image quality assessment based on multiscale geometric analysis," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1409–1423, Jul. 2009.

[6] D. Chandler and S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.

[7] L. Liu, H. Dong, H. Huang, and A. C. Bovik, "No-reference image quality assessment in curvelet domain," *Signal Process., Image Commun.*, vol. 29, no. 4, pp. 494–505, Apr. 2014.

[8] A. Serir, A. Beghdadi, and F. Kerouh, "No-reference blur image quality measure based on multiplicative multiresolution decomposition," *J. Vis. Commun. Image Represent.*, vol. 24, no. 7, pp. 911–925, Oct. 2013.

[9] O. Laligant, F. Truchetet, and E. Fauvet, "Noise estimation from digital step-model signal," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5158–5167, Dec. 2013.

[10] R. Hassen, Z. Wang, and M. M. A. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2798–2810, Jul. 2013.

[11] K. Gu, S. Wang, H. Yang, W. Lin, G. Zhai, X. Yang, and W. Zhang, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1098–1110, Jun. 2016.

[12] L. Li, Y. Zhou, K. Gu, W. Lin, and S. Wang, "Quality assessment of DIBR-synthesized images by measuring local geometric distortions and global sharpness," *IEEE Trans. Multimedia*, vol. 20, no. 4, pp. 914–926, Apr. 2018.

[13] L. Li, D. Wu, J. Wu, H. Li, W. Lin, and A. C. Kot, "Image sharpness assessment by sparse representation," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1085–1097, Jun. 2016.

[14] Y. J. Jung, H. G. Kim, and Y. M. Ro, "Critical binocular asymmetry measure for the perceptual quality assessment of synthesized stereo 3D images in view synthesis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 7, pp. 1201–1214, Jul. 2016.

[15] D. Bokan, G. Velikic, D. Kukolj, and D. Sandic-Stankovic, "Blind DIBR-synthesized image quality assessment based on sparsity features in morphological multiscale domain," in *Proc. 9th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, May 2017.

[16] M. S. Farid, M. Lucenteforte, and M. Grangetto, "Perceptual quality assessment of 3D synthesized images," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, Jul. 2017, pp. 505–510.

[17] Y. Zhou, L. Li, K. Gu, Y. Fang, and W. Lin, "Quality assessment of 3D synthesized images via disoccluded region discovery," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 1012–1016.

[18] S. Wang, C. Deng, B. Zhao, G.-B. Huang, and B. Wang, "Gradient-based no-reference image blur assessment using extreme learning machine," *Neurocomputing*, vol. 174, pp. 310–321, Jan. 2016.

[19] Y. Zhou, L. Li, S. Wang, J. Wu, Y. Fang, and X. Gao, "No-reference quality assessment for view synthesis using DoG-based edge statistics and texture naturalness," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4566–4579, Sep. 2019.

[20] C. Hou and H. Lin, "Stereoscopic image quality assessment based on grid and Weibull statistics," *J. Tianjin Univ. (Sci. Technol.)*, vol. 51, no. 10, pp. 1015–1022, Oct. 2018.

[21] Y. Zhou, L. Li, J. Wu, K. Gu, W. Dong, and G. Shi, "Blind quality index for multiply distorted images using biorder structure degradation and nonlocal statistics," *IEEE Trans. Multimedia*, vol. 20, no. 11, pp. 3019–3032, Nov. 2018.

[22] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, Jun. 1991.

[23] E. Bosc, "Towards a new quality metric for 3-D synthesized image assessment," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 7, pp. 1332–1343, Nov. 2011.

[24] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," *Proc. SPIE*, vol. 5291, pp. 93–104, May 2004.

[25] A. Telea, "An image inpainting technique based on the fast marching method," *J. Graph. Tools*, vol. 9, no. 1, pp. 23–34, Jan. 2004.

[26] Y. Mori, N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "View generation with 3D warping using depth information for FTV," *Signal Process., Image Commun.*, vol. 24, nos. 1–2, pp. 65–72, Jan. 2009.

[27] K. Müller, A. Smolic, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "View synthesis for advanced 3D video systems," *EURASIP J. Image Video Process.*, vol. 2008, Dec. 2008, Art. no. 438148.

[28] P. Ndj, M. Koppel, D. Doshkov, H. Lakshman, P. Merkle, K. Müller, and T. Wiegand, "Depth image based rendering with advanced texture synthesis," in *Proc. IEEE Int. Conf. Multimedia Expo.*, Jul. 2010, pp. 424–429.

[29] M. Koppel, P. Ndjiki-Nya, D. Doshkov, H. Lakshman, P. Merkle, K. Müller, and T. Wiegand, "Temporally consistent handling of disocclusions with texture synthesis for depth-image-based rendering," in *Proc. IEEE Int. Conf. Image Process.*, Hong Kong, Sep. 2010, pp. 1809–1812.

[30] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[31] K. Gu, G. Zhai, X. Yang, and W. Zhang, "An efficient color image quality metric with local-tuned-global model," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 506–510.

[32] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

[33] K. Gu, S. Wang, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Analysis of distortion distribution for pooling in image quality prediction," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 446–456, Jun. 2016.

[34] W. Xue, L. Zhang, and X. Mou, "Learning without human scores for blind image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 995–1002.

[35] A. Moorthy and A. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, May 2010.

[36] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.

[37] Y. Zhang and D. M. Chandler, "No-reference image quality assessment based on log-derivative statistics of natural scenes," *J. Electron. Imag.*, vol. 22, no. 4, Dec. 2013, Art. no. 043025.

[38] D. Sandic-Stankovic, D. Kukolj, and P. Le Callet, "DIBR synthesized image quality assessment based on morphological wavelets," in *Proc. 7th Int. Workshop Qual. Multimedia Exper. (QoMEX)*, May 2015, pp. 1–6.

[39] M. S. Farid, M. Lucenteforte, and M. Grangetto, "Objective quality metric for 3D virtual views," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 3720–3724.

[40] P. H. Conze, P. Robert, and L. Morin, "Objective view synthesis quality assessment," *Proc. SPIE, Electron. Imag. Int. Soc. Opt. Photon.*, vol. 2888, Feb. 2012, Art. no. 82881M.
[41] D. Sandić-Stanković, D. Kukolj, and P. Le Callet, "DIBR-synthesized image quality assessment based on morphological multi-scale approach," *EURASIP J. Image Video Process.*, vol. 2017, no. 1, p. 4, 2016.
[42] E. Bosc, R. Pepion, P. Le Callet, M. Koppel, P. Ndjiki-Nya, M. Pressigout, and L. Morin, "Towards a new quality metric for 3-D synthesized view assessment," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 7, pp. 1332–1343, Nov. 2011.
[43] K. Gu, V. Jakhetiya, J.-F. Qiao, X. Li, W. Lin, and D. Thalmann, "Model-based referenceless quality metric of 3D synthesized images using local image description," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 394–405, Jan. 2018.

YUE ZHAO received the B.S. degree from Qufu Normal University, Shandong, China, in 2018, where she is currently pursuing the M.S. degree. Her current research interests include quality assessment and machine learning.

LAIHUA WANG received the Ph.D. degree in electronic engineering from Tianjin University, Tianjin, China, in 2016. She is currently a Lecturer with the School of Software Engineering, Qufu Normal University, China. Her research interests include computer vision, digital image processing, 3D video processing, and 3D display.

SUMIN QI received the Ph.D. degree in technology of computer application from Soochow University, Suzhou, China, in 2008. She is currently an Associate Professor with the School of Software Engineering, Qufu Normal University, China. Her research interests include computer vision and digital image processing.

WEISHENG WANG received the B.S. degree from Liaocheng University, Shandong, China, in 2019. He is currently pursuing the M.S. degree with Qufu Normal University. His current research includes machine learning and related directions.

QING JIA is currently pursuing the B.S. degree with Qufu Normal University, Shandong, China. Her current research interests include software engineering and software development.