Visual Imaging Method of 3D Virtual Scene Based on VR Technology

Zhao Bing[1](✉) and Zhou Qian[2]

[1] Elites Partners Corporation, Beijing 100027, China
[2] Guizhou Power Grid Corporation, Guiyang 550002, China

Abstract. The traditional 3D virtual scene visualization imaging method has low accuracy and serious center offset in the imaging process. Therefore, a visualization imaging method of 3D virtual scene based on virtual reality technology is designed. In order to reduce the complexity of the imaging scene and reduce the texture interval, texture mapping is used to improve the overall interaction performance of VR technology. In order to improve the reality of the scene, a 3D virtual viewpoint structure is designed by optimizing dynamic collision detection with octree. The visualization of 3D virtual scene is completed by mapping calculation. In order to verify the effectiveness of the design method, an experiment is designed. The results show that the coordinates of the center point obtained by this method are closer to the actual coordinates, indicating that the imaging process is more in line with the actual situation and the imaging accuracy is higher.

Keywords: VR technology · 3D virtual scene · Visualization imaging

1 Introduction

The visual imaging technology of virtual scenes is an important means to realize the interaction between humans and computers and complex data. Its core theory is to realize the real reproduction of virtual scenes or virtual environments and interact with them. The current visual imaging technology of virtual scenes It has been widely used in industrial manufacturing, medical diagnosis, science fiction scene design, virtual combat environment simulation, seismic inversion, oil and gas exploration and many other fields [1]. In the medical field, foreign scholars integrate binocular vision positioning technology and brain tumor resection, stereo vision image acquisition sensor is used to track the operation area, and reconstruct the three-dimensional contour of the local surface through calculation, which greatly reduces the difficulty of operation; while the binocular vision three-dimensional reconstruction technology is used in gastrointestinal appearance inspection to improve the accuracy of operation matching. Chinese scholars have studied the independent litchi picking system. The system uses binocular vision technology, image processing uses fuzzy c-means method to segment the stem and fruit, accurately judge the position of picking point, and then guide the manipulator to pick litchi. Binocular vision is applied to the terminal guidance of autonomous underwater

W. Fu et al. (Eds.): ICMTEL 2021, LNICST 388, pp. 373–384, 2021.
https://doi.org/10.1007/978-3-030-82565-2_31

vehicle (auk). Because the shooting environment is underwater, the nonlinear transformation matrix caused by underwater refraction phenomenon is linearized to approximate, which reduces the range of epipolar variation and achieves high accuracy positioning. Harbin University of technology applies binocular vision to the soccer robot to realize the full autonomous navigation of the soccer robot. The difference is that two cameras are installed: one camera is fixed on the top of the robot, and the other camera that can rotate horizontally is installed on the middle and lower part of the robot. In this way, the detection range is larger, so that the ranging accuracy and processing speed of the robot in the field of vision can match The best way is to adjust the posture in time when encountering obstacles in the process of competition. In the visual imaging technology of virtual scene, the model construction and model rendering of virtual scene are the key links. At present, there are two methods to solve the problem of visual modeling of virtual scene in the world. One is the modeling and rendering of virtual scene based on geometric model based on computer graphics, which is also called virtual scene modeling and rendering based on computer graphics; The other is to use the image acquisition and analysis of 3D environment to build virtual scene, which is based on image [2]. In order to meet the high-precision requirements for visual imaging of virtual scenes such as seismic inversion and oil and gas exploration, many effective imaging methods have been proposed. However, the traditional imaging method has low accuracy and serious imaging shift. Therefore, a visual imaging method of 3D virtual scene based on VR technology is designed.

2 3D Virtual Scene Visualization Imaging Method Based on VR Technology

2.1 Reduce Scene Complexity

In the construction of 3D scene model based on VRML, virtual reality technology is used to describe the roaming virtual environment, which requires a large number of solid surface triangles. Due to the limitation of hardware conditions supporting roaming, the real-time interaction will be stuck or even overloaded, resulting in failure. Therefore, the complexity of the scene must be reduced to improve the overall interactive performance [3]. This article uses texture mapping technology to control the complexity of the scene. Texture refers to the surface details of an object, which is mainly divided into color texture and geometric texture. For example, the non-three-dimensional patterns on indoor furnishings belong to the color texture; the geometric texture includes some three-dimensional folds. The texture mapping is carried out between image space, object space and texture space, as shown in the following Fig. 1.

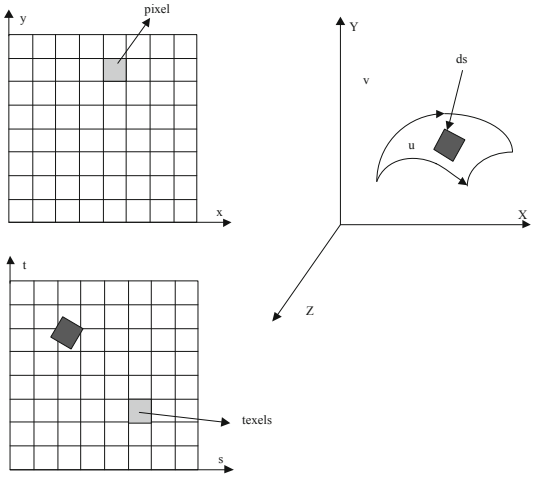

Fig. 1. Mapping process of texture in space

Texture mapping is to map a two-dimensional texture image in a two-dimensional array to a three-dimensional solid surface, and change the light intensity distribution of the surface through a modification process. On the indoor wooden floor model, photos or pictures can be used as texture for mapping, which can save the polygonal surface of the wooden stripes, and the detailed feature description of the wooden floor will not be distorted [4]. The surface area DS corresponding to the pixel on the object surface is mapped to the texture region. The results of partial subtraction are shown in the Table 1 below.

Table 1. 3D virtual scene reduction result table.

Object name	File name	Original file size	Number of original file lines	File size after modification	The number of lines in the modified file
wall	Kz	23.6KB	1065	10.3KB	1065
Interior furnishings	Kz	123.6KB	2640	78KB	1883
Outdoor light	Sxll	724KB	11365	613KB	5610

In the above table, the main reductions made include: removing the spaces caused by shrinkage in the wall, making the supporting parts of the indoor furnishings into prototype nodes, and changing the large amount of repeated parts into prototype nodes in the indoor light. So far, the complexity of the scene has been reduced.

2.1.1 Optimize Dynamic Collision Detection

In VR environment, it is necessary to render the interior space decoration scene. In this process, real-time interactive collision processing will occur to enhance the authenticity of the system for interior decoration scene display [5]. Collision processing mainly includes collision detection, collision determination and collision response. Collision detection is mainly used to determine whether the moving object will collide with other objects when moving an object in three-dimensional space. When it is judged as the relevant information of collision, the operation is determined according to the specific collision situation between two objects [6]. The system in this paper adopts the octree space division technology, and octree is a very effective management method in 3D scene management. The space division process of the octree is shown in the following Fig. 2.

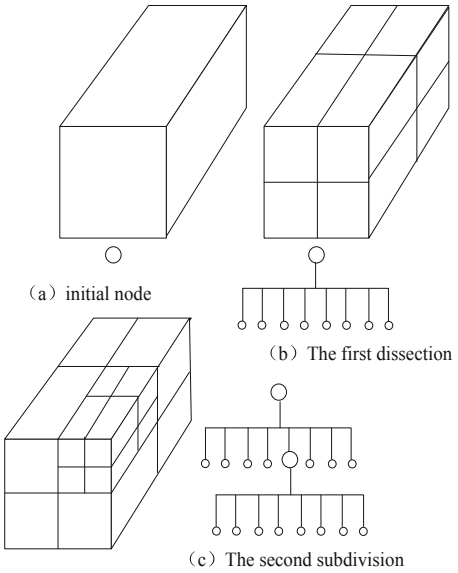

Fig. 2. Schematic diagram of octree subdivision

Assuming that the indoor scene is simplified to a cuboid box, the cuboid is equally divided into 8 small cuboids with all the same shape, and the scene polygons are traversed to determine the positional relationship between the polygons and the small cuboid in the scene, and add it to the node representing the small cuboid In the polygon list, view the number of polygon faces in each non-empty small cuboid. These small cuboids are used as the non leaf nodes of the octree, and are divided recursively. The point in three-dimensional space is represented by non-homogeneous coordinate

(wx, wy, wz, w), which is not unique. To transform a point in space from one coordinate system to another coordinate system, a homogeneous transformation formula must be applied:

$$T = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \tag{1}$$

Until the division reaches a certain depth or the number of polygons contained in the sub-node space is within a given threshold range. If a child node obtained by the division is empty, set it as a leaf node and stop the division [7]. In the process of subdivision, if a polygon exists in two or more small box boxes, the polygon can be divided according to the normal parts in different node cuboids, and the small polygon obtained from the segmentation can be added to the polygon list of corresponding node cuboids to increase the number of polygons in the scene. Aiming at the loose octree data structure in the system, this paper completes the intersection query through the octree scene manager, roughly judges the intersection of objects in the local space where the cuboid of the target node is located, and excludes most objects that will not collide. Create collision geometry and implement the corresponding algorithm. Optimally name all objects that need to collide, and manually add judgment conditions in the background code.

2.2 Synthesis of 3D Virtual Viewpoint Structure

The traditional two-dimensional virtual viewpoint synthesis method only uses the horizontal disparity map to perform virtual synthesis of the viewpoint images adjacent to the reference image in the horizontal direction. In the three-dimensional virtual viewpoint synthesis method mentioned in this chapter, it is necessary to use both horizontal and vertical disparity maps to generate The virtual viewpoint image adjacent to the reference image horizontally, vertically and diagonally [8]. The structure diagram of 3D virtual viewpoint synthesis synthesized in this paper is as follows: (Fig. 3).

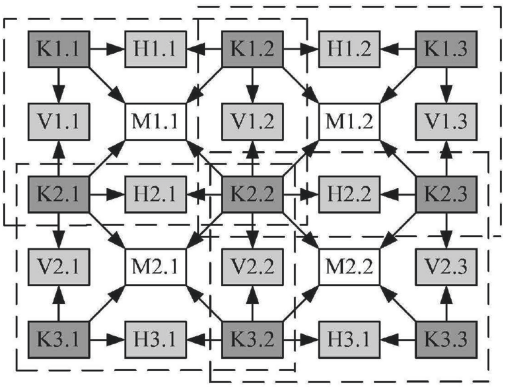

Fig. 3. Structure of 3D virtual viewpoint synthesis

This figure takes a 3 × 3 multi-viewpoint image array to generate a 5 × 5 multi-viewpoint image array as an example. The figure divides the multi-viewpoint images into four categories: K, H, V and M. Among them, the K-type images are actually collected The images of class H and class V are obtained by virtual synthesis of two horizontally and vertically adjacent images of class K respectively. The M class image is located in the center of the 2 × 2 K class image sub-array, and the four K class images are simultaneously Perform virtual synthesis to obtain [9]. The virtual synthesis order of this method is calculated by 2 × 2 class k image subarray from top to bottom and from left to right. The class k image in the above figure is divided into four subarrays. When each subarray is virtual synthesized, the order of each virtual viewpoint image is h, V and m. It can be seen from the above figure that there is an overlap area between adjacent sub-arrays, and this overlap relationship is manifested in the sharing of the disparity map of the overlap area during the virtual synthesis of M-type images. Due to the huge amount of calculation to generate the disparity map, in order to avoid repeated calculations, this chapter will cache the shared disparity map.

In order to generate horizontal and vertical parallax maps, the method of generating the corresponding initial virtual view image from the right view is as follows: define the right view and right parallax map with the size of $C \times R$ pixel as $I_R(i,j)$ and $D_R(i,j)$ respectively, where $i = 0, 1, ..., C-1$ and $j = 0, 1, ..., R-1$ are the location coordinates of the pixel and the parallax value in the right view and the right parallax map, and the coordinate origin is located in the lower left corner of each image. The corresponding generated initial virtual viewpoint image I_{VR} can be expressed as:

$$I_{VR}(i + \alpha \times D_R(i,j), j) = I_R(i,j) \tag{2}$$

In the above formula, α represents the ratio of the distance between the virtual viewpoint to the right viewpoint and the distance between the left and right viewpoints. In the same way, the initial virtual viewpoint image corresponding to the left view can be calculated by using the left view and the left disparity map. The calculation formula is:

$$I_{VL}(i-(1-\alpha) \times D_L(i,j), j) = I_L(i,j) \tag{3}$$

Where I_L and D_L are the left view and the left parallax map respectively, and I_{VL} represents the initial virtual viewpoint image generated by the left view [10]. Since the disparity map cannot accurately extract the edges of objects located in different depth planes, in the generated initial virtual viewpoint image, there are pixels of other depth planes at the edge of the same depth plane, and these pixels are mostly located around the holes in the image. This chapter enlarges the image hole by 1 pixel to eliminate the influence of these confused pixels on image quality. The method to merge the two initial virtual viewpoint images generated separately into one image is:

$$I_{VH} = (1-\alpha) \times I_{VL} + \alpha \times I_{VR} \tag{4}$$

Where: I_{VH} is the merged virtual viewpoint image. For the small holes still existing in the merged image, the hierarchical interpolation method is used in this paper. The specific method is as follows: the unfilled pixels are classified according to the number

of filled pixels around them. The more the number is, the more priority is given. In the processing, the average value of all the filled pixels within the range of 7×7 is taken as the pixel value of the point.

2.3 Visual imaging of 3D virtual scene

In the three-dimensional virtual viewpoint synthesis structure diagram, the virtual viewpoint image in the oblique direction of each subarray is not on the connecting line of any two horizontal or vertical adjacent reference images. Therefore, the initial virtual viewpoint image in oblique direction cannot be generated only by relying on one parallax map of the reference image. To solve this problem, this chapter proposes a method to obtain the initial virtual view image by using the horizontal and vertical parallax map of the reference image at the same time. The method of generating the corresponding initial virtual view image from the upper right view is as follows: firstly, the vertical disparity map of the upper right view and the vertical parallax map of the upper right view are used the specific formula is as follows:

$$I_{VRTM}(I + \alpha_H \times D_{RTH}(i,j), j) = I_{RT}(i,j) \tag{5}$$

Where: I_{RT} represents the upper right view, D_{RTH} represents the horizontal disparity map of the upper right view, I_{VRTM} is the generated middle view, and α_H is the ratio of the horizontal distance between the virtual viewpoint to the upper right viewpoint and the left and right viewpoints. Then, use the obtained intermediate disparity map to map the intermediate view to generate an initial virtual viewpoint image:

$$I_{VRT}(i, j + \alpha_V \times D_{VRTM}(i,j)) = I_{VRTM}(i,j) \tag{6}$$

Where: I_{VRT} is the initial virtual viewpoint image generated by the upper right view, and α_V is the ratio of the vertical distance from the virtual viewpoint to the upper right view and the distance between the upper and lower viewpoints. In the process of merging the generated initial virtual view images, this chapter first merges the initial virtual view images generated by two reference images on the same horizontal line, and then merges the generated two images to generate the final virtual view image. The specific formula is as follows:

$$\begin{aligned} I_{V1} &= (1 - \alpha_H) \times I_{VLT} + \alpha_H \times I_{VRT} \\ I_{V2} &= (1 - \alpha_H) \times I_{VLB} + \alpha_H \times I_{VRB} \\ I_{VM} &= \alpha_V \times I_{V1} + (1 - \alpha_V) \times I_{V2} \end{aligned} \tag{7}$$

In the formula: I_{V1} and I_{V2} are the combined images obtained from the two initial virtual viewpoint images located above and below the sub-array respectively, and I_{VM} is the final synthesized virtual viewpoint image. So far, the research on the visualization and imaging method of 3D virtual scene based on VR technology is completed.

3 Experiment

3.1 Experiment preparation

In the experiment, a series of simulation images are generated by using the generated simulation platform, and the imaging accuracy of the system is detected. The measurement target of the system is the three-dimensional coordinate value of space points. The measurement of typical entity models such as point, line, plane, surface, cube, etc., the entity to be tested is added with different mark types and different levels of noise to study their positioning accuracy respectively. The definitions of points, lines, and polygons in OpenGL are not exactly the same as those in mathematics. On the one hand, the coordinates of geometric figures in OpenGL will cause minor errors due to errors in computer floating point operations; on the other hand, the display of graphics is also subject to The display resolution is limited. Here we use experiments to study the rendering accuracy of geometric primitives in OpenGL. In the experiment, 9×9 sub-images in the sub-image array are extracted as experimental data, as shown in the Fig. 4 below.

Fig. 4. Two dimensional multi view image array actually collected

Take the odd-numbered row and odd-numbered column images in the above array as reference images, and use other position images as unknown viewpoint images for virtual synthesis, and calculate the peak signal-to-noise ratio (PSNR) of the synthesis result and the corresponding actual viewpoint image. The following figure shows the subjective effect of some virtual viewpoint images and corresponding actual viewpoint images: (Fig. 5).

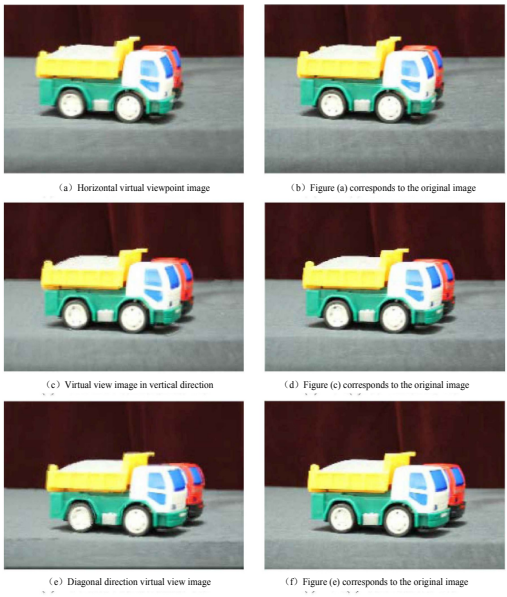

Fig. 5. Subjective effect

In the experiment, the hardware environment is Hisense computer, with Intel (R) Pentium III 996 MHz, 768 m ram, and the card is video quadro2 MXR/ex; the software environment is Windows XP Professional, Visual C++ 6.0, and the display resolution is 1024 × 768 × 32bit. In the process of experiment, we designed some common marks,

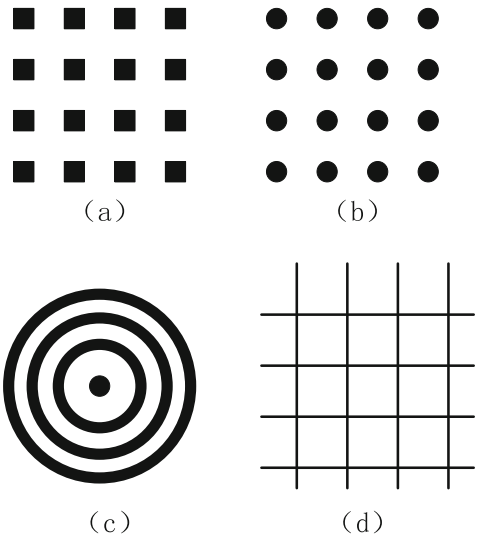

Fig. 6. Simulation mark

such as square point, circle point, stripe and grid. The following is the specific effect of the mark map, as shown in the following Fig. 6.

The accuracy of the generated simulation image is tested. Under different resolutions and imaging functions, the points and lines in any position and direction are verified. At the same time, Gaussian noise and impulse noise of 60dB are added to the image to study the influence of noise on target point positioning.

In this paper, the PSNR of the horizontal virtual view image relative to the actual view image is shown in the Table 2 below.

Table 2. PSNR (dB) of the virtual viewpoint image in the horizontal direction.

Specific location	Column 2	Column 4	Column 6	Column 8
Line 1	29.3	28.0	29.1	29.4
Line 3	30.4	28.3	30.1	30.4
Line 5	30.4	28.9	30.4	30.7
Line 7	29.1	28.7	30.2	29.4
Line 9	28.3	27.3	30.7	29.3

The PSNR of the virtual viewpoint image in the vertical direction relative to the actual viewpoint image is shown in the following Table 3.

Table 3. PSNR (DB) of virtual viewpoint image in vertical direction.

Specific location	Column 2	Column 4	Column 6	Column 8
Line 1	28.5	27.9	26.6	27.5
Line 3	28.5	29.1	26.4	28.3
Line 5	27.3	28.5	28.4	29.8
Line 7	29.5	28.4	26.9	28.4
Line 9	29.5	28.5	19.3	29.0

The PSNR of the virtual viewpoint image in the diagonal direction relative to the actual viewpoint image is shown in the following Table 4.

It can be seen from the above three tables that the horizontal and vertical virtual view images have higher PSNR than the oblique virtual view images, but the subjective quality of the oblique virtual view images is not significantly worse than that of the horizontal and vertical directions. The main reason for this phenomenon is that the images used in this chapter are real scene images, and the distance between adjacent shooting viewpoints Distance is not exactly equal, Therefore, there is a slight error between the virtual synthesized viewpoint position and the corresponding actual acquisition viewpoint position. This error only affects the PSNR calculation result and will not cause

Table 4. PSNR of the virtual viewpoint image in the diagonal direction (dB).

Specific location	Column 2	Column 4	Column 6	Column 8
Line 2	23.4	22.8	23.3	24.5
Line 4	23.5	24.1	24.2	24.0
Line 6	22.1	23.0	22.9	23.4
Line 8	22.0	22.8	23.5	23.4

the deterioration of the display effect, and the error has greater influence on the virtual viewpoint image in the diagonal direction. The impact on the horizontal and vertical directions.

3.1.1 Analysis of experimental results

Under the above experimental conditions, the imaging centers of the two imaging methods are determined, and the results obtained are shown in the following Table 5.

Table 5. Comparison of center point coordinates.

Number of imaging experiments	The coordinates of the center point in this method	Traditional method center point coordinates
1	(150,150)	(137,126)
2	(150,150)	(141,150)
3	(149,149)	(159,162)
4	(200,100)	(157,132)
5	(151,159)	(134,159)
6	(149,153)	(132,161)
7	(150,150)	(173,159)

In this paper, the actual coordinates of the imaging center point are set as (150, 150). According to the coordinate results of the imaging center point in the table above, the coordinates of the center point obtained by the 3D virtual scene visualization imaging method based on VR technology are closer to the actual coordinates, which shows that the method in this paper has higher accuracy in the imaging process and is more in line with the actual situation.

4 Concluding remarks

With the in-depth development of industrial manufacturing, medical diagnosis, science fiction scene design, virtual combat environment simulation, seismic inversion, oil and

gas exploration and many other fields, the visualization imaging of virtual scene has become a global hot topic. Because the current imaging methods can not meet the application requirements of various fields, a virtual scene visualization imaging method based on common focus is proposed The visual imaging results of the pseudo complex surface interface model verify the effectiveness and feasibility of the proposed method, and provide effective guidance for the further development of oil and gas exploration, seismic simulation and other fields.

References

1. Hu, C., Wang, L., Li, Z., et al.: Inverse synthetic aperture radar imaging using a fully convolutional neural Network. IEEE Geosci. Remote Sens. Lett. **17**(7), 1203–1207 (2020)
2. Martel, J.N.P., Müller, L.K., Carey, S.J., Dudek, P., Wetzstein, G.: Neural sensors: learning pixel exposures for HDR imaging and video compressive sensing with programmable sensors. IEEE Trans. Pattern. Anal. Mach. Intell. **13**(02), 1–14 (2020)
3. Xiaokang, Z., Qiuyu, Z., Peng, R.: Research on panoramic imaging system based on VR miniature satellite. Electron. Measur. Tech. **41**(05), 33–37 (2018)
4. Yang, S., Wang, Y., Shi, B., et al.: Construction of anatomical and pathological simulation laboratory based on VR and 3D prisnting technology. J. Zhengzhou. Railway. Vocational Tech. Coll. **30**(01), 31–33 (2018)
5. Guo, W.: The application of 3D laser virtual reality technology in the reconstruction of ancient buildings. Laser. J. **39**(12), 102–105 (2018)
6. Chen, Z., Liu, Q., Yin, W., et al.: Visualization of petroleum hydrocarbon content in latosol based on hyperspectral imaging technology. Spectrosc. Spectral. Anal. **38**(09), 2916–2922 (2018)
7. Liu, S., Lu, M., Li, H., et al.: Prediction of gene expression patterns with generalized linear regression model. Front. Genet. **10**, 120 (2019)
8. Fu, W., Liu, S., Srivastava, G.: Optimization of big data scheduling in social networks. Entropy **21**(9), 902 (2019)
9. Xiuju, X.: 3D visualization technology of collapse column advanced detection by mine transient electromagnetic method. China. Coal. **44**(10), 60–64 (2018)
10. Liu, S., Bai, W., Zeng, N., et al.: A fast fractal based compression for MRI images. IEEE. Access. **7**, 62412–62420 (2019)
11. Bagi, R., Mohanty, S., Dutta, T., et al.: Leveraging smart devices for scene text preserved image stylization: a deep gaming approach. IEEE Multimedia **12**(05), 1–12 (2020)
12. Xu, X., Ma, Y., Sun, W.: Towards real scene super-resolution with raw images. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2020)