

# Multimedia VR Image Improvement and Simulation Analysis Based on Visual VR Restructuring Algorithm

Xiangyang Xu

Henan Police College, Zhengzhou, Henan, 450046, China

E-mail: xxy@hnp.edu.cn

**Keywords:** visual VR reconstruction algorithm, image optimization, virtual reality technology, median filter

**Received:** October 25, 2023

*Due to the advancement of science and technology, the application of virtual reality (VR) technology is more and more extensive, and people can truly immerse themselves in the virtual space through virtual reality. Relying on the visual VR reconstruction algorithm, this paper deals with the problems of "burring" and insufficient compression of relatively simple video imaging devices. Using virtual reality as a foundation, the multimedia effect of the video image is processed, and according to six operation modules, a system combining virtual reality technology is designed. From the aspect of determining the relationship between video image data and color, it is classified into three types: binary image, pseudo-color image, and grayscale image, and the grid of each point is defined and quantified. The extreme value filtering algorithm is used to perform a sorting calculation on the image pixels in the filtering window so as to improve the image effect with the threshold value suitable for filtering processing. Simulation results show that the VR visual restoration algorithm has a higher compression ratio and higher optical efficiency and can effectively support multimedia VR image improvement and simulation analysis.*

*Povzetek: Študija se ukvarja z analizo multimedijskih VR slik z uporabo algoritma za vizualno rekonstrukcijo VR, ki naslavlja težave z zamegljenostjo in nezadostno kompresijo slik, povečuje kompresijsko razmerje in optično učinkovitost za podporo izboljšanju VR slik.*

## 1 Introduction

Under the influence of rapid social and economic development, people's pursuit of virtual space is getting higher and higher [1-4]. On the one hand, the immersion type makes users be in a virtual environment; that is, the physical environment is no longer felt but integrated into a new virtual environment; on the other hand, the interactive type enables users to control the surrounding environment in real-time, making people interactive with the surrounding virtual environment and be creative. Therefore, in the environment of virtual reality, it is first necessary to solve the construction of the virtual environment, that is, how to construct it. At the same time, the perfection of the virtual 3D space is the first sign of the user's experience [5-12].

The related technology of virtual reality is based on the human-centered perspective, viewing the video from 360° to the surroundings, which is no longer limited by time and space, so that people can fully experience the real virtual space. Different from traditional video images, the resolution of virtual reality video images is higher, and the bandwidth resources and storage resources occupied are higher. Therefore, its transmission, storage, anti-interference, and other capabilities are poor, and complex digital processing cannot be conducted [13-15]. How to effectively process multimedia VR images, this paper relies on the visual VR reconstruction algorithm; facing the problem of processing the related effects of virtual and real video graphics images, we can use filtering to calculate the threshold, which can effectively remove the noise in the video image, to explore the improvement and simulation analysis of multimedia VR images.

## 2 Related works

Table 1: Literature survey

Reference	Key Findings	Methodologies	Outcomes
[16]	The paper examined the main findings of designing interactive VR classrooms, with a specific focus on categorizing educational activities.	They investigated the incorporation of deep learning algorithms and utilized a quantitative regression analysis methodology.	They demonstrated the impact of assessing teaching quality on improving the learning experience.

[17]	The study examined significant discoveries concerning occlusion in hand posture estimation, emphasizing the use of the Skeleton-Difference Loss Function and the Object-Manipulating Loss Function.	The research utilized approaches that specifically targeted the training of deep learning models.	The experimental results indicated the flexibility and exceptional efficiency of the suggested system across many circumstances.
[18]	The article examined the advancement of visual effects in landscaping graphics, emphasizing the application of deep belief networks as classifiers.	The article employed a three-fold cross-validation methodology, a deep belief network learning process, a wavelet deep belief network model, and a weighted k-nearest neighbor algorithm.	The results indicated an improvement in recognition accuracy and classification effectiveness, suggesting possible applications in garden image recognition technologies.
[19]	The paper presented significant findings on the development of virtual interactive models, with a focus on improving the user experience.	The article utilized approaches focused on the integration of technology to research cultural heritage.	The research revealed advancements in user experience, cultural investigation, and conservation.
[20]	The paper included deep learning integration, multidimensional assistance, and the broad commercial use of a particular system or technology.	The paper examined the approaches associated with the use of supporting technologies, with a specific emphasis on the impact of deep learning.	The paper examined the results of widespread commercial applications, the growing popularity of VR products, and the incorporation of deep learning theory.
[21]	The paper examined significant discoveries on the broader uses of VR in architecture design, with a particular focus on advanced rendering techniques.	The research examined techniques that relate to Immersive Rendering and Deep Learning Training, explicitly emphasizing the Camera Velocity Rendering Method.	The study investigated the practicality and efficacy of enhanced animation routes in mitigating VR sickness, demonstrating improved results in the domain of VR.
[22]	The integration of VR technology for complete quality improvement was presented.	The methods for building digitization and VR technology applications in high-end construction projects were examined in this study.	The consequence of the article focused on "Enhanced Designer-User Interaction" and offered recommendations for the advancement of the sector.
[23]	The study examined significant discoveries concerning the emotional influence of video games, highlighted the advancements in design approaches, and discussed the challenges encountered in VR.	The paper described the techniques used to develop emotionally intelligent virtual avatars, with a specific emphasis on the implementation of emotional avatars from Bernardo Agents.	The article demonstrated enhanced narrative perception, a favorable influence on presence, and versatility in its application to different virtual worlds.
[24]	The research presented significant discoveries about image transformation technology, specifically focusing on the construction and equalization of Grey Level Histograms.	The study examined approaches connected with modeling and 3D Technology in VR, with a specific emphasis on Image Transformation Technology.	The paper's result involved addressing interface difficulties and improving visual effects that were in line with human features.
[25]	The study examined essential findings regarding the incorporation and	The paper explored techniques for reducing	The research exhibited improved detection performance achieved by

	interaction of multimedia data, with a specific focus on resolving interference problems in multimedia networks.	interference in multimedia networks by utilizing compressed coding and decoding technology.	effectively implementing compression technology, supported by positive testing outcomes.
--	--	---	--

The need for standardized tools and workflows for creating high-quality VR content, along with difficulties in capturing and producing content that maximizes VR capabilities, contribute to challenges such as low resolution leading to pixelation. To address these issues, we propose the implementation of the Visual VR Restructuring Algorithm.

### Dataset

Everyday Objects in Context database [26]: The COCO Dataset, an extensive image collection, consists of over 20,000 carefully annotated images across 81 distinct image types. This dataset was utilized for both training as well as assessment purposes. This step can be regarded as a top-down data training stage, in which user-labeled information is used for supervised salient recognition of objects. Furthermore, our issue is due to the need for more knowledge regarding the specific item or object class. Therefore, we must establish a general conspicuous object by relying on global properties. Figure 1 depicts the Dataset sample images.



Figure 1: Dataset sample images

## 3 Research methods

### 3.1 Virtual reality technology

VR technology is composed of modules such as feedback, detection, sensors, controllers, modeling, etc. The specific composition is depicted in Figure 2.

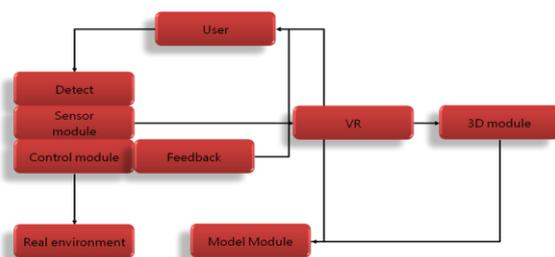


Figure 2: System composition of virtual reality technology

Among these six modules, which are entirely different but related, the sensor module is linked to the user by the detection and feedback modules. It communicates with the 3D module via the control module.

### 3.2 Visual VR reconstruction algorithm

3D Stereo Matching Correlated with Phase: If two images with a size of  $m$  in  $E(m)$  and  $H(m)$  sequences have the same time, the calculation of the discrete Fourier function is shown in formula (1) and formula (2):

$$E(l) = \sum_{m=-N}^N e(m)X_M^{lm} = B_E(l)f^{i\theta_E(l)}, \quad (1)$$

$$H(l) = \sum_{m=-N}^N h(m)X_M^{lm} = B_H(l)f^{i\theta_H(l)}, \quad (2)$$

where  $m = -N, \dots, N$ ,  $M = 2N + 1$ ,  $X_M = f^{-\frac{l2m}{M}}$ ,  $B_E(l)$  &  $B_H(l)$  are the “image” amplitude data of the description; the two values  $f^{i\theta_E(l)}$  &  $f^{i\theta_H(l)}$  denote the phase areas of the  $u$ , then it can be obtained that the calculation of the image phase difference after normalization processing is as shown in the formula (3):

$$\hat{Q}(l) = \frac{E(l)\overline{H(l)}}{|E(l)H(l)|} = f^{i\theta(l)} \quad (3)$$

Among them is a conjugate complicated number description  $\theta(l) = \theta_e(l) - \theta_H(l)$ . The inverse transformation of the discrete Fourier function can be obtained by  $Q$ , as shown in formula (4):

$$\hat{q}(m) = \frac{1}{M} \sum_{l=-N}^N \hat{Q}(l)X_M^{-lm} \quad (4)$$

Phase correlation visual VR reconstruction under the averaging method: When images of low quality are used, the accuracy of the corresponding 3D stereo adaptation method is biased [7-9]. At this time, the averaging method is used to improve the accuracy of binocular vision

adaptation, utilizing the particular procedure illustrated in Figure 3.

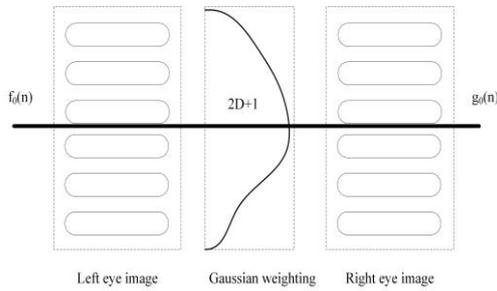


Figure 3: Visual VR 3D matching under phase correlation

As described in Figure 3 above, the 3D VR matching of visual images based on phase correlation, we can conclude that the human eyes are symmetrical. The lateral configuration can be applied to the optical system. The obtained sequence of sample images  $f(m)$  of the numerical value is shifted by  $L$  distance units along the  $y$ -axis to get  $f_e(m)$ . Subsequently, a line  $h_e(m)$  corresponding to this value is calculated, and a specific phase-related numerical processing is performed on  $f_e(m)$  and  $h_e(m)$   $r_{e(m)}$  obtained so that the final mean sequence value is shown in formula (5):

$$\hat{q}_{ave}(m) = \frac{1}{2c+1} \sum_{k=-c}^{k=c} \hat{q}_k(m) \quad (5)$$

3D Matching Structure of CTF:  $e_s$  represents the binocular image, displayed on the idea of the first level, in formula (6), it corresponds to the average value of the pixels illustrated in the four relatively close areas of the previous level  $e_{s-1}$ , as shown in formula (6):

$$e_{s-1}(w, z) = \frac{1}{4} \sum_{j=0}^1 \sum_{i=0}^1 e_{s-1}(2w + j, 2z + i) \quad (6)$$

After the denoising method,  $J_w^i$  the grayscale arrangement of the image in the visual VR remodeling is made. To obtain the dispersion of the characteristics of the surrounding area of the target individual,  $w=p,e$ , and the grayscale transformation formula is shown in Equation (7), Equation (8):

$$T_d = [T_0, T_1, \dots, T_{R-1}]_{binary} = [\sum_j^{R-1} T_j \times 3^j]_{dce} \quad (7)$$

$$T_j = \sum_i^{X \times X} J_w^i \quad (8)$$

It can be seen that  $W$  is the transformation step size,  $Y$  is the gray value of the area near the target range, and the

denoising function is constructed as shown in formula (9, 10):

$$w(l + 1) = T_j(l)_w + x_j(l), j = 1, 2, \dots, n, \quad (9)$$

$$y(l) = G_j(l)_w(l) + u_j(l), j = 1, 2, \dots, n \quad (10)$$

The above formula (9)  $u_j(l)x_j(l)$  represents the pixel noise value of different target individuals in the range area, and the mean value of  $G_j(l)_w$  and  $T_j(l)_w$  is 0, and the variance represents the state that this value  $T_j(l)$  is uniformly distributed on the image and constitutes a fuzzy set  $v = \{v_{jk}\}$ . After that, the noise-removed output is obtained, and the texture characteristics are analyzed, as shown in formula (9, 10):

$$J_{GSM} = J(D^M; C^M | t^M) = \sum_{j=1}^M J(D_j; C_j | t_j) = \sum_{j=1}^M (g(C_j | t_j) - g(C_j | D_j; t_j)) = \sum_{j=1}^M (g(h_j; D_j + U_j | t_j) - g(U_j)) \quad (11)$$

Considering the difference between the reconstructed image ranges, the pixel grayscale range is formed in the direction of the gradient to obtain the alternative formula for image denoising by describing the spatial texture characteristics  $v^{(m)}(w, z, c)$ , as shown in formula (12) and formula (13):

$$v^{(m+1)}(w, z) = v^{(m)}(w, z) + \delta v_1^{(m)}(w, z). \quad (12)$$

$$v_1^{(m)}(w, z) = N \Delta_w v^{(m)}(w, z) + M \Delta_z v^{(m)}(w, z, c) \quad (13)$$

Then, formula (13) is applied to the “original image for high-pass filtering.” After processing, the texture of the “image” is enhanced again, as shown in formula (14):

$$d(w, z) = \sum_{j=-m}^m \sum_{i=-n}^n g(j, i) h(w - j, z - i) \quad (14)$$

Among them,  $d$  is the high-frequency range selected by the feature,  $h$  denotes the image reconstructed, and  $g$  represents a  $3 \times 3$  rectangular matrix based on a high-pass filter, as shown in Equation (15):

$$h = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (15)$$

After that, an image  $e_1$  with enhanced texture quality can be obtained according to equation (16):

$$e_1(w, z) = e(w, z) + d(w, z) \quad (16)$$

Where  $(w, z)$  denotes the coordinates for the direction of the image,  $e_1$  denotes the pattern image after details are enhanced,  $e$  denotes the original image, and  $d$  denotes the high-frequency range selected.

### 3.3 Modification of multimedia vision vr reconstruction algorithm

#### (1) Algorithm peak search

In order to reduce the maximum peak range, recalculate the link representing the strengthening of low-pass filtering  $V_1, V_2$  and calculate a peak value, as shown in formula (17):

$$g(l_1, l_2) = \begin{cases} 1, & |l_1| \leq V_1, |l_2| \leq V_2 \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

#### (2) Error Identification and peak relocation

The peak  $\alpha$  in the random construction level is collected and sorted into a set. If the edge of the limited range of level 1 is the threshold  $\alpha_{th}$ , then  $r_k(n_k)$  it can be calculated. To confirm that this point is a cluster point, the phase correlation peak  $\alpha$  must be  $\geq$  the threshold  $\alpha_{th}$ , as shown in equations (18) and (19):

$$r_{kd}(n_k) = r_k(n_k), \quad (18)$$

$$c_{kd}(n_k) = n_k - r_k(n_k) \quad (19)$$

In order to confirm that this point is an outlier,  $\alpha < \alpha_{th}$  it is assumed that the arrangement is arranged according to the size of the middle value taken. All issues between the  $5 \times 5$  area in the near range excluding this point are selected, and the  $c$  value of this point is defined; see Formulas (20) and (21) are shown as:

$$c'_1(n_k) = (c_1^{med}, c_2^{med}), \quad (20)$$

$$q'_k(n_k) = n_k - c'_k(n_k) \quad (21)$$

Secondly, a new peak value is obtained by using the phase correlation calculation, assuming  $\alpha > \alpha_{th}$ , as shown in formula (22) and formula (23):

$$r_{kd}(n_k) = r'_k(n_k) \quad (22)$$

$$c_{kd}(n_k) = n_k - c'_k(n_k) \quad (23)$$

On the contrary,  $c_{kd}(n_k) = c'_k(n_k)$  it is made

#### (3) Algorithm correction

Let the relative reconstruction point be  $(w_r, z_r)$ , the sample direction has the coordinates  $(x_l, y_l)$ , the calculated peak value is  $\alpha$ , and the visual aberration of the two eyes is  $c = y_l - y_r$ .

If  $C_{min} < C < C_{max}$  and  $\alpha \geq \alpha_{th}$  are entered into the stereo model, where  $C_{max}$  and  $C_{min}$  are the maximum and minimum parallax values in sequence?

### 3.4 Video image digitization

In this design, video image digitization mainly uses the quantization method. And make corresponding assumptions, use equidistant sampling to obtain a nearly coherent image  $e(w, z)$ , and set it as a rectangular array of  $N \times M$  so the following formula (24) can be obtained:

$$e(w, z) = \begin{bmatrix} e(0,0) & e(1,0) & \dots & e(0,N-1) \\ e(1,0) & e(1,1) & \dots & e(1,N-1) \\ \dots & \dots & \dots & \dots \\ e(M-1,0) & e(M-1,1) & \dots & e(M-1,N-1) \end{bmatrix} \quad (24)$$

Every element is an independent discrete variable; the right side of formula (4) shows a video image of a number, and each element in the data set is described as a corresponding pixel [10].

In the actual calculation process, in order to make  $Z$  and  $r$  an array of real numbers and integers, during the collection of samples, they are converted into network format as the flattened data of the image. The grid of each node is calculated and finally determined according to the Cartesian coordinate system.

During the actual calculation process,  $Y$  and  $R$  are transformed into arrays of real numbers and integers. This conversion occurs while collecting samples, where they are represented in network format as the flattened data of the image. Subsequently, the grid of each node is computed and ultimately determined based on the Cartesian coordinate system

During the whole process of digitally converting the idea, first, we determine the N and M dimensions of the image and the distinct grayscale values. H in pixels. When calculated, these values are usually rounded to an integer power of 2. Therefore, the expression tested in the snapshot can be expressed by Equation (25):

$$\begin{cases} N = 2^n, \\ M = 2^m, \\ H = 2^l \end{cases} \quad (25)$$

Suppose the value range of the discrete grayscale is set to be between 0 and 10, and the distribution is in a uniform state. In that case, the bits required to store the digital video image can be retrieved by the following formula, as shown in formula (26):

$$b = M \times N \times K \quad (26)$$

If M=N, then it can be shown as formula (27):

$$b = N^2 k \quad (27)$$

### 3.5 Video image processing effect optimization

Analysis of one-way multi-stage median filter algorithm: The median filter should replace the median value of the area between two adjacent points with a point in the numerical interval and finally calculate the median value using the following method [11-12]:

Let  $w_1, w_2, \dots, w_n$  be a set of m values, which are sorted according to their size, and get the following formula (28):

$$w_{j1} \leq w_{j2} \leq w_{j3} \leq \dots \leq w_m \quad (28)$$

Among them, when the odd number is m, the following formula (29) is obtained:

$$z = \text{Med} (w_1, w_2, w_3, \dots, w_n) = w_j \frac{(m+1)}{2} \quad (29)$$

When the even number is m, see the following formula (30):

$$z = \text{Me} (w_1, w_2, w_3, \dots, w_m) = \frac{1}{2} \left[ w_j \frac{(m+1)}{2} + w_j \frac{(m)}{3} + 1 \right] \quad (30)$$

From the calculation results of the above formulas, it can be found that the result of the effect of the image is similar to the simple root mean square in the results obtained from the 3\*3 window, and y represents the median value of the sequence.

During the process of this calculation, let  $w(n, m)$  be a complex video image in the original noise and represented as a grayscale point value of  $(n, m)$  in the coordinates. First, select the rectangle where the window  $L=2m+1$  is located and divide the window into four independent windows, in which integer m is positive. The process is described as follows, see calculation equations (31)-(34):

$$X_1(n, m) = \{w(n, m + j), -M \leq j \leq M\} \quad (31)$$

$$X_2(n, m) = \{w(n + j, m), -M \leq j \leq M\} \quad (32)$$

$$X_3(n, m) = \{w(n + j, m - j), -M \leq j \leq M\} \quad (33)$$

$$X_4(n, m) = \{w(n + j, m - j), -M \leq j \leq M\} \quad (34)$$

According to the calculation, the schematic diagram of the MLM filter in Figure 4 can be obtained:

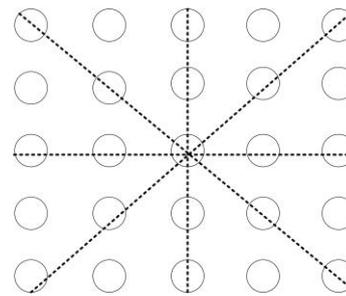


Figure 4: MLM filter

A set of twisted one-dimensional image windows along the horizontal or vertical direction is represented by  $X_1, X_2, X_3, X_4$  respectively, and then  $Y_1(n, m), Y_2(n, m), Y_3(n, m)$  and  $Y_4(n, m)$  denote the median of  $X_1, X_2, X_3, X_4$  (4 windows), as shown in equations (35)-(38):

$$Y_1(n, m) = \text{med} [w(j, i) \in X_1(n, m)], \quad (35)$$

$$Y_2(n, m) = \text{med} [w(j, i) \in X_2(n, m)], \quad (36)$$

$$Y_3(n, m) = \text{med} [w(j, i) \in X_3(n, m)], \quad (37)$$

$$Y_4(n, m) = \text{med} [w(j, i) \in X_4(n, m)] \quad (38)$$

Where  $V_{min}(n, m), V_{max}(n, m)$  and denote the max and min of  $Y_1(n, m), Y_2(n, m), Y_3(n, m)$ , and  $Y_4(n, m)$  (mean). Thus, equations (39) and (40) can be obtained:

$$V_{max}(n, m) = \max[Y_1(n, m), Y_2(n, m), Y_3(n, m), Y_4(n, m)], \tag{39}$$

$$V_{min}(n, m) = \min[Y_1(n, m), Y_2(n, m), Y_3(n, m), Y_4(n, m)], \tag{40}$$

Based on the above formula, the derivation method of the multi-stage median filter of the single term is shown in formula (41):

$$z(n, m) = med [V_{min}(n, m), V_{max}(n, m), w(n, m)] \tag{41}$$

Image optimization by extreme median filtering based on thresholds: In this paper, points are introduced to establish median extremum filtering and enhance the arrangement of image pixels in the window used for filtering. The image range is pre-determined and segmented into the fine details of the image edges, noise influence, and flat range optimization [13]:

First, arrange the pixel points  $W[x_{j,i}]$  in the window to find the  $\min(W[x_{j,i}])$  point and the  $\max(W[x_{j,i}])$  point; this point  $\max(W[x_{j,i}])$  represents the maximum value point and  $\min(W[x_{j,i}])$  refers to the minimum value point. After that, compare this point  $x_{j,i}$  with  $\max(W[x_{j,i}])$   $\min(W[x_{j,i}])$  and. The results suggest that if the two points are entirely different, no filtering will be performed on the original value; on the contrary, if the two points get the same value state, the program can be started using a pre-judged calculation method.

If  $f(w, z)$  is the grayscale of point  $(w, z)$  in the image, and  $h(w, z)$  is that of the pixel in  $(w, z)$  (adjacent range), operator  $Z$  is selected and applied to  $e(w, z)$  and  $h(w, z)$ .  $Z=Z(e, h)$  can be obtained. Then, continue with the next step according to the different  $Z$ . Here the way of  $Z$  is shown in formula (42):

$$Z = \sum_{j=0}^S z(e(w, z) - h_j(w, z)) \tag{42}$$

The following formula (43) can be obtained:

$$Z(w) = \begin{cases} 1, & |w| \leq S, \\ 0, & |w| > S \end{cases} \tag{43}$$

After calculation, the value of  $i$  in the formula is shown in Figure 5.

In Figure 5, we can see that the point  $f(w+1, z)$  denotes point 0 adjacent to end  $(w, z)$ , and these values circle the point  $(w, z)$  in turn until the point  $f(w+1, z-1)$  becomes the seventh point.

5	4	3
6	(w, z)	5
7	0	1

Figure 5: i value distribution.

According to the above description, it can be obtained that the threshold value  $T$  represents a constant threshold value. If the visual effect of the image is in a good state, it will not be affected by significant noise pollution, and the distribution change has not become large, then in the calculation, the minimum value of  $T$  value should be taken in the process. Otherwise, it may cause errors because the threshold selection result is not accurate [14-15].

Suppose the final selected result is too high. In that case, the noise will be misunderstood as a helpful signal point during image processing, and most of the noise is retained during processing, reducing filtering efficiency and visual effects. Instead, this can happen if the threshold ( $f$ ) is too low. A helpful signal point would be seen as contamination noise, which would make the image even more blurry, and signal noise would make the visual effect much lower. According to the above calculations, it can be judged that the following conclusions can be further obtained:

- 1) If the gray value in pixels is infinitely equal to or close to it, i.e.,  $y$  is equal to 0, then this point can be regarded as an isolated point of the median noise filter.
- 2) When  $1 \leq y \leq 4$ , that is, the grayscale values of 1 to 4 pixels are equal to or very close, this point is regarded as the peripheral detail node of the part and is not processed.
- 3) When the  $Y$  value is no less than 4, i.e., more significant than the grayscale of four pixels, and is equal to or very close to this value, it can be considered that the point is in a flat area, and the issue has not been processed. The whole operation process is shown in formula (44):

$$y_{ij} = \begin{cases} med(W[x_{ij}]), & Y = 0 \\ x_{ij}, & Y > 0 \end{cases} \tag{44}$$

## 4 Experiment result texture feature analysis noise ratio

If the gray level co-occurrence matrix is defined by a value  $\rho(d, \varphi)$  as the probability that the gray value is set to  $j$  from a point whose gray value is set to  $i$ , the possibility that the gray value at a point leaving a relatively unchanged position is set to  $j$ . In this position,  $d$  is selected to be equal to 1, and the importance of  $\varphi$  is set to  $0^\circ, 45^\circ, 90^\circ,$  and  $135^\circ$ , which are representative angles. In this way, a contrast ratio and a calculation of the entropy value can be established, as shown in Equation (45):

$$\begin{cases} I(d, \varphi) = \sum_k k^2 \left[ \sum_i \sum_j \rho(i, j|d, \varphi) \right] \\ H(d, \varphi) = - \sum_i \sum_j \rho(i, j|d, \varphi) \log \sum_i \sum_j \rho(i, j|d, \varphi) \end{cases} \quad (45)$$

Based on the results of the calculation of  $k=I-J$ , in the process of optimizing the entire video image, the value of contrast is a significant factor used to measure the texture grooves of the image. The deeper the groove texture, the correspondingly greater the contrast between the images. Entropy can measure image information. If there are many fine textures, the value of entropy becomes larger. For the evaluation after video image optimization processing, the evaluation index represents the contrast to the Noise Ratio (CNR) that may be contained in the image, as shown in formula (46):

$$CNR = \frac{2(\mu_t - \mu_b)^2}{\sigma_t^2 + \sigma_b^2} \quad (46)$$

In formula (45), the noise target area value in the image is represented by  $\mu_t$  and  $\mu_b$  represents a mean value described as the background area. The noise target value and the standard deviation of the background are represented by these two data  $\sigma_t, \sigma_b$ , respectively. From the value of the calculation result, if the CNR value obtained above is higher, it means that the image has yet to reach the optimal effect.

### 4.1 Experimental result

The recommended task is executed on CUDA 9.0, Python 3.6, and Tensor Flow 1.9.0, Python software, and is required to be installed alongside Python to carry out the procedure.

In this study, the field-programmable gate array (FPGA), digital signal processing (DSP), and simulation system-based video image optimization [27] are compared with image processing techniques. The simulation results are shown in Figure 6. The proposed

algorithm in this study has the best filtering effect, as shown in Figure 6 because it can determine the most appropriate threshold value. Other algorithms are capable of optimizing, but they are unable to produce outcomes that meet expectations. It could be because calculating the image enhancement standard value is challenging.

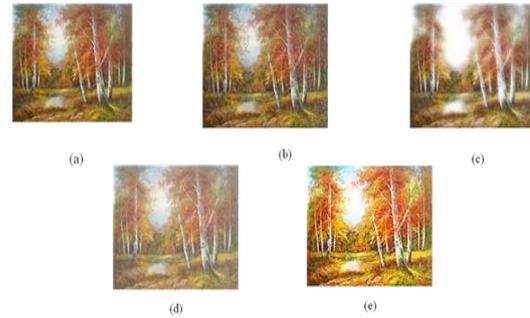


Figure 6 : An analysis of the impact of several techniques on image optimization: (a) source image; (b) simulator; (c) FPGA; (d) DSP; (e) the proposed algorithm

The results of comparing the proposed algorithm's performance are shown in Figure 7. Compared to other techniques, the images processed using the algorithm described here have lower contrast, entropy, and noise. These outcomes demonstrate the effectiveness of this strategy in adjusting contrast.

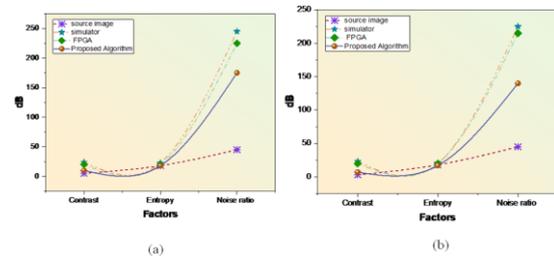


Figure 7: Results of comparing algorithm performance: (a) initial test outcome; (b) second-round test result.

The presented VR-based simulation system for video image processing optimization is compared to DSP-based and FPGA-based techniques in high-brightness images. The simulation results are shown in Figure 8.

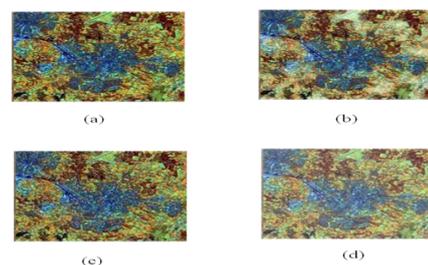


Figure 8: Evaluation of enhanced processing for brighter images: (a) source image; (b) simulator; (c) FPGA; (d) the proposed algorithm

Figure 9 depicts the Comparative algorithm performance outcomes. The proposed algorithm presented in this paper has lower contrast, entropy, and noise values compared to previous approaches. This suggests that it is capable of effectively adjusting contrast parameters, preserving image details, and reducing glare intensity. The proposed algorithm developed in this paper outperforms existing methods in terms of image quality.

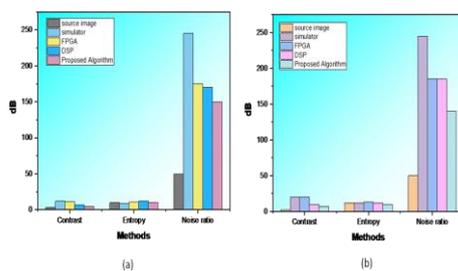


Figure 9: Comparative algorithm performance outcomes: (a) initial test outcome; (b) second-round test result.

The loss metric is used to quantify the prediction error of a model with the goal of reducing the difference between predicted and actual values. Figure 10 depicts the outcome of loss.

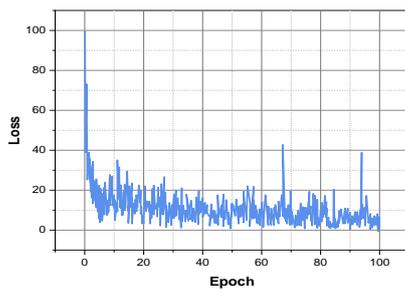


Figure 10: Outcome of loss.

## Discussion

Field-Programmable Gate Array (FPGA) algorithms and Digital Signal Processing (DSP) algorithms face limitations in terms of resources, including logic cells, memory, and interconnects, making scalability and complexity challenging. When implementing a proposed visual restoration algorithm, achieving a balance between computational efficiency and preserving image quality poses challenges.

## 5 Conclusion

The study introduced a VR reconstruction algorithm for increasing the resolution of VR content to provide a clearer and more detailed visual experience. We gathered a COCO dataset for training models to detect and segment objects in VR environments. This can be useful for applications like virtual object manipulation or scene understanding—the proposed method results in an image with the best visual effect. Throughout the future rounds, the discriminant and generator will mutually enhance their learning process, resulting in improved quality of samples and resolution, as well as enhanced VR image improvement, often involving high-resolution content and complex data. Streaming such content in real-time may face challenges related to data transmission speeds and bandwidth limitations, affecting the overall user experience. In future research, improvements may focus on enabling collaborative VR experiences, allowing multiple users to interact seamlessly within a shared virtual environment, enhancing the social aspect of multimedia VR.

## Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of interest

The authors declare no conflicts of interest

## Funding statement

This study did not receive any funding in any form.

## References

- [1] Bannas P, Li Y, Motosugi U, et al. Prior Image Constrained Compressed Sensing Metal Artifact Reduction (PICCS-MAR): 2D and 3D Image Quality Improvement with Hip Prostheses at CT Colonography[J]. *European Radiology*, 2016, 26(7):2039-2046.
- [2] Garg M, Naik T R, Pathak C S, et al. Significant improvement in the electrical characteristics of Schottky barrier diodes on molecularly modified Gallium Nitride surfaces[J]. *Applied Physics Letters*, 2018, 112(16):163-173.
- [3] Vincenzo P, Brodie J P, Terry B, et al. A SLUGGS and Gemini/GMOS combined study of the elliptical galaxy M60: wide-field photometry and kinematics of the globular cluster system[J]. *Monthly Notices of the Royal Astronomical Society*, 2015, 450(2):12-20.
- [4] Vrchota P, Prachar A, Smid M. Improvement of Computational Results of NASA Airliner Model

- by Wing Modal Analysis[J]. *Journal of Aircraft*, 2017, 54(4):1-9.
- [5] Hak, Gu, Kim, et al. VRSA Net: VR Sickness Assessment Considering Exceptional Motion for 360° VR Video[J]. *IEEE Transactions on Image Processing*, 2019, 28(4):1646-1660.
- [6] Iyer V R , Sheedy S P , Gunderson T M , et al. Procedure-Related Pain During Image-Guided Percutaneous Biopsies: A Retrospective Study of Prevalence and Predictive Factors[J]. *American Journal of Roentgenology*, 2019, 213(4):1-7.
- [7] Chen H H , Singh V R , Luo Y . Speckle-based volume holographic microscopy for optically sectioned multi-plane fluorescent imaging[J]. *Optics Express*, 2015, 23(6):7075-7084.
- [8] Ncetan K , Celik I O , Obeid A , et al. VR-Caps: A Virtual Environment for Capsule Endoscopy[J]. *Medical Image Analysis*, 2021, 70(7):101-110.
- [9] Evelyn S , Apolo S F , Ignacio V R , et al. Image-Guided BrachyAblation (IGBA) en hepatocarcinoma. Descripción de la técnica y reporte del primer caso en Chile[J]. *Revista medica de Chile*, 2019, 147(6):808-812.
- [10] Gevaert O , Mitchell L A , Achrol A S , et al. Errata: Glioblastoma multiforme: Exploratory radiogenomic analysis by using quantitative image features (*Radiology* (2014) 273, 1, (168-174) DOI: 10.1148/radiol.14131731)[J]. *Radiology*, 2015, 276(1):56-63.
- [11] Ross A S , Bruno M J , Kozarek R A , et al. Novel single-use duodenoscope compared with 3 models of reusable duodenoscopes for ERCP: a randomized bench-model comparison[J]. *Gastrointestinal Endoscopy*, 2019, 91(2):521-530.
- [12] Aguilera V R , K Apaza Edoya, Pereira B , et al. Clinical study with short implants - Relation among insertion torque, osseointegration and bone loss[J]. *Clinical Oral Implants Research*, 2019, 30(19):457-464.
- [13] Thies J , Zollhfer M , Stamminger M , et al. FaceVR: Real-Time Gaze-Aware Facial Reenactment in Virtual Reality[J]. *ACM Transactions on Graphics*, 2018, 37(2):1-15.
- [14] Kozlov A A , Abdullaev S D , Flid V R , et al. Algorithm and criterion of quality for assessing the packing of polymer microspheres[J]. *Russian Journal of Physical Chemistry A*, 2016, 90(9):1835-1838.
- [15] Zhen, Wang, Qian, et al. The image variations in mastoid segment of facial nerve and sinus tympani in congenital aural atresia by HRCT and 3D VR CT[J]. *International Journal of Pediatric Otorhinolaryngology*, 2015,3(1):59-67
- [16] Chen, W., Liu, X., Qiao, L., Wang, J. and Zhao, Y., 2020. Construction of virtual reality-interactive classroom based on deep learning algorithm. *Wireless Communications and Mobile Computing*, 2020, pp.1-9.
- [17] Wu, M.Y., Ting, P.W., Tang, Y.H., Chou, E.T. and Fu, L.C., 2020. Hand pose estimation in object-interaction based on deep learning for virtual reality applications. *Journal of Visual Communication and Image Representation*, 70, p.102802.
- [18] Zhang, H. and Min, X., 2020. Optimization and simulation of garden image visual effect based on Particle Swarm and wavelet threshold. *IEEE Access*, 8, pp.154390-154403.
- [19] Liu, J. and Chen, Y., 2021, March. Research on scene fusion and interaction method based on virtual reality technology. In *Journal of Physics: Conference Series* (Vol. 1827, No. 1, p. 012010). IOP Publishing.
- [20] Lin, Q., 2020. Application and development of virtual reality technology in artificial intelligence deep learning. In *IOP Conference Series: Materials Science and Engineering* (Vol. 740, No. 1, p. 012151). IOP Publishing.
- [21] T. Fukuda, M. Novak, H. Fujii, Y. Pencreach, and L. C. Fu, "Virtual reality rendering methods for training deep learning, analysing landscapes, and preventing virtual reality sickness," *International Journal of Architectural Computing*, vol. 1, no. 2, Article ID 147807712095754, 2020.
- [22] Zhu, Z. and Du, Y., 2021. Research on interior design optimization based on virtual reality technology. In *Journal of Physics: Conference Series* (Vol. 1746, No. 1, p. 012063). IOP Publishing.
- [23] Geslin, E., Bartheye, O.O., Schmidt, C., Tcha-Tokey, K., Kulsuwan, T., Keziz, S. and Belouin, T., 2020. Bernardo autonomous emotional agents increase perception of VR stimuli. *Network and Communication Technologies*, 5(1), pp.11-25.
- [24] Li, L., 2021. Visual information enhancement method of multimedia human-computer interaction interface based on virtual reality technology. *International Journal of Information and Communication Technology*, 19(2), pp.127-142.
- [25] Yang, S., 2021, February. Intelligent Improvement Measures for the Broadcasting and Hosting Major based on Multimedia and Virtual Reality. In 2021

- Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 447-451). IEEE.
- [26] Mao, W., 2022. Video analysis of intelligent teaching based on machine learning and virtual reality technology. *Neural Computing and Applications*, 34(9), pp.6603-6614.
- [27] Cui, L., Zhang, Z., Wang, J. and Meng, Z., 2022. Film Effect Optimization by Deep Learning and Virtual Reality Technology in New Media Environment. *Computational Intelligence and Neuroscience*, 2022.

